ENAC

Institut mermoz

In cooperation with

AIRBUS

Radio Navigation

New programme

EASA
European Aviation Safety Agency

ATPL Training

062

062

Radio Navigation

Version 2

Intentionally left blank

CONTENTS

062 RADIO NAVIGATION

01

BASIC RADIO PROPAGATION THEORY

01 BASIC PRINCIPLES

1.1 - Electromagnetic waves

In theory, radio waves travel in space (likened to a vacuum) at the speed of light.

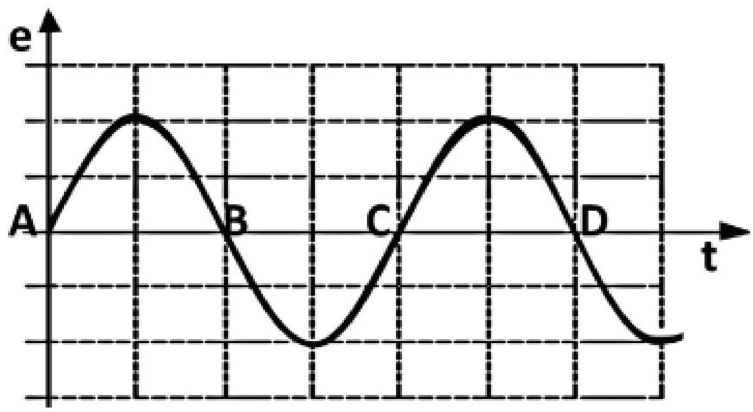
This speed is commonly denoted C and is equal to 300 000 km/s, which is the upper limit for the speed of light in a vacuum (C is also called Einstein's constant).

$$C = 300\,000 \text{ km/s}$$

Strictly speaking, the movement of the wave in a complex environment (the atmosphere or the ionosphere) happens at a speed lower than C called phase velocity (Vc), which depends on the relative permittivity and permeability of the environment.

Just for information $Vc = 1/ (C^2 . \varepsilon r . \mu r)^{½}$

We call a cycle, a complete series of values of a periodical process.



AC represents a cycle
BD represents a cycle

When a full cycle occurs over a period of 1 second, its frequency is 1 Hz.

$$1 \text{ Hz} = 1 \text{ cycle/s}$$

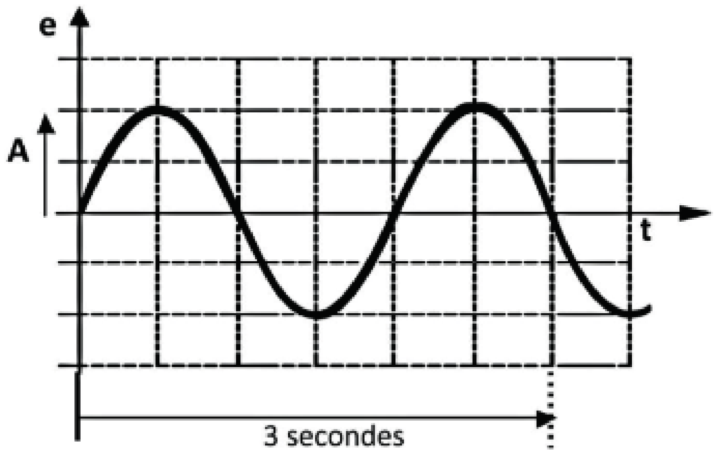
1.2 - Frequency, Wavelength, Amplitude, Phase Angle

1.2.1 - Frequency

Frequency F is the number of cycles (or periods) occurring during a period of 1 second. It is expressed in Hertz (Hz).

The period T is the time in seconds that it takes for one cycle to repeat itself identically and in the same direction.

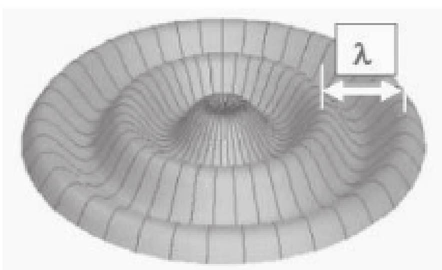
The period is the reciprocal of the frequency: $T = 1/F$.



We have here one cycle and a half in a 3 second period of time, therefore one cycle is equal to 2 seconds.
We can say T = 2s.
The frequency of this signal is 0,5 Hz.
Note on the diagram the amplitude A.
Amplitude is the maximum deflection in an oscillation or a wave.

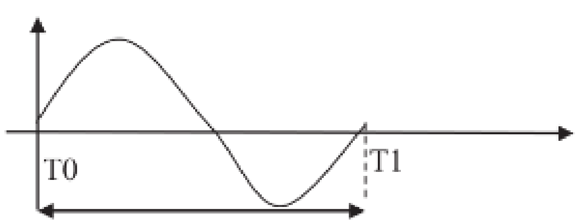
1.2.2 - Wavelength

If you throw a stone on a still lake, you will observe from the point of impact of the stone a series of concentric waves that spread towards the shore at a speed V (that depends on the propagation environment).



The wavelength λ is the distance between two ridges of a wave.

The same is true for the propagation of an electromagnetic wave.

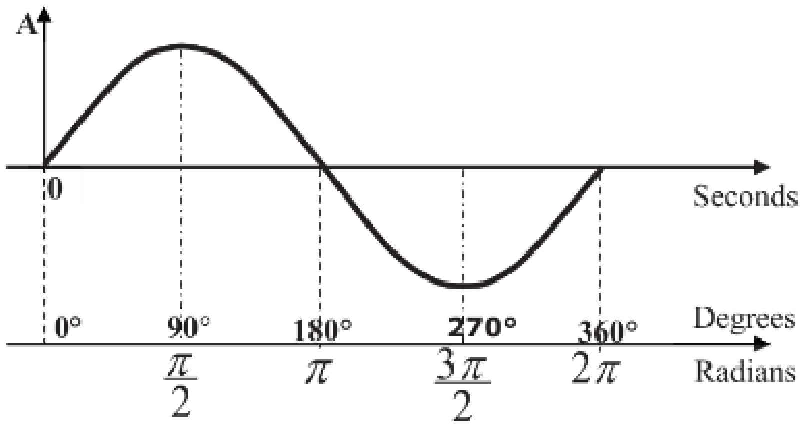


We call wavelength λ, the physical distance travelled by a wave from the start of the emission at To until the end of the cycle at T1.
This distance is calculated in physics with the formula:
Distance travelled = velocity x time, so for a wave λ = C.T and since T = 1/F,
a wavelength is calculated with the formula λ = C/F .

This formula applies if the wave travels at the speed of light C but in reality the wave travels with a phase velocity Vp that depends on the refraction index, thus λ = C/F becomes λ = Vp/F (the explanation of refraction will come later) but we will retain λ = C/F for exam purposes.

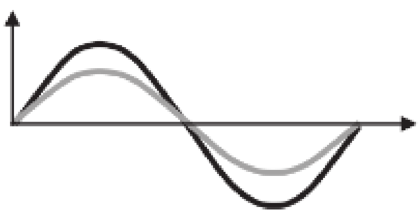
1.2.3 – Phase angle

Instead of presenting the amplitude A of a wave as a function of time, it is more convenient to present it as a function of the angle α reached by the alternative signal at a certain time t.
This angle will be expressed in degrees or in radians.



We call the phase of a signal, the angle reached at a time t.

This concept is useful to visualize if two signals vary in time in the same way.
If two signals reach their maximum amplitude simultaneously, we will say that they are in phase.
If two signals do not reach their maximum amplitude simultaneously, we will say that there is a phase difference or a shift.



2 signals in phase

2 signals with a phase difference: the phase difference is 90 degrees; The black signal is phase advanced on the grey signal because it reaches its maximum amplitude first.

We express the phase difference by an angle noted Φ.
In the example above the phase difference, or shift, Φ is 90 degrees.

1.3 - Electromagnetic frequency bands

Demonstrations of electromagnetic radiation are numerous, and apply to varied fields such as cosmic rays, optics, infrared radiation, and radio waves.

The following table shows the radio frequency spectrum divided into frequency bands.

Very Low Frequency	VLF	3-30	kHz
Low Frequency	LF	30-300	kHz
Medium Frequency	MF	300-3000	kHz
High Frequency	HF	3-30	MHz
Very High Frequency	VHF	30-300	MHz
Ultra High Frequency	UHF	300-3000	MHz
Super High Frequency	SHF	3-30	GHz
Extremely High Frequency	EHF	30-300	GHz

Each band possesses its own propagation properties.

Because of their wide range, UHF and SHF bands are divided into sub-bands.

	SHF				UHF		
BAND	Ka	Ku	X	C	S	L	P
WAVELENGTH in cm	1,00	2,00	3,10	5,60	9,60	23,00	68,00
FREQUENCY (Ghz)	35,00	14,00	9,60	5,30	3,00	1,30	0,44

Sub-bands X and C are used for weather radars; sub-band L is used for GPS or ATC radar.

1.4 - Carrier and Modulation

1.4.1 – Carrier Wave

A high frequency wave (from VLF to EHF) does not possess any information but it travels well in space.
It is a succession of repetitive cycles that do not vary, but to carry information there needs to be a variation of the medium.

Example:

A lamp that is permanently turned on does not provide any information (other than the fact that it is on).
It emits a continuous wave which is the light (the medium).
This light has a frequency which gives the color of the light.
My voice will not carry far enough to transmit a vocal information over a long distance.
I will use light and make it vary (or modulate) following an established code.
So, by decoding the variations of a medium it is possible to retrieve information.

The same applies to a high frequency wave.
It will be used as a carrier to transport information and will be called a carrier wave.
The high frequency signal is called CARRIER (it does not contain information).
The low frequency signal is called MODULATING SIGNAL (it is the information).

1.4.2 - Modulation

It is the technical term that designates the process of impressing and transporting information by radio waves.

1.5 – Kinds of modulation

A wave can vary in amplitude, in frequency, and in phase, or in a combination of these properties.

A sine curve, also called sinusoid, is defined by three parameters: its amplitude, its frequency and its phase.
It has the following equation:

$$A (\sin \omega t + \varphi) = A (\sin 2 \pi F t + \varphi)$$

Each of the three parameters of the carrier wave are proportional to the signal that needs to be transmitted.
This provides three basic types of modulation:

- Amplitude Modulation ;
- Frequency Modulation ;
- Phase Modulation.

There is also a Pulse modulation that will be discussed in the radar course.

Remember that modulations can be of three kinds: A, F, P.

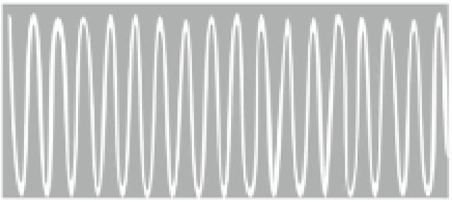
1.5.1 - Classification of modulated signals

The different kinds of modulation that are used have been codified by the ITU (International Telecommunication Union) with a 3 symbol code.
The signal modulation codes used for aeronautical applications in this course are:

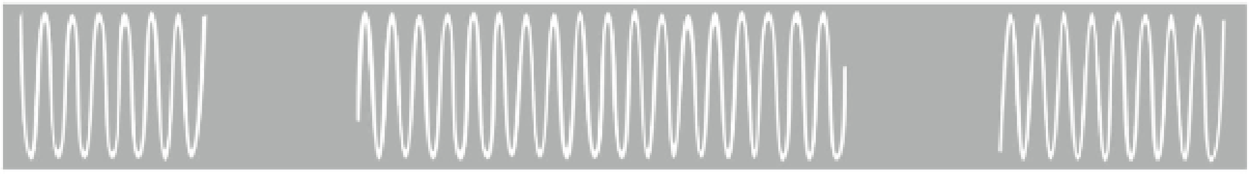
- N0N (N zero N)
- A1A
- A2A

A signal A3E corresponds to a modulation used in VHF COM.

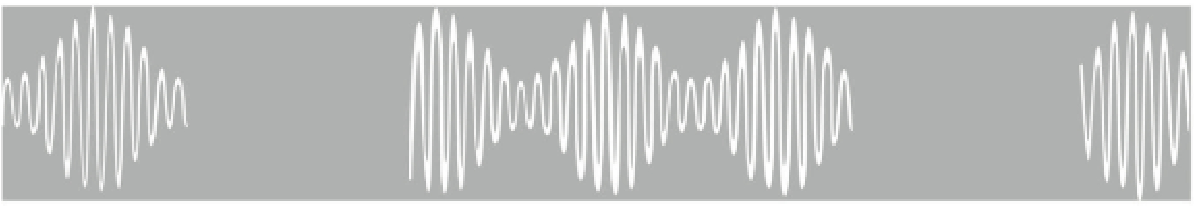
N0N (pure non modulated carrier);



A1A (pure non modulated carrier, cut out to follow the rhythm of the Morse code, it is a telegraphic information but it is not directly audible – we will see that in the ADF course);

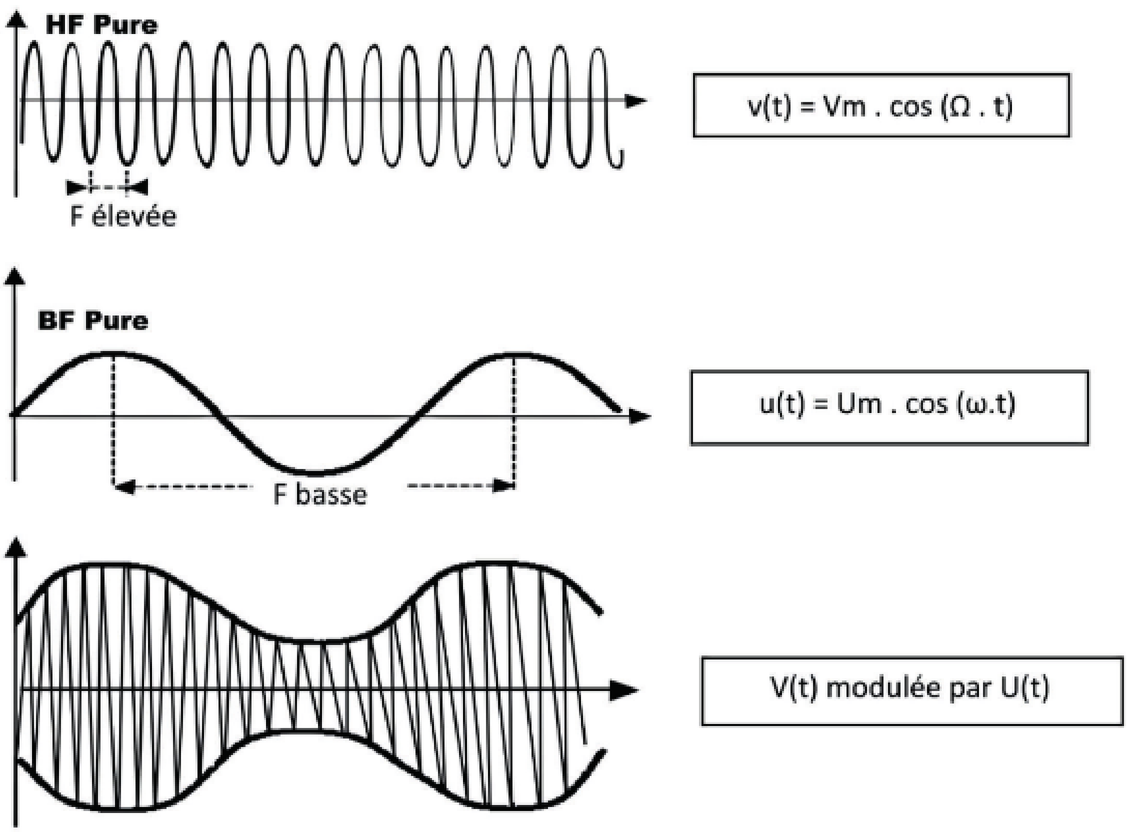


A2A (pure carrier modulated by an audible signal, cut out to follow the rhythm of the Morse code, it is a telegraphic information).

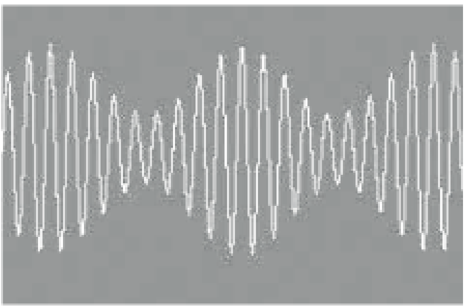


1.5.2 - Amplitude Modulation

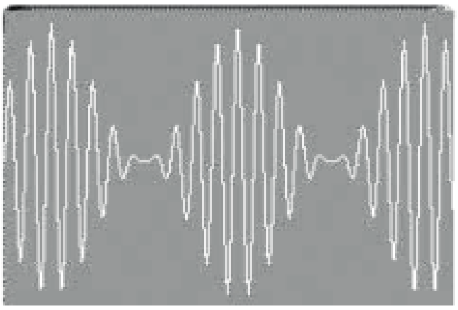
In amplitude modulation, the amplitude of the HF signal, called carrier, is modulated at the rhythm of the information u(t), called modulating signal.



The modulating signal can (depending on the electronic modulation circuit creating the multiplication of the two signals) make the carrier wave vary in higher or lower proportions.
This ratio between the maximum value and the minimum value of the modulated carrier is called modulation index and is expressed as a percentage.



50% Modulation

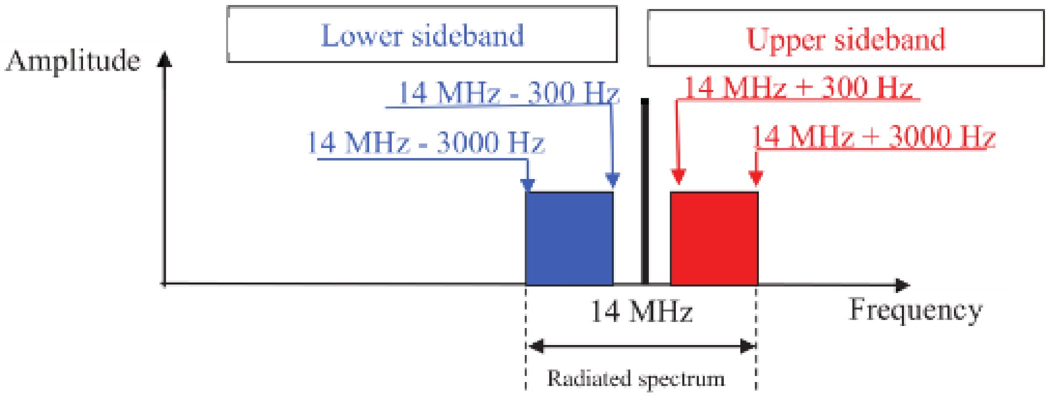


100% Modulation

In amplitude modulation, the information is contained in the varying amplitude of the carrier wave.

Spectrum of an amplitude modulated carrier:

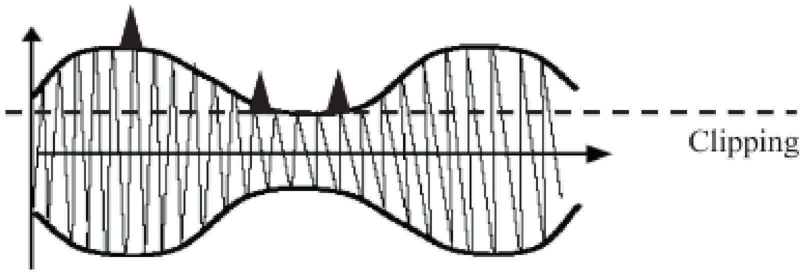
The transmitter of a high frequency signal (the carrier) modulated in amplitude by an information (a voice for example) radiates the carrier signal in space, the carrier signal plus the voice and the carrier signal minus the voice.
The figure below shows the modulation of a 14 MHz frequency by a voice ranging from 300 Hz to 3000 Hz.



We see that the information (the voice) is radiated twice.
In order to lower the spectral congestion and allow more stations to operate in a designated frequency bandwidth, it is possible to transmit only the upper or the lower sideband.
This is known as Single SideBand transmission (SSB).

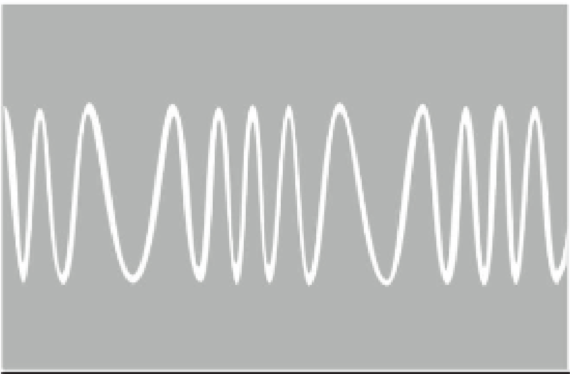
HF communication and VOLMET signals use SSB.

One of the disadvantages of amplitude modulation is that it is sensitive to parasites (industrial and atmospheric).
If we clip a signal at the level of the parasites in order to suppress them, we lose information.



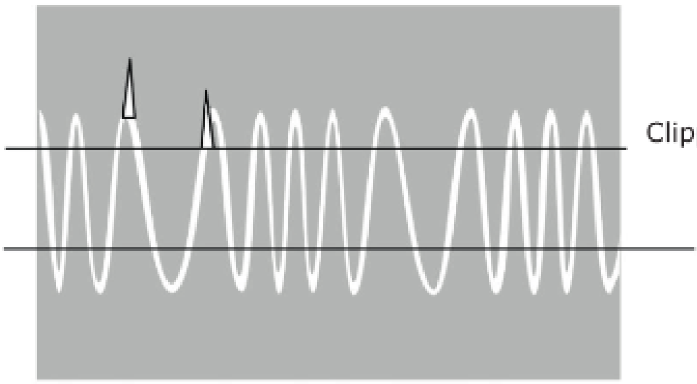
1.5.3 – Frequency Modulation

In this type of modulation, the amplitude A remains constant. It is the frequency of the carrier signal that is made to vary (in greater or lesser proportions), at the rhythm of the modulating signal.



carrier F frequency modulated by f

One of the advantages of frequency modulation is that it is insensitive to statics.

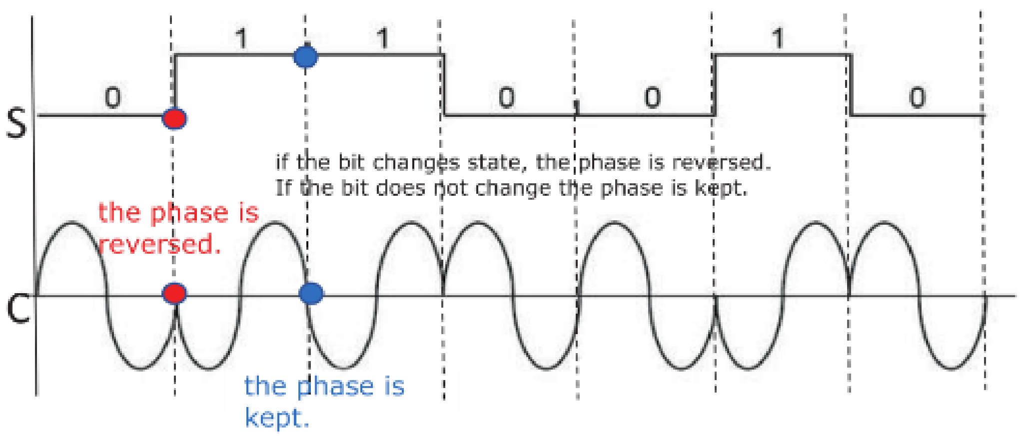


If statics develop on the transmitted signal, it can be clipped without losing information.
The information being contained « in the elasticity » of the period, it will remain even after the signal is clipped.

Low altitude radio altimeter uses this kind of modulation (see 022.12.08.00).

1.5.4 – Phase Modulation

In this kind of modulation, the signal S is encoded as variations of phase of the carrier C.
Amplitude and frequency remain the same.



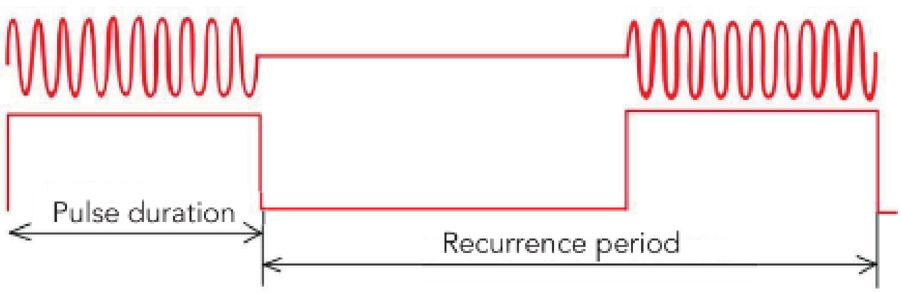
In the example above, the phase varies abruptly by 180° with each change in the state of the modulating signal.

This kind of modulation is used to transmit the GPS signal.
Phase modulation is suited for data transmission.

The example above shows a phase modulation with two states that allow only a low output of information (2 states only = binary information, and the name BPSK for Binary Phase Shift Keying) for this phase modulation.
There are numerous other kinds of phase modulations, but also phase modulations combined with amplitude modulation, which allow the simultaneous transmission of more information and a higher output.

1.6 - Pulse Characteristics

A pulse is the transmission over a short period of time of an electromagnetic signal.
The transmission below is an example of pulse transmission.
Many systems use pulse modulation (radar, DME, transponder).



The transmission is characterized by the amplitude (peak value), the duration and the recurrence period of the pulse.
A large pulse contains more power that a narrow pulse.
The power is equal to the peak value of the pulse multiplied by its width.
The continuous power is the average power on a full transmission cycle (the duration of the pulse plus the time before the new transmission of a pulse).
It therefore depends on the recurring frequency of the pulse and on its peak value.

02 ANTENNAS

2.1 - Characteristics

2.1.1 - Antenna – Definition

We have determined that an alternative electrical signal can contain information.
Now we want this information to be transmitted over a distance, without using a physical medium.
The electrical signal must be transformed into an electromagnetic signal that will travel in space.
This is the role of the antenna.

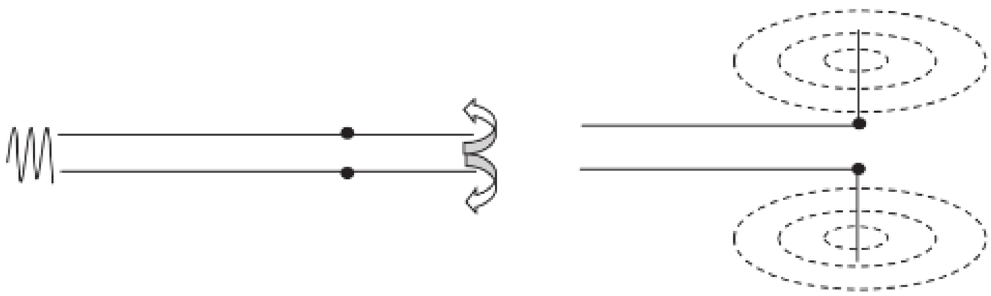
Antennas transform energy:

- upon transmission, they transform electrical energy into electromagnetic energy;
- upon reception, they transform electromagnetic energy into electrical energy.

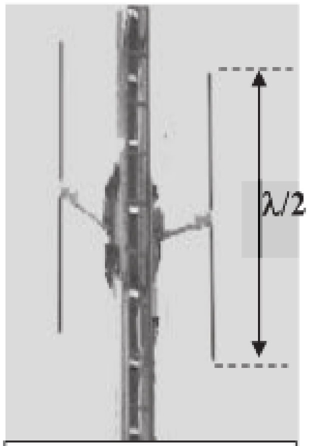
Antennas are reciprocal, they have symmetrical properties which means that the same effects occur during the transmission and the reception of radio waves.
For this conversion to happen, the frequency of the electrical signal must be sufficiently high (there is no exploitable radiation below 20 000 Hz).
Indeed, radiation comes from the changes in acceleration of the electrons in the medium.
For an antenna to radiate part of the energy received, its length L must be must be in a correct ratio (λ, $\lambda/2$, $\lambda/4$...) with the wavelength of the signal to be transmitted.

2.1.2 - Dipole

The simplest type of antenna is a dipole, working in half of a wave.



It is an open transmission wire with ends that are bended at the appropriate length (half of a wavelength).



Set of 2 dipoles

2.1.3 - Transmitting antenna

This antenna is fed alternative current generated by a transmitter (for frequencies of 20 000 Hz and above), and radiates in space part of the energy that it is fed.

2.1.4 – Receiving antenna

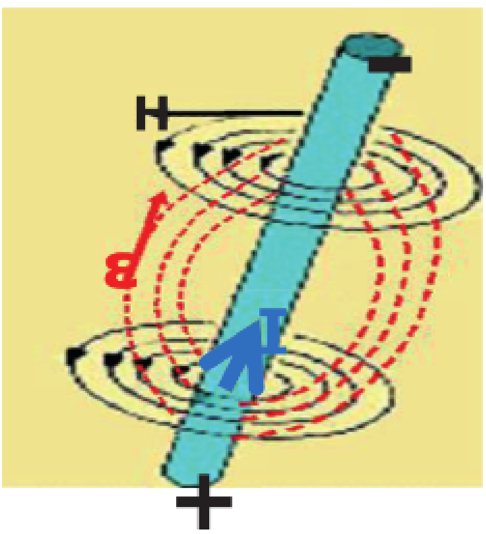
This antenna, immersed in a wave that is generated remotely, is induced by the wave and receives the electrical current representative of the inductive wave.
This electrical signal is then treated by a receiver.
It is the case of two parallel conductors, one is powered, the other is not.
The conductor that is powered radiates and induces the other conductor.

2.1.5 - Electromagnetic wave

Alternative current going through a circuit produces an alternative magnetic field, noted H in the medium surrounding the circuit.
The differences in electrical potential between the different points of the circuit produce in the surrounding space an alternative electrical field noted ε.

Those two magnetic fields are inseparable from each other.

Their production in space is made easier when the frequency is high.
The direction of these fields follows the rules of electromagnetism and electrostatic.

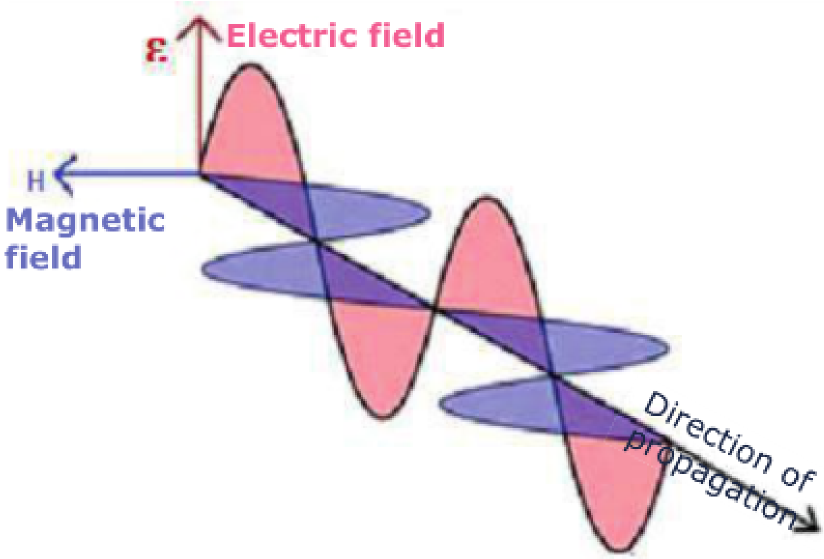


At the instant t, a current I flow from + to – creating an electric field ε parallel to the wire. At the same instant a magnetic field develops around the wire.

The value of the field at a point in space is designated by a vector (H or ε) tangent to the lines of force.
H et ε are perpendicular:
H is expressed in A/m;
ε is expressed in V/m.
The electrical field is parallel to the conductor.
The magnetic field is perpendicular to the conductor.

The combination of those two inseparable fields produces in space an electromagnetic wave that travels at the speed of light.

The production of the electromagnetic wave is a result of the movements of the electrons in the medium.
Despite the fact that nothing moves, the variations of ε and H are felt remotely at the speed of light.



The wave travelling in space (here represented in a given direction) is the composition at any time of two vectors H and ε, orthogonal with each other and with the direction of propagation and travelling in phase.

2.2 - Polarisation

2.2.1 - Definition

The polarisation of an electromagnetic wave describes the orientation of the plane of oscillation of the electrical component of the wave with regards to its direction of propagation.
On the above schematic: the electromagnetic wave has a vertical polarisation.

Since the electrical field is parallel to the conductor radiating the wave, a vertical dipole transmits a wave with a vertical polarization, and a horizontal dipole transmits a wave with horizontal polarisation.

2.2.2 - Types of polarisation

Polarisation can be linear, meaning that the vectors keep a fixed direction (vertical or horizontal) while travelling, or circular in which case the vectors rotate around the propagation axis while staying orthogonal.
Imagine the bullet of a rifled barrel gun.

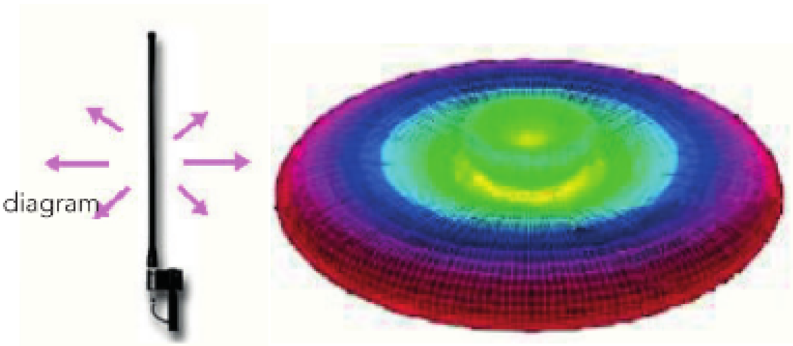
2.3 - Types of antennas

Antennas radiate an electromagnetic wave in space but depending on their technology they do not necessarily illuminate the same areas in space.
A dipole antenna or the whip antenna below, have an omnidirectional radiation.

Others have directional radiation.
This is obtained with a combination of additional elements put close to the radiating antenna (reflector).
These elements reduce the waves in some areas of space and increase them in other areas.
We say that these antennas have a gain because in the chosen direction the transmitted signal is equal to the one that we would obtain with a more powerful transmitter.

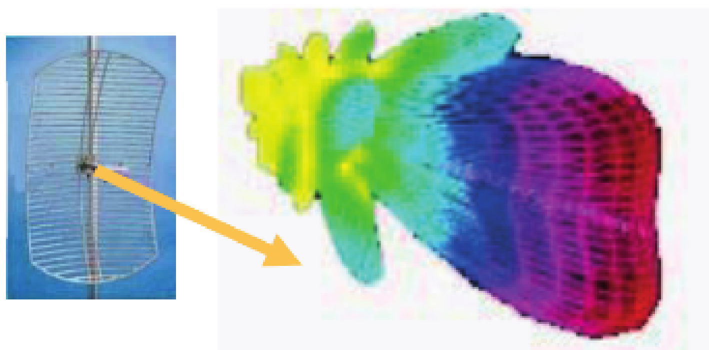
Vertical whip antenna

Omnidirectional radiation diagram



Antenna with a reflector

Directional radiation diagram



To be able to materialize the areas of space that are illuminated by an antenna, we draw radiation diagrams horizontally and vertically.

The field that is remotely radiated by an antenna is expressed in Volts/meter and is measured with an instrument called field meter (it is often expressed in µV/m because of the weakness of the field).
Moving around the antenna we can trace the points with the same field.
Joining these points gives us an idea of the radiation of the antenna.

A directional antenna that concentrates more energy in a narrow beam offers a greater range in a chosen direction.
In addition, by radiating in a chosen direction it will create less disturbance to neighboring transmissions.

Upon reception, the use of such an antenna will pick-up the weaker signals and limit disturbance form nearby transmissions.

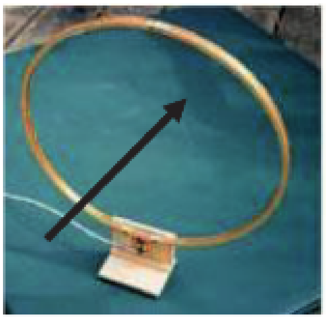
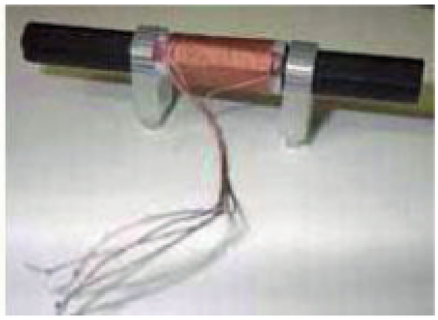
There are many types of antennas that offer more or less directionality:

- Loop antenna used for the ADF;
- Array antennas used for the Localizer and the ILS for example;

- Parabolic antenna used for radar;
- Slotted planar array used for weather radar;
- Patch antenna used for the GPS receiver;
- Helical antenna used for the GPS satellite.

A directional antenna will compensate for the attenuation that a signal undergoes on its path.

2.3.1 - Loop antenna

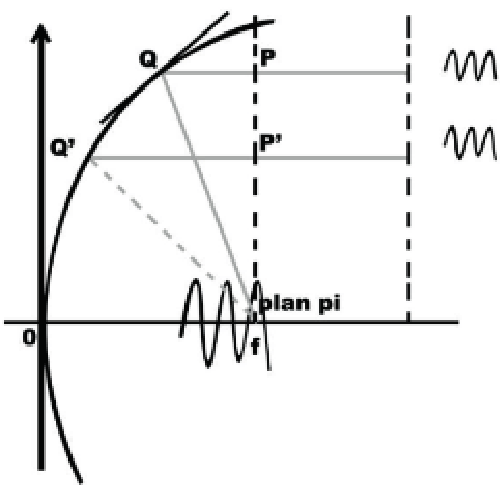



It consists of a single coil (flat) or a variable number of coils, winded into a ferrite frame to reinforce permittivity, depending on the frequency being used.
This antenna is sensitive to the magnetic component of the electromagnetic signal.
Its diagram is a figure of eight (see NDB/ADF course).

These antennas are used in the LF and MF bands, and even in the lower HF frequencies.

2.3.2 - Parabolic antenna

One of the geometric properties of the parabola is that it is a set of points, where for all points the sums of the distances (fQ+QP, fQ'+fP') are equal.
Thus, any signal transmitted from point f (the focus of the parabola) is reflected on the parabola and comes out as P, P', etc... having travelled a strictly identical distance.
We have on a plane Pi a large quantity of the same signals, that are in phase and that add up to produce a stronger signal.
At reception, a signal arriving at plane Pi will produce by reflection a stronger resulting signal at the focal point of the parabola.



2.3.2 - Flat slot array antenna

Theory shows that we can obtain the equivalent of an antenna by cutting a slot on an indefinite plane of a perfect conductor.
The slots can be in straight lines or in loops.
Each edge of the slot can be assimilated to a conductor.
Electrical currents run in opposite direction on the conductors and produce a magnetic field perpendicular to the plane of the slot, and that add up to each other.
The slot has a dimension of λ/2 or λ/4 and enters into resonance.
A high number of slots produces an array and allows a strong directional effect (slotted planar array of modern radars, MLS).

Slotted planar array of a radar
(See antenna A340 in radar chapter)

A flat slot array antenna displays an increased directivity compared to a parabola antenna and thus allows the same range with lesser power.

It is possible with this technology to drive the phase of the signal going through each slot and obtain a mobile beam without moving the antenna (in a limited sector).

Parabolic and flat slot array antennas are used for UHF et SHF (frequencies used by radars).

2.3.3 - Helical Antenna

This type of antenna used with UHF and SHF displays a directivity according to length and diameter.

It radiates a wave with a circular polarisation according of the direction of its spiral.
In a right-hand side or left-hand side circle.

This antenna is installed on the transmitters of the GPS satellites.

It is worth noting that an antenna is connected to a transmitter (receiver) with a supply line responsible for transferring electrical energy to the antenna, which in turn will convert the energy into an electromagnetic field.
Up to 3 000 MHz this link is done with a coaxial cable (like a television cable but of a higher quality), and above by a tube called waveguide (as with radars in particular).

2.3.4 – Masking effect

When an object is on the trajectory of the wave there is a mask effect attenuating of the received wave. (Sometimes there can be total diminishment.)
The masking effect varies according to the frequency being used, the distance between the antenna and the obstacle, and the size and height of the antennas.
Example: a very small size antenna like the DME antenna (see chapter on Distance Measuring Equipment) using an ultra high frequency can easily be momentarily hidden by the wing during a turn.

2.3.5 – Antenna location

The location of the antennas on an aeroplane is crucial to avoid as much as possible the masking effects and the multiple wave trajectories created by reflection.
Of course, antennas must be unobstructed but we must also take into consideration the environment and the direction of the radio link.
A satellite antenna (COM or GNSS) will inevitably be on the roof.
An ILS antenna or a weather radar will be located in the nose of the aeroplane.
Antennas are sensitive to the environment, for example an ILS antenna located in the nose of a large aircraft could be disturbed by the nose gear being retracted or extended.
This would require adjustments.
One solution is to practice antenna diversity which consists of installing two antennas at different places and letting the system choose the best antenna depending on the circumstances.

Another important point for VHF COM is to have an operating antenna in the case of a belly landing or of a ditching.
That is why the VHF COM antenna that stays powered in emergencies is always located on the roof of the aeroplane.

03 WAVE PROPAGATION

3.1 – Structure of the ionosphere

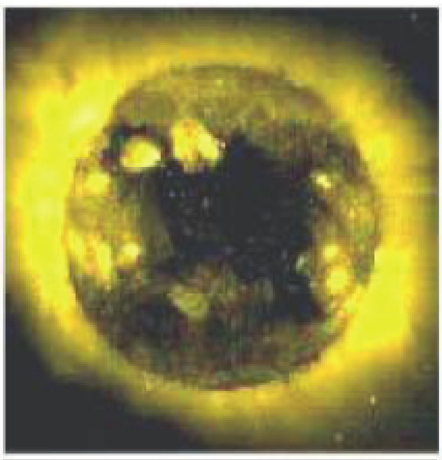
The ionosphere is a region of the earth's upper atmosphere (from 60 to 400 km above the earth surface) that contains free electrons.
The ionisation causes losses due to absorption, and refractions or reflections.
The two major causes of the ionisation of the upper atmosphere are:

➢ Photo-ionisation:

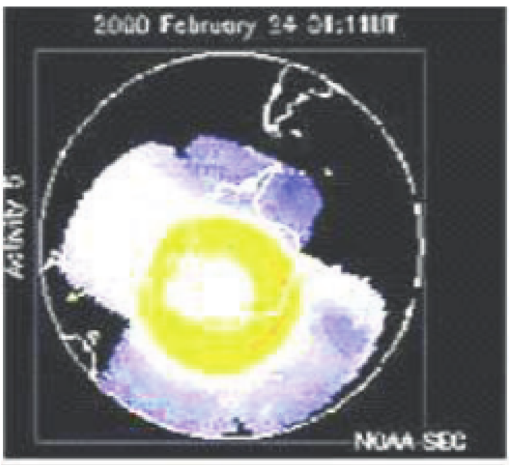
The energy of the photons of solar light is sufficient to pull off the outer electrons from some gaseous atoms in the upper atmosphere.
This is a variable phenomenon with day/night, summer/winter cycles.

➢ Ionisation by solar particles:

Solar flares project into space bursts of electrified material particles (ions or electrons).
These particles, trapped by the magnetic field of the earth, move around the globe (aurora borealis close to the poles because of the concentration of forces of the earth's magnetic and thus a very strong ionisation of space which illuminates like a neon light).
Solar activity follows an 11-year cycle, giving very different ionisation rates over time.
This phenomenon of varying intensity is permanent.



Solar emission



Magnetosphere

The properties of the ionosphere will vary with:

- the time of day
- the seasons;
- solar activity.

This ionisation phenomenon is more significant during the day and increases with altitude because of the thinning of the air.

For a high frequency wave, losses are low because the free electrons that vibrate at the frequency of the electromagnetic field that crosses it have a low amplitude movement (because of their inertia) in comparison to the average free movement between atoms.
The probability of impacts is therefore low.
We could say that the wave goes through a « coarse-mesh net ».
VHF and higher frequency waves will (with some attenuation) get through the ionosphere.

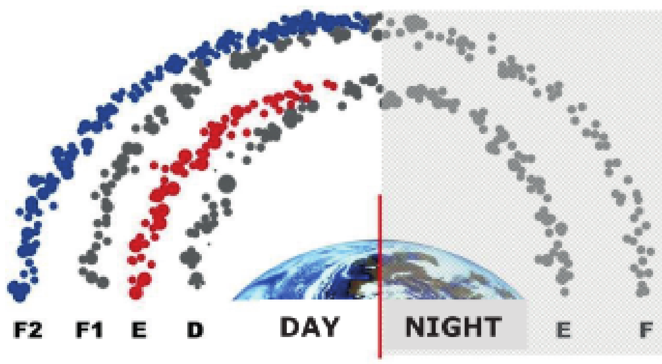
For low frequencies, the amplitude of the movement increases and the losses are more significant (frequent impacts).
We could say that the wave goes through a « tight-mesh net ».

HF and lower frequency waves will not get through the ionosphere.

Layers

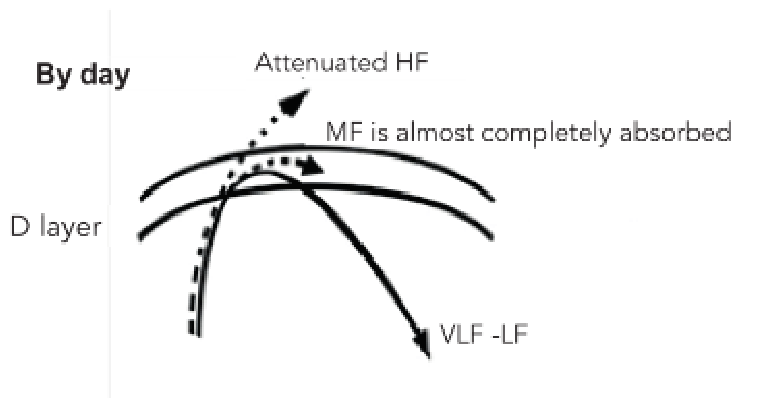
The ionosphere is not a homogenous environment.
It is subdivided in 3 different layers named D, E and F.
These layers play a key role in the propagation of HF and lower frequency electromagnetic waves.
It is difficult to predict disturbances such as solar eruption, magnetic storms, and aurora borealis will modify these layers.

During the day, layer F separates in layers F1 and F2.
During the night, layer D disappears and layers F1 et F2 reunite.



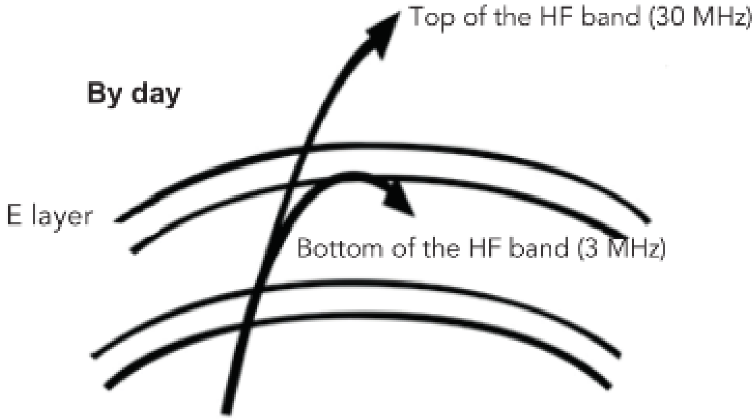
Layer D:

Located between 60 and 90 km above the surface of the earth;
High altitude variations between day and night when it combines with layer E;
In reality, it exists only during the day. This D layer reflects VLF waves and the lower frequencies of the LF band, the other frequencies cross it but are slightly attenuated.

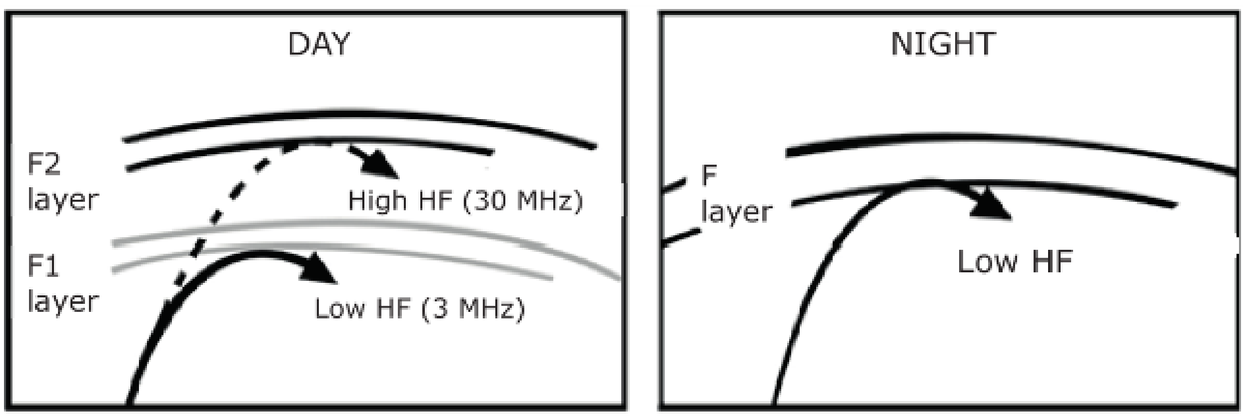


A wave received on the earth after it bounces (reflects) on the ionosphere is called a « sky wave ».

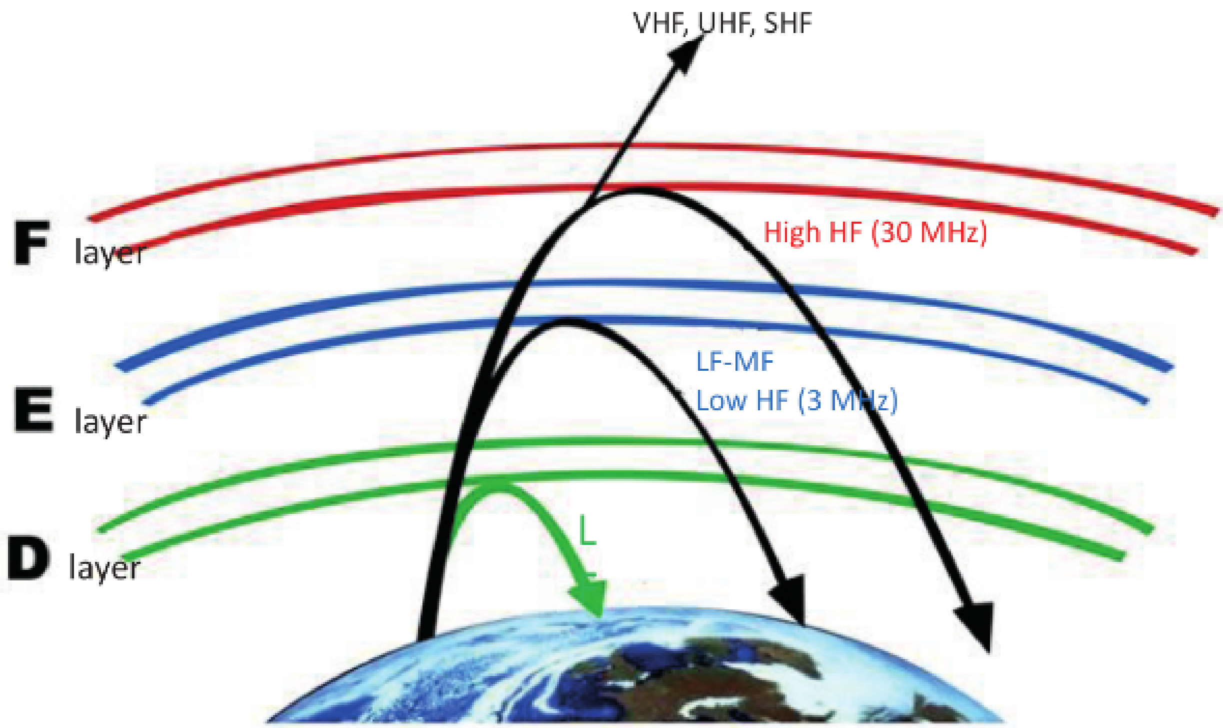
Layer E:
Located between 100 and 110 km above the surface of the earth;
Very stable altitude;
Highly variable ionisation depending on solar light; high ionisation during the day, light persistence during the night;
Reflects LF – MF and low HF waves (primarily during the night because they are attenuated during the day when crossing layer D).



Layer F:
Average altitude 250 km above the surface of the earth;
During the day it splits into layer F1 and F2;
Reflects LF, MF and HF waves.



HF communications will be done on lower frequencies during the night.



SUMMARY OF IONSPHERIC PROPAGATION

3.2 - Ground waves

Bodies that are electrical conductors are very poor propagators of electromagnetic waves due to:
attenuation of the speed of phase of the wave (see 062-01-01 – A);
deviation of the direction of propagation;
absorption of the energy.
A conductor surface acts as a mirror for the waves, nothing gets through, everything is reflected.

Insulating bodies are excellent propagation environments.
The wave going through them is practically unmodified.
They are comparable to a glass in optics.

The ground can be likened to a semi-conductor environment with characteristics that depend on:

- the nature of the ground:
- its moisture level;
- its temperature; (with low frequencies, the absorption is significant in the arctic regions).

Waves that travel close to the ground are called ground waves or surface waves.

Losses due to the interaction of the wave with the surface of the earth increase when the frequency is high, and the conductivity of the surface is low (they are more significant on the ground than on sea).

This is why the transmission range of surface waves is high for VLF, LF and even MF waves, but low for HF and higher frequency waves.

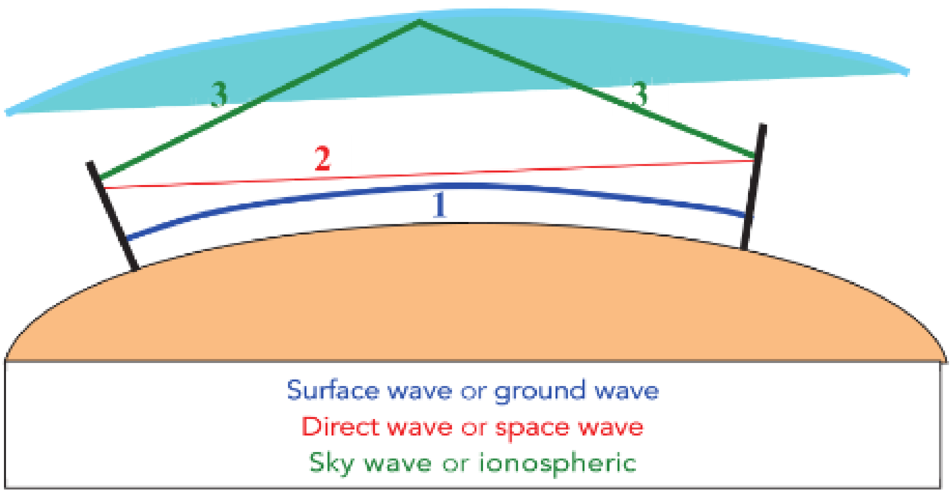
With very low frequencies, there is significant penetration of the wave in the ground; they are particularly used in special military communications.
Example: sub-marines (a few tens of meters).

3.3 - Space waves

We call space waves or direct waves the waves which travel directly between two antennas through the atmosphere (tropospheric layer).
Depending on their frequency these waves will be subjected to certain phenomena that will be described later.
VHF, UHF and SHF type of propagation has a quasi-optical range.

3.4 - Propagation with the frequency bands

Between two points, a wave can follow numerous paths.



A wave transmitted by an antenna could follow all of these paths but:
HF and higher frequency waves are rapidly absorbed by the ground;
VHF and higher frequency waves go through the ionosphere and do not return to the earth.

In conclusion:

VLF, LF, MF waves and to some extent HF waves, travel as ground waves (surface waves) and as sky waves (ionospheric).

VHF, UHF, SHF and EHF waves travel essentially as space waves.

Nota:
Whatever the path that they take, waves always travel following a great circle of the earth (orthodromy).

3.5 - Doppler principle

The Doppler-Fizeau effect is the phenomenon where the frequency of an electromagnetic wave F will increase or decrease if there is relative motion between the transmitter and the receiver.
It also appears when the wave is reflected on an object in relative motion to the transmitter or the receiver.

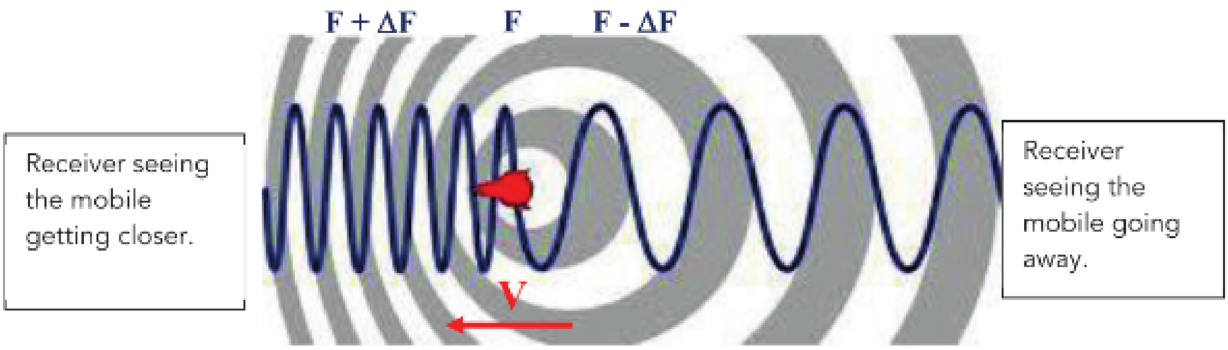


Figure above: a mobile traveling at a velocity V transmits a wave with frequency F, the receiver seeing the mobile getting closer hears the wave at a frequency F+ΔF, while the receiver seeing the mobile going away hears the frequency F-ΔF.

> If the transmitter and the receiver are moving, Δf is proportional to the relative velocity between the transmitter and the receiver.
> If the transmitter only is moving, the receiver measures a velocity which is the velocity of the transmitter.
> If the receiver is moving towards a stationary transmitter, Δf is proportional to the speed of the receiver.

Not on the syllabus:

For a stationary receiver, a wave transmitted at a frequency F0, will be received at a frequency F0 if the transmitter is stationary.

$$F_0 = \frac{C}{\lambda_0}$$

If a receiver is moving towards a transmitter at Vs, the wave will reach it at velocity C to which will be added its own speed Vs; it is as if the wave were travelling at C + Vs.
The transmitted frequency Fo is then received as:

$$F = \frac{C + Vs}{\lambda_0}$$

Inversely if a receiver is moving away from a stationary transmitter, the transmitted frequency Fo is then received as:

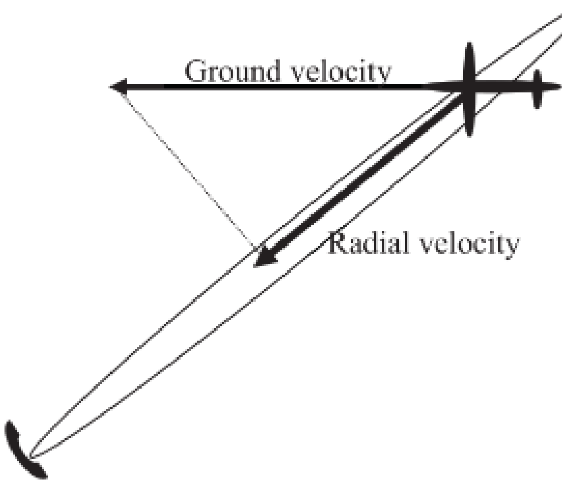
$$F = \frac{C - Vs}{\lambda_0}$$

The analysis of the frequency deviation (F – Fo) will allow the deduction of velocity Vs.

$$\frac{\Delta F}{F} = \frac{Vs}{C}$$

The deducted velocity vector is the velocity vector parallel to the propagation direction of the wave.

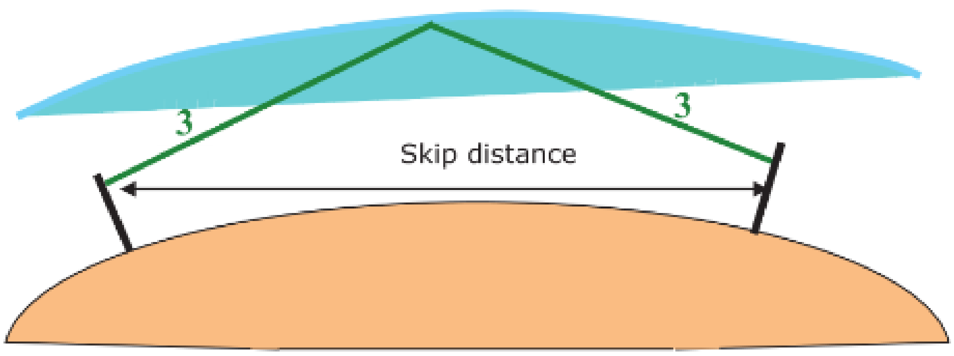
We will measure Δf proportional to the radial speed of the mobile if we are not on the path of the mobile.



A calculator will be able, without difficulty, to recalculate from the radial velocity the velocity on the path and anticipate the trajectory of the mobile.

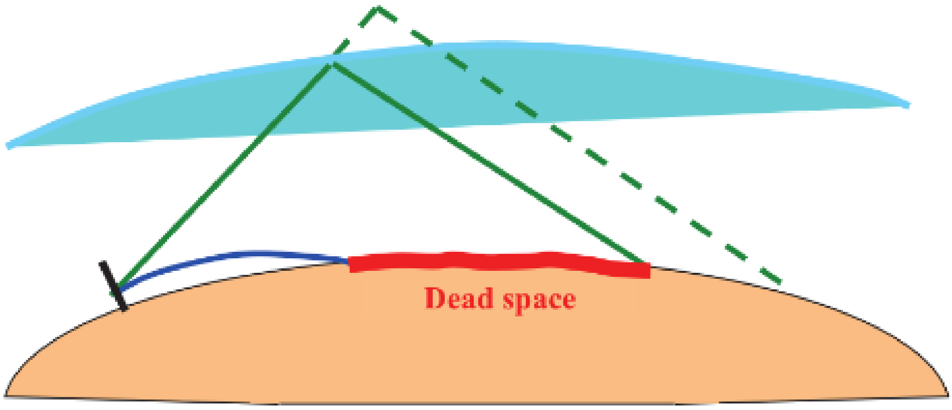
3.6 - Factors affecting propagation

3.6.1 - Skip distance



Skip distance is the distance between the transmitter and the point on the surface of the earth where the first sky wave arrives after rebounding on the ionosphere.
This distance can be many thousands of Nm and even be a trip around the earth after many rebounds.

3.6.2 - Dead space



For a given frequency, we call dead space the area located between the maximum range of the surface wave and the beginning of the reception area of the sky wave.
If rebound occurs on a higher layer, the dead space increases.

3.6.3 - Fading

This phenomenon is caused by the simultaneous reception of both a surface wave and a sky wave.
The two waves having followed paths of different lengths arrive out of phase and cancel each other out.
In extreme cases they arrive in phase opposition and neutralize each other out giving way to what is called fading.

This phenomenon is maximal at dawn and dusk, when the ionosphere is subjected to its strongest and quickest variations.
Note that the reception of two interfering sky waves will produce the same result.

3.6.4 - VHF-UHF propagation

VHF and higher frequency waves do not reflect on the ionosphere but travel through it.
The VHF surface wave is very rapidly absorbed.
VHF communication is therefore done by sky waves.
For this, transmission and reception antennas must be within sight of each other.
We are thus limited to an optical range, but actually because of a particular phenomenon (atmospheric refraction) described below, the waves travel further and we can state that these waves have a quasi-optical range.

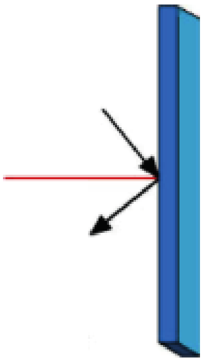
3.6.5 - Physical phenomena

Electromagnetic waves are subjected to, no matter their frequency, more or less 4 phenomena according to their frequencies and the area where they travel:

- Reflection;
- Diffraction;
- Refraction;
- Absorption.

a) Reflection

We will distinguish reflections on obstacles and on the ground from ionospheric reflections, treated later, because they have different causes.
What we call ionospheric reflection is not strictly speaking a reflection but rather a succession of refractions leading the wave back to earth.

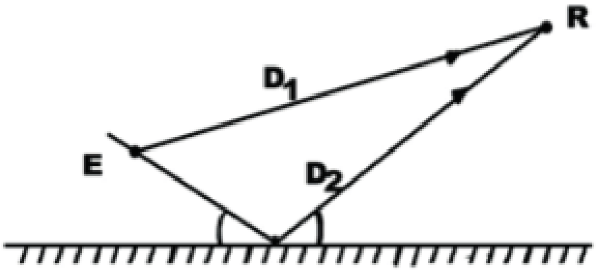


Conductive bodies reflect waves well.
Reflection, on an obstacle or on the ground, generates a change in the phase of the wave.
The strength of the reflected wave depends on the following parameters:

- Reflecting properties of the obstacle;
- Angle of incidence;
- Frequency f: for a given obstacle, the higher it is the more significant the reflection will be.

These reflections are insignificant at low frequencies but significant at very high frequencies.
For example, the primary radar uses this phenomenon.
We will see that there must be a correlation between the wavelength being used and the size of the object to be detected.

In contrast, reflections will cause multipath errors in radio navigation because the receiver will receive a direct wave and a reflected wave with different phases (reflection of the ILS for example).



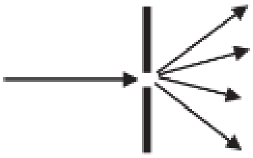
The direct wave emitted at E follows the path D1.
The reflected wave follows the path D2> D1, so it arrives at R with a slight delay (phase shift).

R receives the combination of the two signals.

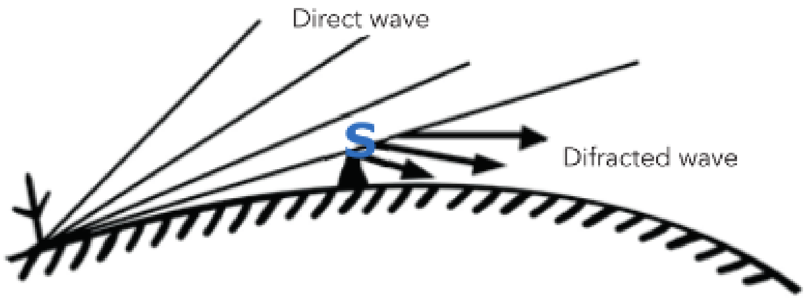
b) Diffraction

It is the physical phenomenon that allows hertzian waves skimming the edges of an obstacle of smaller dimension or equal to the wavelength, to bypass it.

This phenomenon is demonstrated in optics by the Huygens principle and Young's experiment.

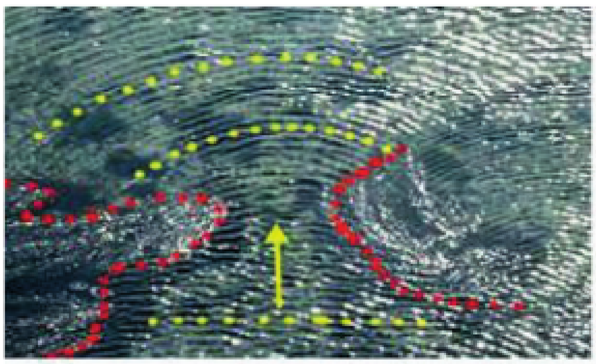


Taking their wavelength into consideration, low frequencies (VLF, LF, MF) have better diffraction behind natural obstacles (hills, mountains) and this allows propagation well beyond the horizon for VLF, LF and MF.

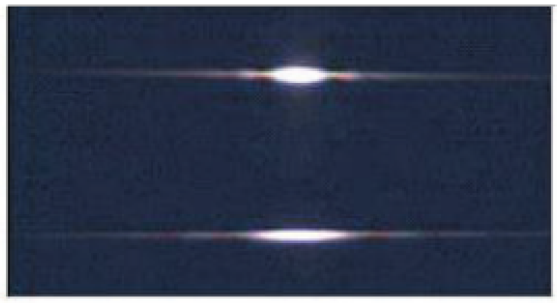


Everything happens as if S had become a new place of emission.

More visible examples:



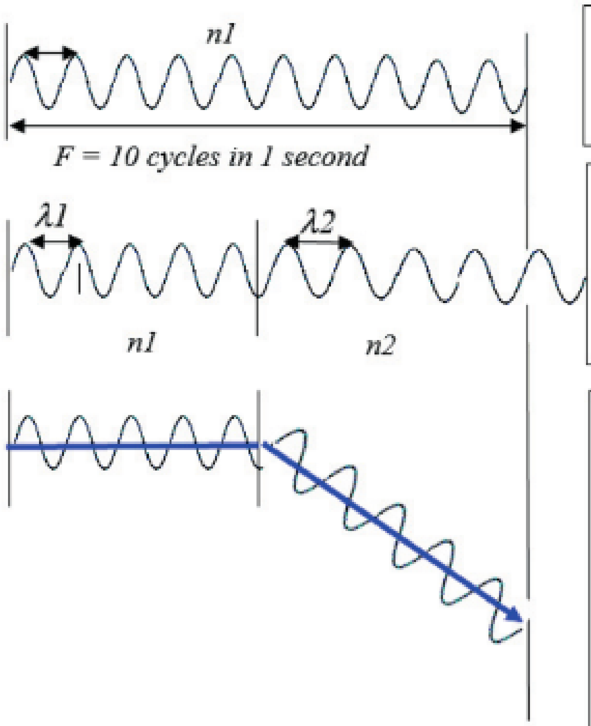
Wavelets are parallel to each other, but fan-shaped when they cross small passages between surface algae.



If we interpose on the path of light coming from a star, a vertical wire, or else a fine slit, we observe a phenomenon of diffraction of light.

c) Refraction

A radio wave is refracted (meaning it changes direction) when it passes from medium M1 to medium M2, for example from one air mass to another air mass with different characteristics (temperature, pressure, moisture, degree of ionisation).
Here is a very simplistic explanation of coastal refraction for the NDB/ADF as of why the wave accelerates and changes direction.
We have studied that a wave moves with a phase velocity (Vp) which is a function of the refraction index (n) of the propagation medium (Vp = C/n).
Thus, we can figure out that if a wave goes from a medium with index n1 to a medium with index n2, the phase velocity will go from Vp1 to Vp2 and therefore accelerate or slow down.
The wavelength ($\lambda = Vp/F$) will then vary but the frequency will not vary because it depends only on how it is generated by the transmitter.



Propagation of a wave with frequency F =10 cycles / second in a medium with index n1 and phase velocity Vp1 = C/n1

Propagation of a wave in a medium with index n1 then in a medium with index n2<n1.
Phase velocity has increased (Vp2) and wavelength has also increase $\lambda = Vp/F$.
Over 10 cycles the wave travels further.

But the frequency cannot change, it is still 10 Hz.
The 10 cycles must always fit into a 1 second duration.
The only solution is to increase the travel distance by curving the wave.

Fermat's principle states :

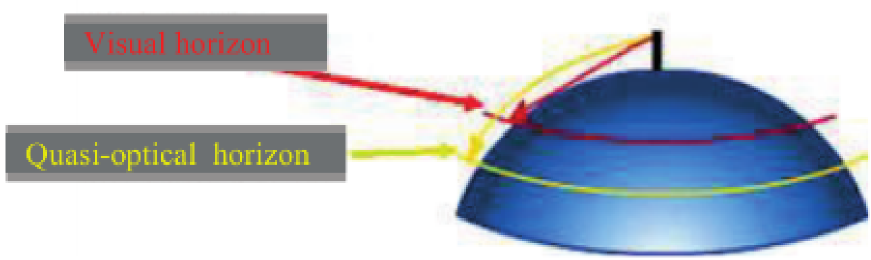
« Light (a wave) propagates from one point to another point on the path that can be travelled in a stationary time. »

Refraction can happen on the horizontal plane and on the vertical plane.

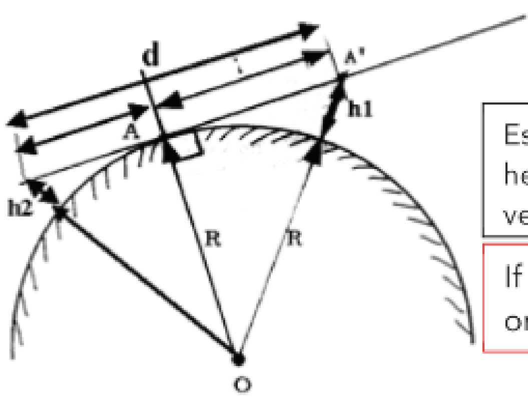
Consequences

- Vertically:

The radio horizon is further by 15% than the visual horizon.



VHF and higher waves propagate on earth (in general) up to the quasi-optical horizon.



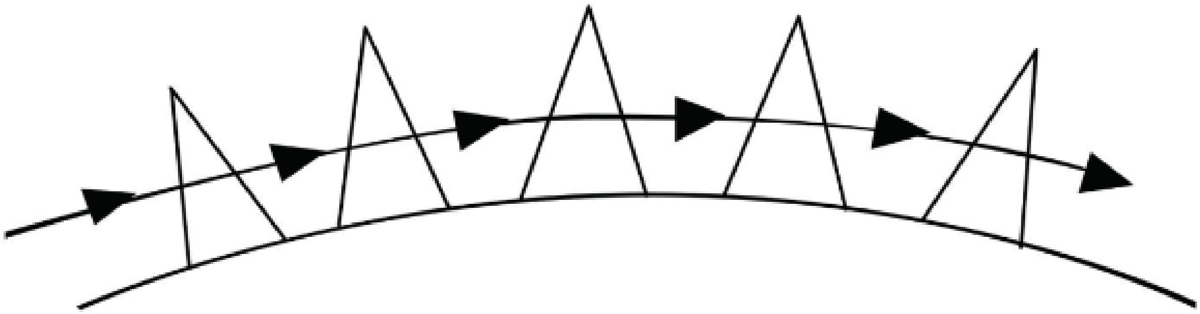
$$d_{NM} = 1,23\sqrt{h_{ft}}$$

Estimated range in NM for an antenna located at a height h and a receiver located at sea level (or vice versa).

If the antennas are at height h1 and h2 (figure on the left), we get:

$$d_{NM} = 1,23\left(\sqrt{h_1} + \sqrt{h_2}\right)$$

In some weather conditions, the atmosphere can act as a series of optical prisms, and successive refractions can bring VHF waves to propagate following the curvature of the earth.
This can bring considerable ranges, totally unusual for VHF and UHF.
This phenomenon is called super refraction (or ducting).

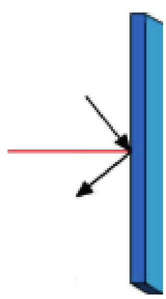


- Horizontally:

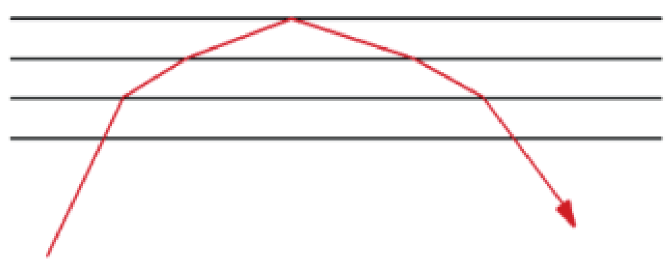
A radio wave that crosses the coast obliquely will deviate horizontally because the ground and the sea have different dielectric constants.
(We will see that this generates an ADF bearing error called coastal error, or littoral error, in the radio navigation course).

- In the ionosphere:

When reaching the ionosphere, HF radio waves are refracted by the different layers of the ionosphere.
For some frequencies, this will lead the waves to return to the earth as if they were reflected (sky waves).
This is referred to as ionospheric reflection (but it is actually a succession of refractions).



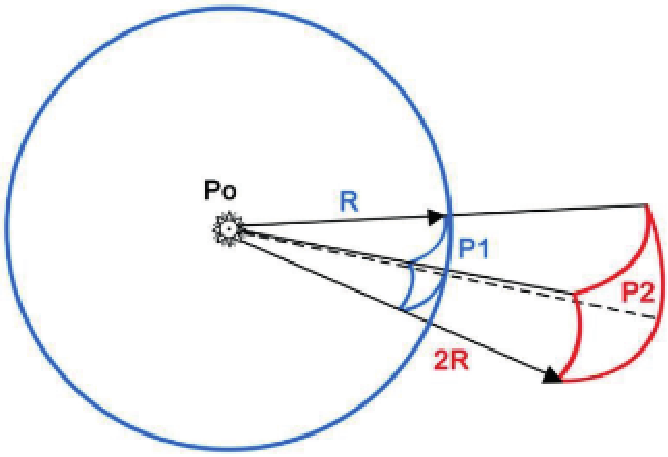
Reflection on obstacles follows Descartes' law.



"Reflection" on the ionosphere is a cascade of refractions that varies with ionization index n of the encountered layers.

d) Attenuation

The level of electromagnetic waves travelling in a vacuum decreases as they propagate.
Imagine an isotropic antenna (a virtual antenna, a point is space that would radiate evenly in all directions the power Po that it is supplied).
The transmitted wave would then propagate as a bubble that would inflate, inflate ...



We demonstrate that a power Po remains constant but is spread out on an increasing surface and thus the surface density of power decreases.

It decreases as the square of the distance.

Therefore, if we want to double the radio range, it will require multiplying the power of the transmitter by 4.
For a radar, a target at a distance D receives a power that has already been decreased by a factor of D^2 and returns an echo that will again decrease by the square of the distance between the target and the radar receiver.
Hence to double a radar range will require multiplying the radar power by 16.
But waves propagate in mediums different from a vacuum and, depending on the medium, absorption will occur that will attenuate even more the signal.
The use of directive antennas (antennas with gain) compensate the attenuation.

e) Absorption

Part of the electromagnetic energy is absorbed by the medium through which it propagates.

- Atmospheric absorption (troposphere):
  Low for VLF, significant for SHF in particular through clouds and rain.
  In general, the atmosphere is an excellent propagator.

- Ground surface absorption:
  Part of the energy is refracted and penetrates the ground.
  Propagation on the sea is better than on the ground.

- Ionospheric absorption:
  Wave attenuation in ionised layers of the atmosphere.
  The intensity depends on the duration of travel of the wave and on the electronic density n of the layers.

Intentionally left blank

02

RADIO AIDS

01 GROUND DIRECTION FINDING (DF)

There are two types of direction finders:

- Ground direction finders, VDF studied in this chapter;
- Direction finders onboard aircraft, studied in the ADF chapter.

1.1 – VHF Directional Finder (VDF) Principles

1.1.1 - Objective

A VHF directional finder is a reception system that allows air traffic control to determine the area of position of an aircraft (which is a straight line), when it makes a transmission on its VHF radio communication system (VHF COM).
It is therefore the ground station that determines the bearing of the aircraft.
The bearing of the aircraft is displayed on a visual system to which the directional finder transmits the value of the measured azimuth, as long as the VHF COM that the pilot is using is transmitting.
The directional finder brings safety to flights that are not under radar surveillance, and VFR traffic.
Directional finders are also used to ensure surveillance of air traffic in order to ease approach operations, and to identify "calling" traffic on their radar screen.

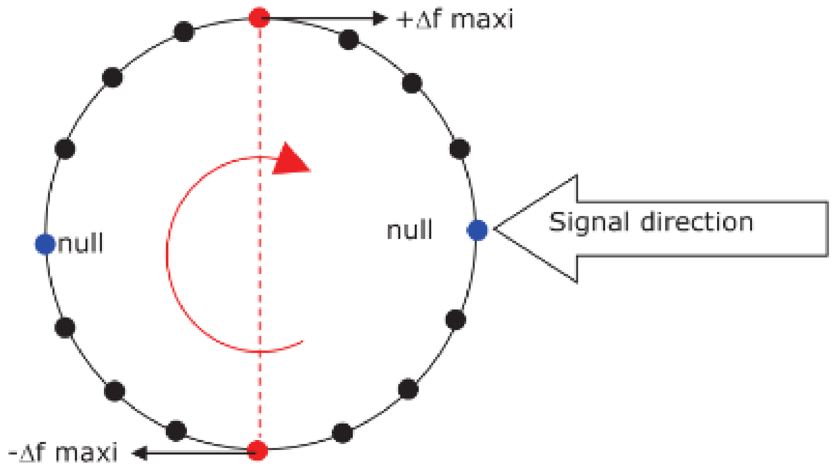
1.1.2 - Characteristics

The VHF directional finder uses, for civil aviation, the VHF frequency band reserved for communication (VHF COM from 118 to 136 MHz).
As a consequence, it is a short or medium distance communication aid.
(See the Wave Propagation course- VHF/UHF characteristics- quasi-optical range).

VHF Directional Finder determines the area of position of a moving body from a GROUND station.
The bearing can be magnetic or true, depending on the antennas setting.
The information is then communicated to the pilot by the GROUND station operator via the VHF communication system.
Hence, it is not an autonomous navigation system.
The ground station can only deal with one aircraft at a time (saturation for one aircraft).

1.1.3 - Operation

The purpose of the VDF is to determine from which direction the signal is coming when an aircraft makes a VHF COM transmission.
Nowadays, the VDF uses the Doppler effect to carry out this operation.
Very simply put, a large number of antennas are positioned in a circle and commute two by two at high speed.
This amounts to constantly having two rotating antennas, one moving closer to the arrival direction of the signal and the other moving away from it.
The antenna moving closer is subjected to a positive Doppler effect (Δf), the other antenna moving away is subjected to a negative effect.
The analysis of these Doppler effects detects the direction from where the signal is arriving.

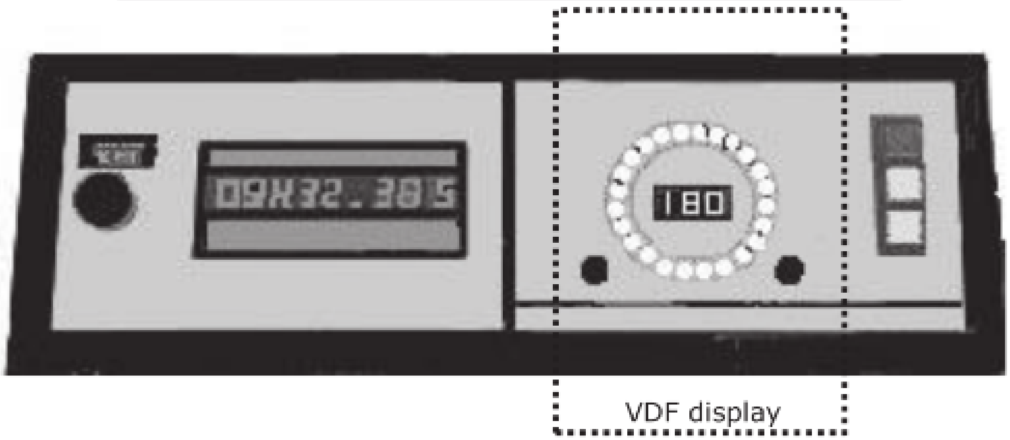


VDF Doppler antenna

1.1.4 – Presentation

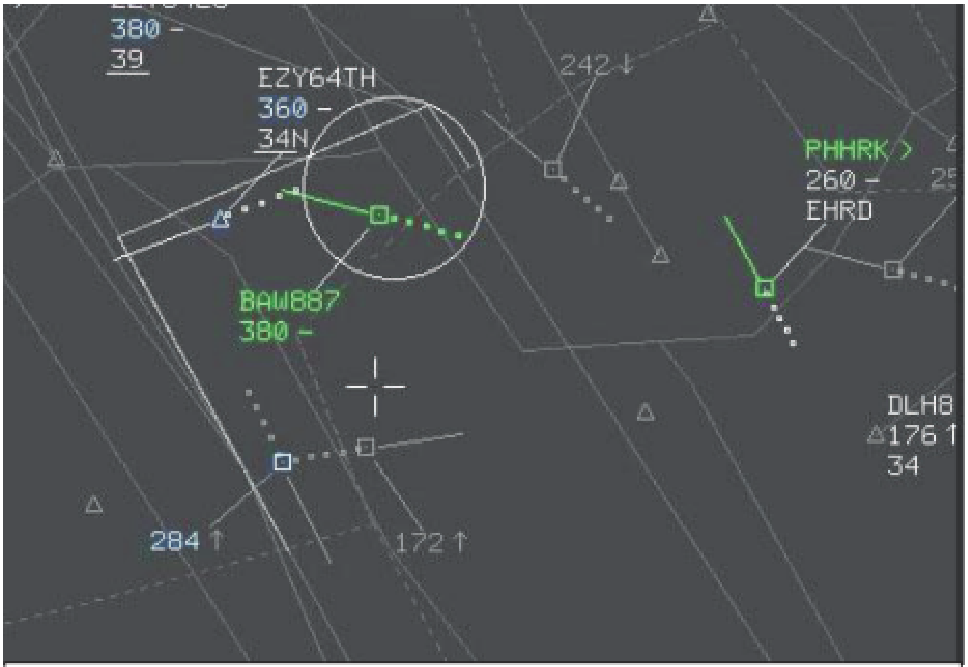
- Digital display

Example of a control panel equipped with a digital display

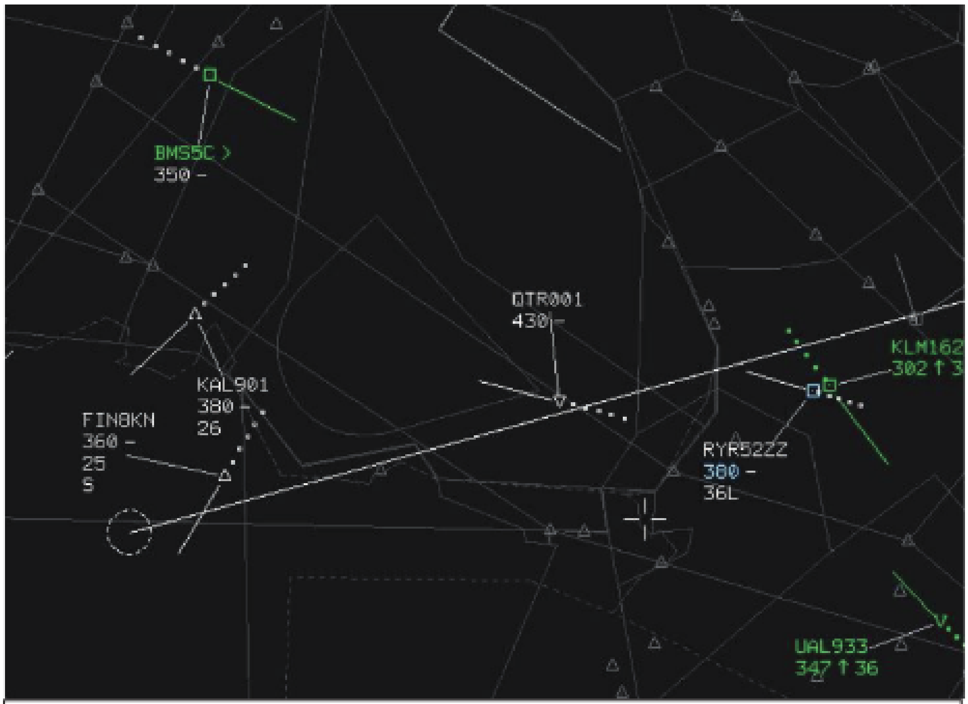


The bearing information is transferred to a liquid crystal display that gives a digital readout, or on a group of LEDs (light-emitting diodes) distributed over 360°, the illuminated LED showing the transmission azimuth.
A switch allows choosing between a QDM or a QDR readout.

This display is often obsolete, and the information is now shown on a computer screen.
It is possible to couple the VDF to the surveillance radar in order to illustrate the target representing the calling aircraft.



Maastricht Control – Calling target is circled (BAW887)

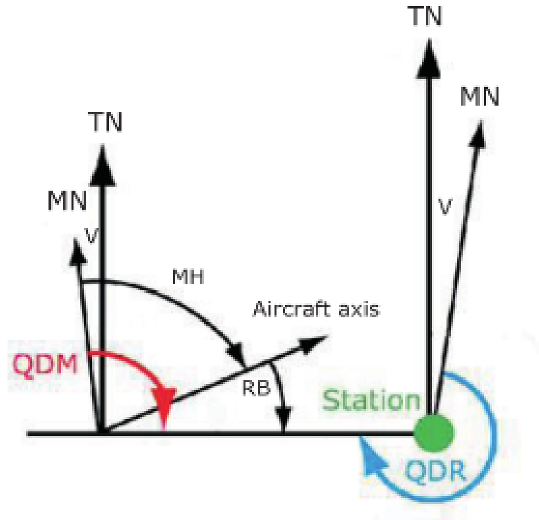


Maastricht Control – the line points to the calling target which is outside of the screen range

1.2 - Presentation and interpretation

- Review of the different bearing definitions

TN True North;
MV Magnetic Variation;
MN Magnetic North;
MH Aircraft Magnetic Heading;
RB Relative Bearing.
V Magnetic variation



QDM Magnetic bearing of the station by the aircraft;

QDR Magnetic bearing of the aircraft by the station

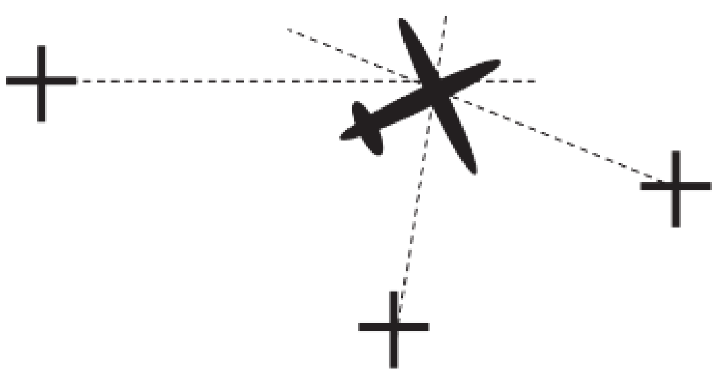
called RADIAL;

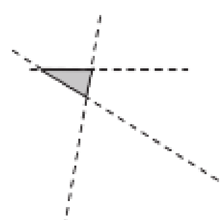
We note: QDM = MH + RB.

Note: QDM is not necessarily synonymous of QDR ± 180 (it is true only if the magnetic variation of the station and the aircraft are the same).

A VDF does not provide a position.
It provides a position area which a straight line.

Several VDF are necessary to obtain a position, by overlapping their lines of position (triangulation).





Triangle of uncertainty due to the imprecision of VDFs

This method can be used in case of EMERGENCY (121,5 MHz).
Some states have a network of stations intended for that purpose.

1.3 - Coverage

The frequency band that is used only allows for a quasi-optical range.

$$D_{NM} = 1{,}23(\sqrt{H_{FT}} + \sqrt{h_{FT}})$$

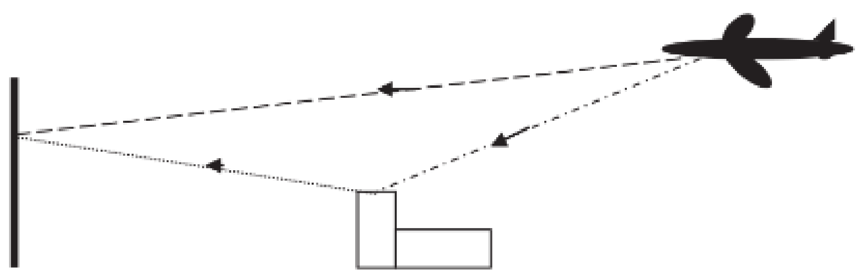
H, aircraft height in feet

h, station height in feet

1.4 - Errors and accuracy

1.4.1 – Reflection error

The signal coming from the aircraft can travel directly to the station but can also be reflected by the ground, or by obstacles close to the aircraft and the station.
In that case the receiver receives a direct signal and a reflected signal.
The reflected signal arriving later than the direct signal (longer travel), a phase shift that can amount to the fading of the signal can occur and be a source of error.



1.4.2 – Polarisation error

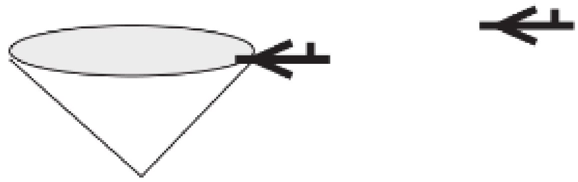
It occurs when the aircraft approaches the station.
Precision is best achieved when the aircraft is high and far from the station (within the limits of the operational range).

1.4.3 – Synchronous transmissions

Because of the operating principle, if many aircraft call the station simultaneously, signals coming from different directions interfere, and it is impossible to have a proper bearing.
It is the synchronous transmissions error.

1.4.4 - Other errors

VDF and aircraft antennas have a vertical polarization, errors can occur when the aircraft is pitching or banking.
Above the station, there is an area of great inaccuracy (cone of silence).



1.4.5 - Precision

VDFs are divided into the following accuracy classes:

Class A	±2°;
Class B	±5°;
Class C	±10°;
Class D	more than 10°.

In general the maximum precision available is that of class B: ± 5°.
When a controller provides a bearing, the class is specified.
For example, « QDM 070, class B ».

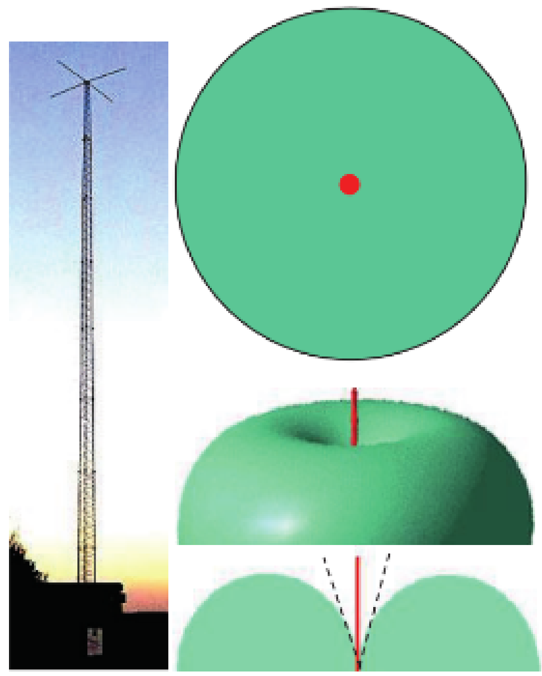
02 NON-DIRECTIONNAL RADIO BEACON (NDB)/AUTOMATIC DIRECTION FINDING

2.1 – Principles

2.1.1 - NDB

The NDBs (and Locator, the principle is rigorously the same) are LF/MF radio beacons.
(They use surface waves, see course on wave propagation).
They are the ground equipment used by the radio direction finders that are installed onboard the aircraft (ADF – Automatic Direction Finding).
The ICAO Annex 10 stipulates that an NDB station has an automatic ground monitoring system.

a) Antenna



The NDB uses an antenna with a vertical polarization, radiating omnidirectionally in the horizontal plane.
Approaching the vertical of the antenna, there is an area where the radiation becomes nonexistent, and thus the signal is not exploitable.
This area, because of its shape, is called: cone of silence (45 to 60° aperture angle).
LOs use the term cone of confusion instead of cone of silence because effectively at low altitude this cone of silence is practically nonexistent.
The term cone of confusion comes from the fact that very close to the station the angular deviation between the QDMs becomes so low that it is impossible to differentiate them.

b) Characteristics of the transmitted signal

The NDB transmits a N0N signal allowing an onboard measure of the relative bearing, AND, a A1A or A2A signal for the transmission of the identifier.

<u>N0N- A1A combined signal:</u>
This is used in areas with a high density of beacons, garbled by radio stations or severely affected by atmospheric parasites.
A1A modulation requires the use of a BFO (Beat Frequency Oscillator) at reception to hear the identifier (the BFO is developed later).

<u>N0N- A2A combined signal:</u>
This is used in areas with a low density of beacons.
The modulation frequency is 400 or 1020 Hz.
This type of modulation does not require a BFO to hear the identifier.
The identifier is transmitted in general 2 times per minute and is composed of 2 or 3 letters (plus 1 digit for oceanic stations).

VHF direction finders used as approach and holding aids (Locator) transmit their identifier 8 times per minute instead of 2 times per minute.

c) Frequency band

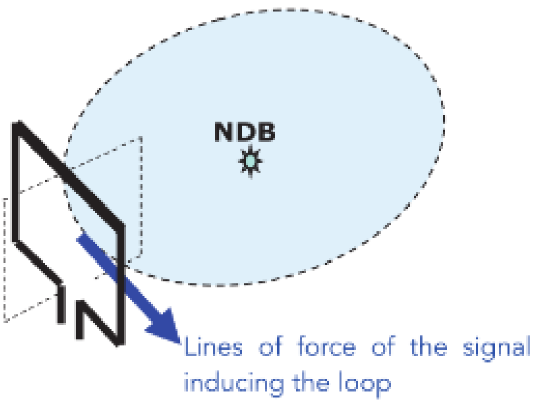
The frequency band assigned to aeronautical NDBs is 190 kHz to 1750 kHz (LF and MF frequency band).
This frequency band is also used by commercial radio stations.

d) Power

It varies considerably according to the use of the beacon.
It is 10 to 25 W for a locator (used for a hold or combined to an ILS) for a range of 10 to 25 NM, and up to more than 2KW for an en-route NDB (range 250 to 300 NM).

2.1.2 - ADF

The ADF (Automatic Direction Finder) is a LF/MF receiver installed onboard an aircraft. An internal system produces, from the signal received from the NDB, an information of relative bearing distributed to applicable indicators.
To detect the arrival direction of an NDB signal, a directive antenna is needed. The loop antenna is best suited for ADF frequencies and features a reduced size.



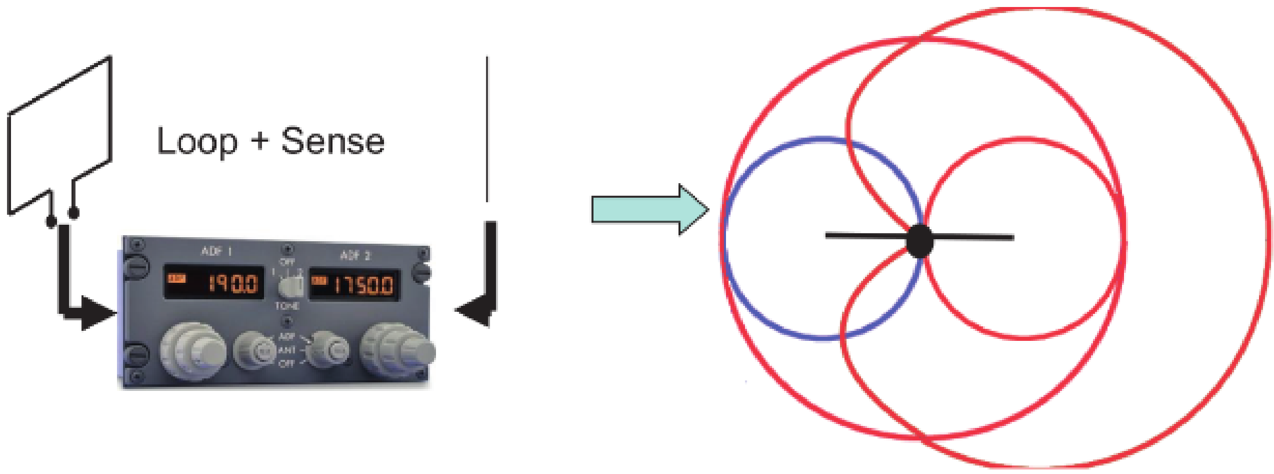
By rotating the loop over 360°, there are two symmetrical positions giving maximum induction in the loop and two positions with almost zero induction.

For reasons of precision in the detected induced signal, it is preferable to use the detection of a null signal (solid line loop on the figure).

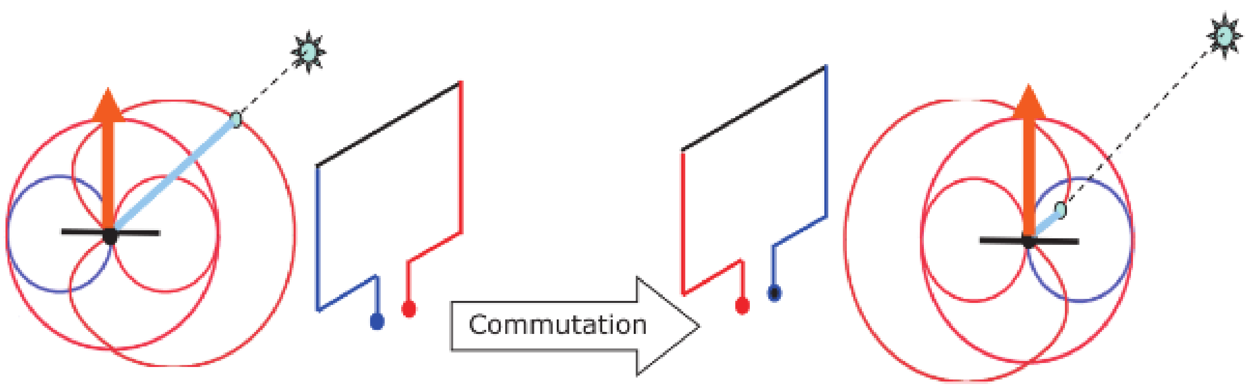
The principle is to rotates the loop until the NDB cannot be heard, we know then that it is in a perpendicular direction to the plane of the loop.

However, there is an ambiguity of 180° since there are two symmetrical positions giving this null signal.

To resolve this ambiguity, a second antenna called sense antenna is installed.
This antenna is omnidirectional.
Combined with the reception diagram of the loop, we obtain an antenna system presenting a marked directivity in only one direction (heart shape diagram called cardioid).

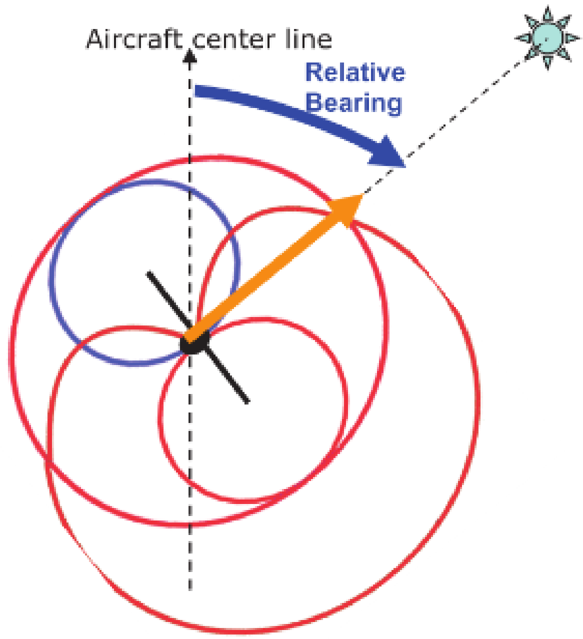


It is possible by commuting the loop to have a directivity to the right or to the left.



Commutation is permanently carried out, automatically and rapidly, by a receiver internal system with the purpose of inverting the phase of the signal of the loop.
The combination of loop and antenna signals produce changes to directivity indicated in the other direction.

Coming back to the two figures above, we see that the loop and the associated amber needle do not indicate the direction of the station.
We also see that, depending on the commutation, the NDB is heard strongly (light blue line on the left figure above) when directivity is on the right, and when the directivity is on the left (light blue line on the right figure above) it is weak.
We know then that the NDB (in our example) is somewhere on the right side of the loop.
A dedicated circuit rotates the loop to the right until obtaining a null loop signal.
The needle associated to the loop then indicates the direction of the station.
The angle between the aircraft center line and the direction of the station is called the relative bearing.



Because the aircraft is not immobile (it is moving towards the North in our example), the loop will soon stop receiving a null induction.
Because of the permanent commutation the NDB will be heard at times loudly, at times weakly at the rate of the commutation.
The servitude system will again lightly turn the loop to find the null induction.
Therefore, the needle indicates permanently and automatically the direction of the station.

Memory:

Initially the loop antenna was located outside the aircraft and was enclosed in covering that generated drag.
This system needed a servo motor that was subjected to become loose and to suffer mechanical wear.

Courtesy of and thanks to Rick Covington, www.airliners.net

Modern ADFs do not possess a rotating loop anymore but use a stationary loop antenna located under the belly of the aircraft (called a Flush antenna).

The stationary loop breaks down the field received by the NDB, transmits the components of the field to a receiver system which recovers the field present outside of the aircraft.
A small internal mobile coil plays the role of the rotating loop.

The NDB does not only induce the loop but also the metal structure of the aircraft which then re-radiates over the loop and the sensor antenna.
This is the source of a bearing error, called quadrantal, which is corrected electrically by the loop.
The re-radiation being variable from one aircraft to the other we cannot transfer a Flush antenna from one aircraft to another one.

Control box:

The control box (here B737) allows:

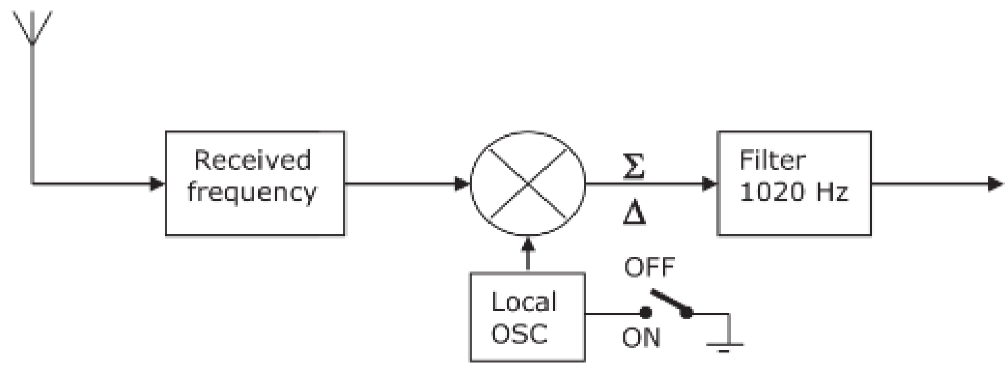
Frequency : 3 concentric knobs allow the selection of the NDB frequency.

Selector: ADF : normal use (loop + sense antenna in use);
ANT : only sense antenna in use; allows adjusting the receiver or listening to the identifier when garbled on the ADF;

TEST: The needle of the indicator turns 90° (in general) when pushed then comes back to its initial position when released.

BFO (TONE): An A1 signal is a pure carrier frequency cut out to the rhythm of the Morse code.

After demodulation of this kind of a wave, no audible signal is available.
A local internal oscillator is introduced with a frequency close to the reception frequency.
By mixing the received frequency with the one of this local oscillator (OSC), we obtain the sum and the difference of the entry frequencies.
A filter allows to recover the useful frequency (set to obtain 1020 Hz).
For an NDB N0N-A2A, there is no need for a BFO in service to hear the identification.



On EFIS aircraft there is no longer a control box, the NDB selection is done through the NAVRAD page of the CDU – FMS.
The system automatically identifies the station and displays its identifier next to the frequency.

The BFO control remains because if the electronic system does not need a BFO to identify an A1A signal anymore, the human ear still does.

2.2 - Instrumental representations

The ADF receiver elaborates the relative bearing of the station (NDB or Locator) and delivers this information to different indicators:
Fixed-card ADF (radio compass);
Moving-card ADF;
Radio magnetic indicator (RMI).

2.2.1 – Fixed-card ADF (radio compass)

The relative bearing is counted from 0 to 360°, or from 0 to 180° Right or Left (here 315 or – 045).

The signal issued by the ADF actuates a needle mounted on a fixed-card graduated from 0 to 360°.
The 0 is here aligned with the lubber line of the aircraft.
The readout at the point of the needle is thus a relative bearing (here – 045°).

This indicator does not directly provide a magnetic bearing to the station (QDM).
It is up to the pilot to carry out mentally:
QDM = Magnetic Heading + Relative Bearing
Namely in our example a QDM = 268° - 45° = 223°

2.2.2 – Moving-card ADF

In order to read directly a QDM (magnetic route to the station), a knob (H or HDG for Heading or OBS for Omni Bearing Selector) allows the manual rotation of the moving-card and the display of the aircraft heading in front of the lubber line index.

The instrument then mechanically adds up MH + RB.

The direct readout at the tip of the needle is a QDM.

Adjacent figure, the pilot has selected a 090 heading on the moving-card.

The relative bearing to the station is -045° (or 315°).

We therefore read a QDM 045 at the point of the needle.

Any change of heading requires a new setting, which is not very convenient.
This is why the RMI ADF was designed.

2.2.2 - RMI

The dial, graduated from 0 to 360° is, this time, enslaved to the heading reference system.

This way, whatever the heading of the aircraft, the instrument mechanically adds up magnetic heading + relative bearing, with no action from the pilot.

The presentation is the same but the Heading knob disappears, as it is no longer needed.

We then read, by the tip of the arrow, the magnetic bearing of the station by the aircraft (QDM) without any action from the pilot.

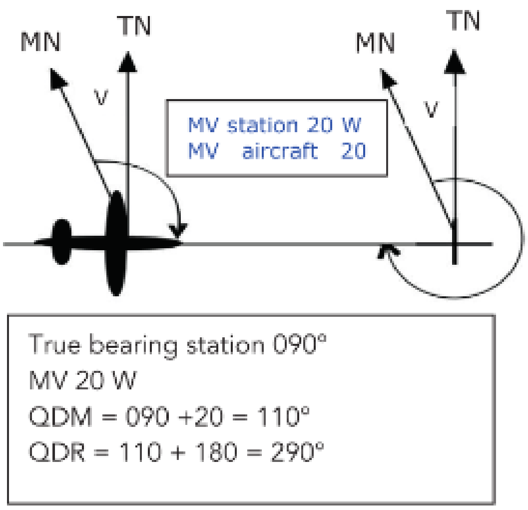
We can read (in general), at the tail of the needle, the magnetic bearing of the aircraft by the station (QDR).

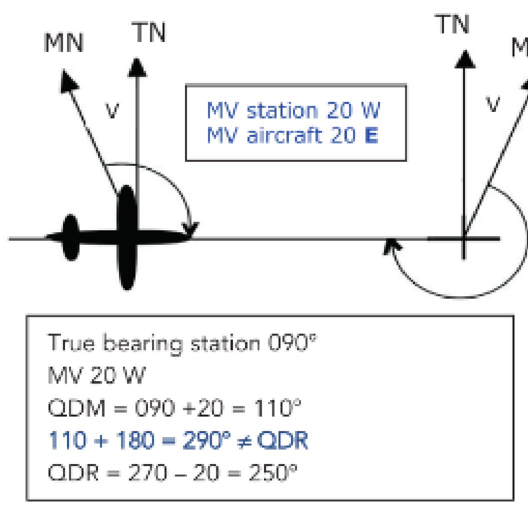
Adjacent figure:
RMI allowing the simultaneous presentation of the bearings of two ADFs.

We will remember that the ADF gives RELATIVE BEARING information and that the RMI provides QDM information namely the magnetic route to follow to reach a station.
To go from the magnetic route to the true route, we use the magnetic variation of the aircraft.
A failure of the heading reference system of the RMI or a mechanical blockage of the moving-card no longer allow the correct sum MH + RB, only the RELATIVE BEARING WILL BE CORRECT (we return to the case of a fixed-card indicator).

Strictly speaking the heading indicated on the RMI is the compass heading, meaning the magnetic heading affected by the compass deviation, and to obtain a true bearing from the station we must use the magnetic variation at the aircraft.

Note on QDM and QDR:
The precedent statement concerning the delivery of a QDR information at the tail of the needle is true, on the only condition that the magnetic variation is the same at the aircraft and at the station.
This is generally true because of use within medium distance, and in areas where magnetic variation does not vary.
It is the case on the figure on the left.
It is not the case on the figure on the right.





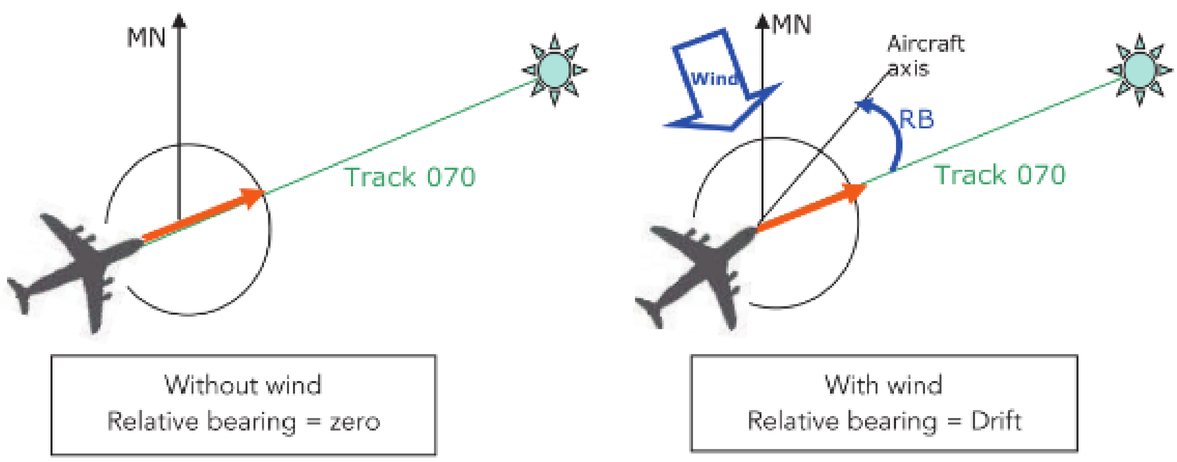
The position area given by an ADF is a straight line, an additional radio navigation mean will be necessary to determine a position (ex: two measures with two different NDBs or better, a second ADF, or a combination ADF-DME).
Unlike other indicating systems like the RMI VOR, the LOC and GLIDE indicators, the ADF does not provide any indication relative to the integrity of the displayed indication (no FLAG or ALARM).
For this reason, during an NDB approach we constantly must listen to the identifier.

2.2.3 - Interpretation

The ADF is used for en-route navigation, to follow a route marked by NDBs, make position overlaps, on approach to an airport (used as an approach aid or LOCATOR), or as a holding fix (for holding pattern).

a) Navigation to an NDB

When there is no wind, heading = track, navigation is done with zero relative bearing.
When there is wind, the drift must be compensated (heading ≠ track) and there is a relative bearing.
Of course in both cases, with an RMI, we read the same QDM, but a different relative bearing.

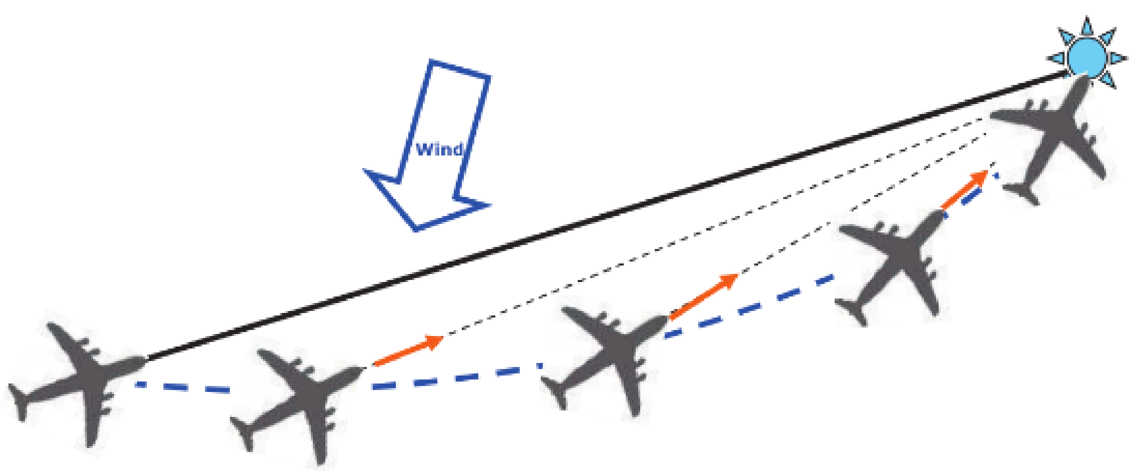


Question example:

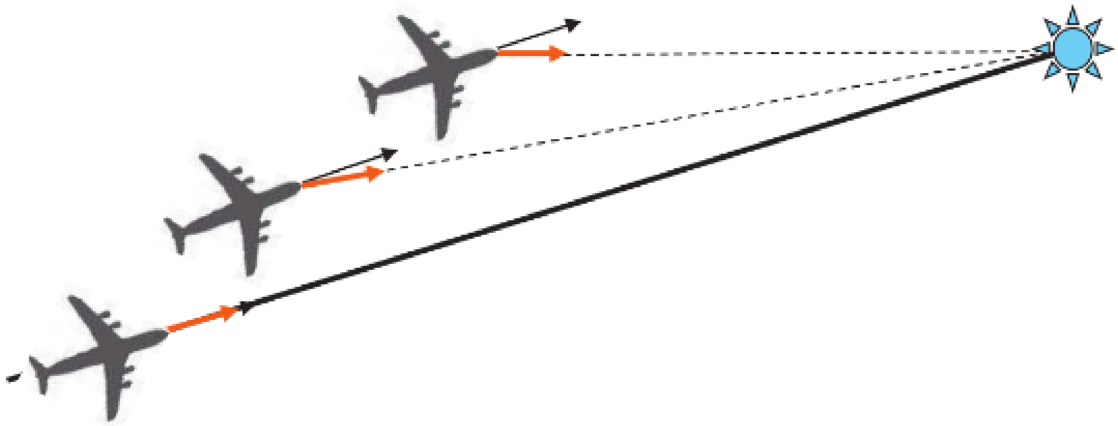
If, flying to a station, in order to maintain the QDM you must decrease the heading, the wind is coming from: 1. the left, 2. the right, etc.
Note that, for practice, it is good to overestimate the drift; in order to come back on the desired QDM, it can just be decreased with the wind bringing us on track (much easier than to come back on track facing the wind).

b) Homing

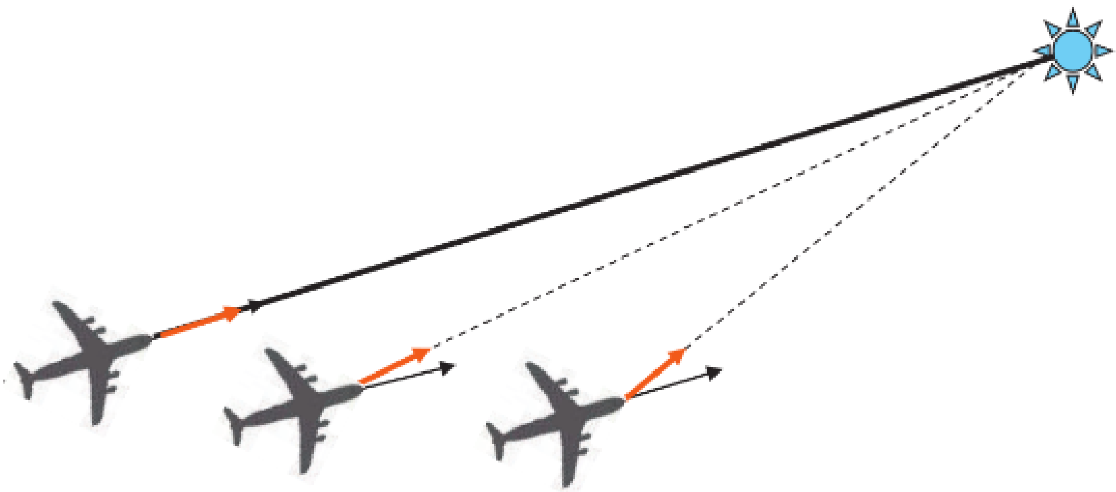
Technique of constantly maintaining zero relative bearing.
An aircraft flying without anticipating the drift and constantly correcting its heading to maintain zero relative bearing follows a particular curve to correct its heading even more as it approaches the station (we do not of course fly this way).
Colloquially speaking, it is called a «bird-dogging».



Constant heading modifications to keep zero relative bearing.



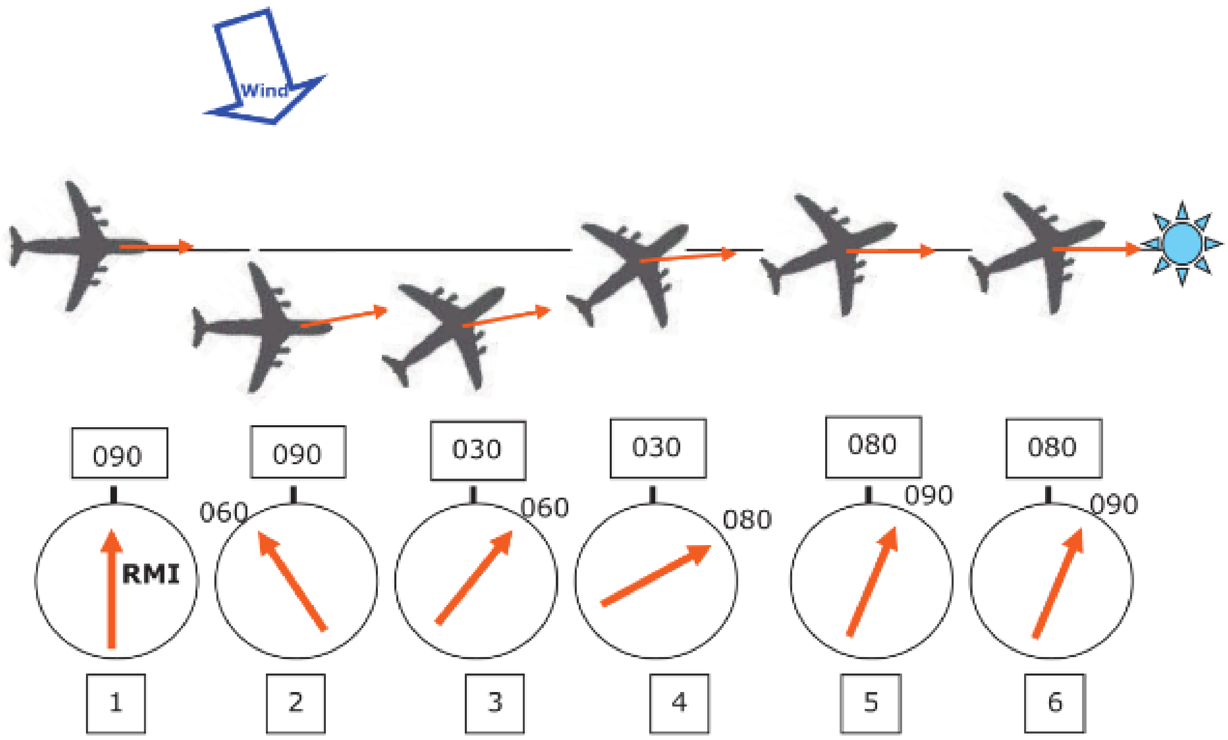
If, with constant heading, the relative bearing increases, the drift is on the left.



If, with constant heading, the relative bearing decreases, the drift is on the right.

c) Determining drift angle and tracking

Given an aircraft wanting to fly towards an NDB on QDM 090° subjected to a wind coming from the left:



1 – With no knowledge of the wind (theoretical case) relative bearing is assumed at zero.

2 - The relative bearing decreases, the drift is on the right, the nose of the aircraft must face the wind (heading decreases).

The new heading must make the needle go to the other side of the lubber line.

(Mnemonic technique: the tip of the needle can only go down).

3 - New heading 030°, QDM is always 060°.

The aircraft faces the wind towards the track to follow 090°.

4 - As the aircraft approaches the track, the relative bearing evolves, and the indicated QDM gets closer to 090°, the heading must be reduced progressively (anticipation) otherwise there is a risk overshooting the 090° track because by the time that the 50° turn is accomplished the track will be overshot.

5 - To stay on QDM 090°, heading will not be 090° because we now know that the wind is coming from the left.

Not knowing the strength of the wind this might not be enough, the relative bearing will then decrease again, we will then fly heading 070° for example.

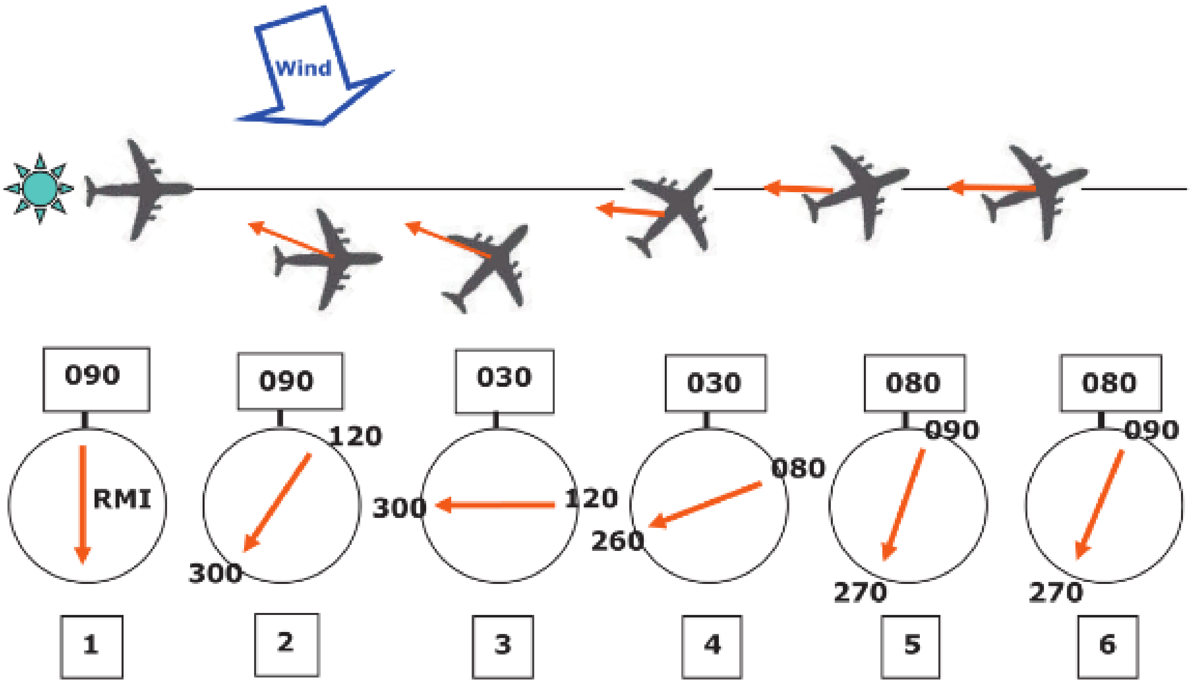
6 - If the drift angle is correctly estimated, the aircraft will stay on track (QDM 090°).

However, approaching the station, the relative bearing will become harder and harder to maintain.

We will maintain an average heading while overflying the NDB, then, one minute after flying vertical (determined by one or many tipping over of the needle) we will resume with the corrections flying away from the station.

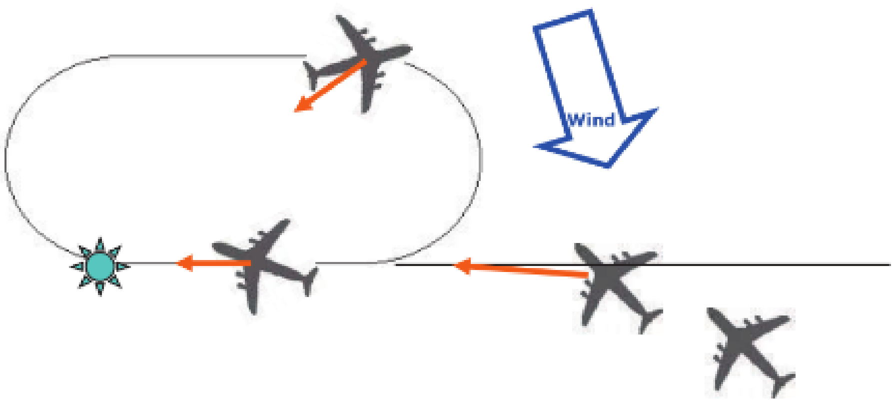
d) Tracking away from an NDB

Flying away on QDR 090°



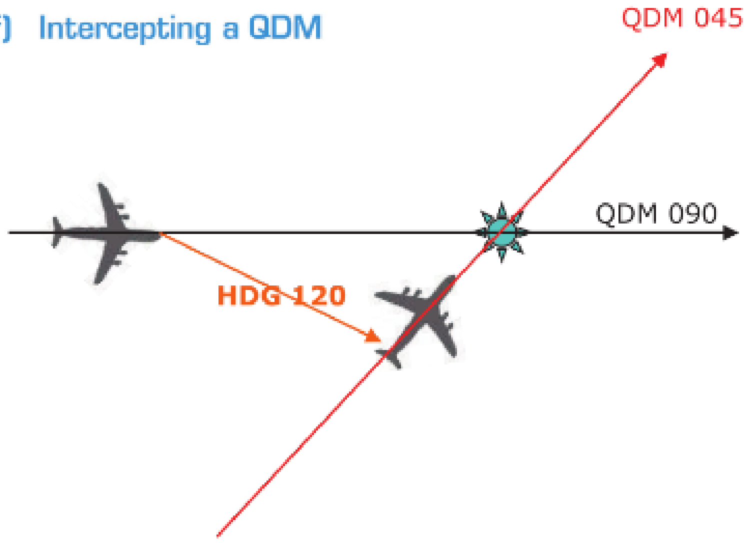
The sequence is the same as when flying towards the station, with the difference that we work with the tail of the needle (which can only go up).

e) Holding



Management of drift is the same as seen before.

f) Intercepting a QDM



As a general rule we open up by:
30° if we are more than 2 minutes from the station,
20° if we are more than 1 minute from the station,
10° if we are less than 1 minute from the station.

g) Approach aid

The management of the horizontal path uses the same tracking techniques to and from the station that we have seen before.
The management of the vertical path is done cross-checking the altitude at which we should pass the beacon and vertical speed/time or other means associated with the descent (these techniques will be developed during practical training).
The inaccuracy of the ADF as an approach aid does not allow to fly precision approaches.
According to ICAO DOC 8168 an aircraft is considered established on the required bearing within ±5° of the center line.

2.3 - Coverage and range

The NDB operates in the LF-MF frequency band.
The surface wave diffracts well behind hills and mountains which explains the long range of NDBs.
The lower the frequency, the longer the range (better diffraction, less attenuation).
The range also depends on the power of the NDB and is higher over the sea (better conductor).
Remember that to double the range of a transmitter we must multiply its power by 4.

Range of the Locator: 10 to 25 NM.
Range of the NDB: at least 50 NM and up to several hundreds of nautical miles.

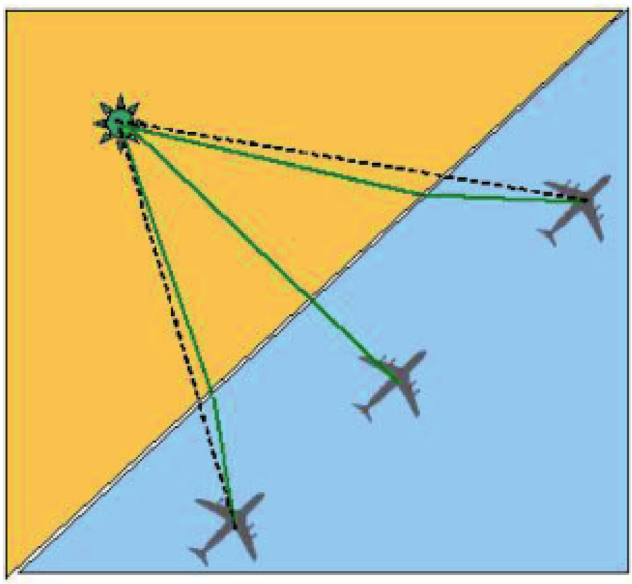
2.4 - Errors and accuracy

2.4.1 – Propagation anomalies

At night LF-MF waves reflect well on the ionosphere and travel a very long distance as sky waves.
The ADF receives signals with different trajectories (surface wave or sky wave from NDBs or distant commercial radio stations), with different phases that create interference and introduce bearing errors.
When the waves arrive in total phase opposition at significantly comparable levels, they neutralize each other and fade.

- Coastal error (or coastal refraction error)

The wave travelling from the NDB to the ADF transits through environments with different refraction indexes (from land to sea) which bend the wave.



Passing the coast, the wave which was travelling over land where the refraction index is n1 now transits to the sea where the refraction index n2 is less than n1.
The phase velocity of the wave equal to C/n (cf. wave propagation) increases and the wave changes direction creating bearing errors.

2.4.2 - Receiver

Of course, performances (sensitivity, selectivity etc.) influence the reception qualities and therefore the accuracy.

2.5 - Factors affecting range and accuracy

2.4.1 – Local thunderstorm

Nearby thunderstorms generate parasites that cause momentary bearing errors.
However, an area of high density of storm cells can produce a bearing error during a considerable time (succession of discharges in the same azimuth).
This error is the greatest among the possible errors of the ADF.
It can reach 90°.

2.4.2 – Mountain effect

LF-MF waves diffract on large size obstacles (hills or mountains) but also reflects on them.
The reception of those reflected signals coming from a direction other than the one of the NDB leads to significant bearing errors mostly during flights at low altitude.

2.4.3 – Accuracy

Accuracy decreases outside of a specified coverage area where the ratio signal/noise becomes too low.
There is therefore a certified operational coverage area, exclusively during the day (because of skywaves at night), that takes into account the cited errors, and guarantees an accuracy of ± 5°.

03 VHF OMNIDIRECTIONAL RADIO RANGE (VOR)

3.1 - Principles

The VOR (VHF Omnidirectional Range) is a short and medium distance omnidirectional radio navigation beacon operating in the VHF frequency band.
The use of this VHF band eliminates disturbances that occur in the MF band (thunderstorm parasites for example, no disturbing skywaves).
Coverage is thus certified by day and by night (unlike the ADF).
In the VOR navigation system, the bearing is given by the GROUND station (QDR) and is detected onboard the aircraft by means of a receiver and an indicator.
The frequency band in use gives a quasi optical range.

There are two types of VOR in operation:
- CVOR (Conventional VOR): old generation;
- DVOR (Doppler VOR): new generation.

Different applications are based on those two VOR technologies:
- En-route VOR: for en-route navigation;
- TVOR (Terminal VOR): used as an aid for approach or departure;
- VOT: VOR station emitting a test signal.

3.1.1 - Frequencies

The VOR frequency band is allocated according to ICAO Annex 10, chapter 4.
Channels are spaced by 50 KHz.
The band goes from 108 MHz to 117,95 MHz with the following precisions:

108 to 112 MHz band:
Shared VOR / ILS band (40 channels);
VORs have a low range (T- VOR, P = 50 W);
Only frequencies which have an EVEN number in the first decimal place are used by VORs.
Ex: 109,8 MEL; 108,25 TSU.

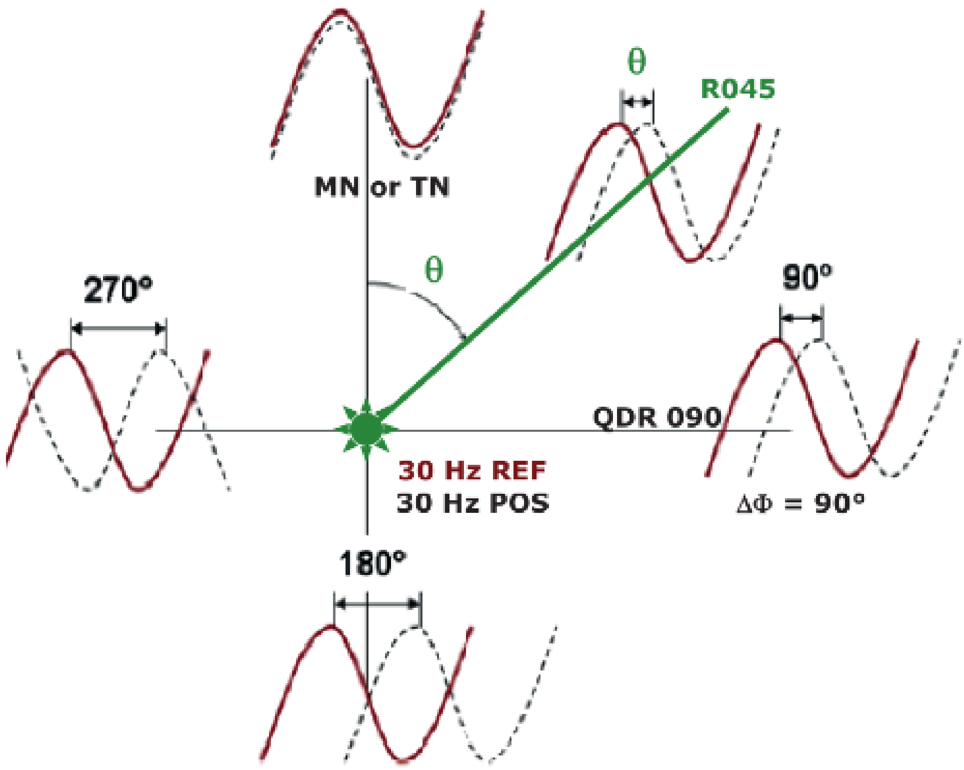
112 à 117,95 MHz band:
VOR only reserved band (120 channels);
P = 200 W (En-route navigation VORs).
Ex: 117,1 LUX; 117,3 RLP; 117,45 BLM; 117,5 KIR.

3.1.2 - CVOR ground station operation

The ground station produces a signal, called position signal, that has a variable phase according to the transmission azimuth.
For this information it must also possess a reference signal to be represent information of magnetic bearing (or true bearing in some cases),
Hence the radiation of a second signal, called reference signal, that has a phase independent from the azimuth.
These two signals have a modulation of 30 Hz.

Onboard the aircraft, a receiver compares the variable phase with the reference phase to deduct on which bearing the ground station sees the aircraft. This is a QDR information.

Note:
In general, position and reference signals are in phase at the MAGNETIC north of the station.
In some cases (areas where the magnetic variation changes very rapidly) there may be VORs with signals in phase at the TRUE north (true bearing of the aircraft by the station).



Reference and position signals as a function of the azimuth

The ground station comprises:

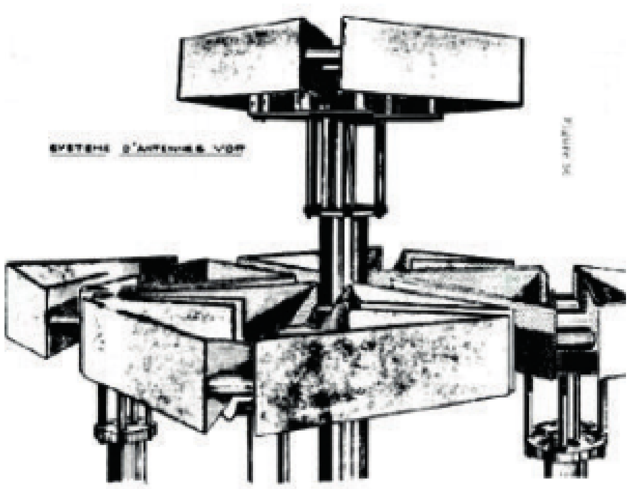
- A VHF transmitter characterized by its frequency and its power;
- A modulation system;
- An aerial system allowing horizontal polarization;

In this way, the ground and the obstacles close to the ground produce an important absorption but less reflection which is a source of error (cf. propagation course).

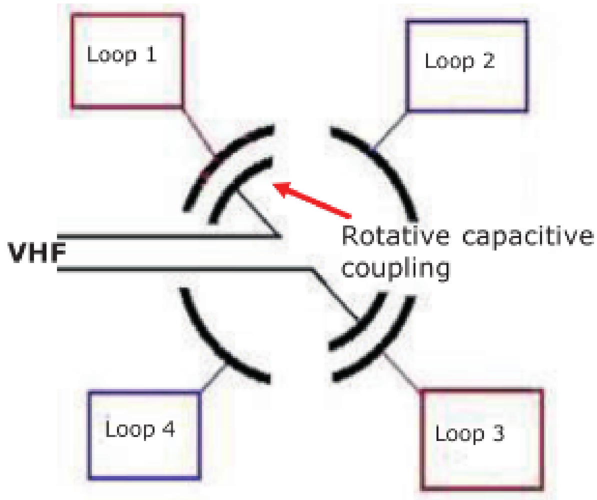
- A surveillance system: for quality and precision of the transmission.

Note from the author:

The LOs would like me to give you this definition of the CVOR: a first-generation VOR station emitting signals by means of a rotating antenna.
I refuse by principle because no antenna ever rotated.
The first generation of VOR transmitted the position signal by mean of a rotative direction finder, but only the power supply of the antenna used rotative capacitive coupling.
The antennas called Alford loop antennas were stationary.



CVOR aerial made of 5 Alford loop antennas



Electrical supply of the 4 loops

The four loops at the base are supplied two by two by the rotative capacitive coupling and radiate a diagram in the form of a figure of eight rotating at 30t/s.
The VHF carrier is received by the VOR receiver as a VHF signal modulated at 30 Hz with a phase that depends on the position of the receiver around the station.
The fifth loop (on the top) radiates the reference signal omnidirectionally.

With the evolution of technology, the antennas became slot antennas powered by an electronic switch.

The four slots of the antenna are powered like the four loops of the Alford antenna but with an electronic commutation, therefore it is stationary, radiating the rotating diagram at 30t/s.
The cylinder in its entirety is the equivalent of the omnidirectional antenna emitting the reference signal.

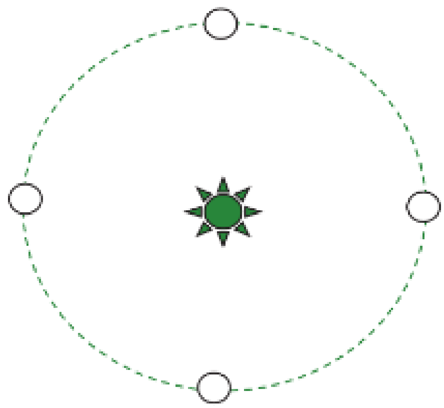
Emission characteristics:

The VOR emits two 30 Hz signals.

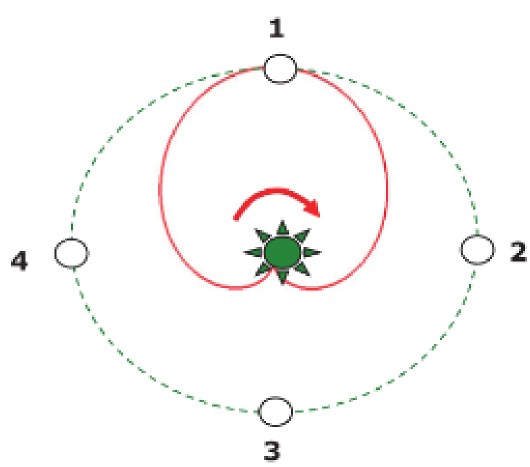
Then, how can we recognize which is the reference signal and which is the variable signal?

To achieve this, two different types of modulation are used.
The reference signal is emitted with a frequency modulation.
The position signal is emitted to be received as an amplitude modulation.

The reference signal, modulated in frequency, is emitted omnidirectionally.
Everyone receives the maximum level at the same time (same phase everywhere).



The position signal (or variable signal) which is a pure non-modulated carrier signal is emitted in a diagram rotating at 30 tours/second.
This rotation makes the signal received a signal that is modulated in amplitude at all points.
At position 1 at an instant t the signal is maximum, then because of the rotation it decreases as it moves away from position 1 until it becomes minimum when reaching position 3 where it will increase again as it goes back to position 1.
Each observer (VOR receiver) placed at position 2, 3, 4 will observe the same phenomenon but at a different time (variable phase).
At position 1 the reference signal and the variable signal are in phase (they reach the maximum at the same time).
At position 2 the variable signal will be maximum when the emission will have turned 90° whereas the reference signal will be maximum at the same time as in position 1.
The signals are out of phase by 90°.
The same thing applies for position 3 and position 4 but with a phase difference of 180° and 270°.



An internal system in the VOR receiver capable of measuring the phase difference between the reference signal and the variable signal will indicate on which radial the receiver is.

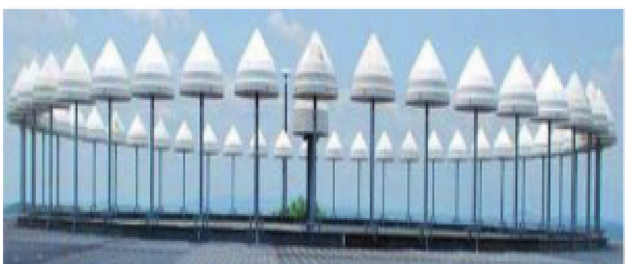
To finish with the ground station, we note that the emitter possesses a modulation circuit in order to transmit the identifier of the station in Morse code.
The identifier modulation frequency is 1 020 Hz.
Another modulation circuit allows transmitting voice messages (A3A), for example for the **ATIS**.

3.1.3 – Doppler VOR (D-VOR)

They are the new generation VORs, with a higher precision and less sensitivity to site errors (conceived to reduce propagation errors).

The difference resides in the modulation technique of the carrier wave.
48 antennas forming a circle emit with high speed commutation (producing a diagram rotating counterclockwise at 30 t/s).
The receiver sees the emitted wave affected by a Doppler effect like a frequency modulation at 30Hz (reception of F0, F0+ΔF, F0, F0-ΔF, etc.).
The central antenna emits omnidirectionally a signal modulated in amplitude at 30 Hz.
We find again our 30REF and 30VAR signals.
Differences are completely transparent to the onboard receiver, and the operation of the navigation instruments (CDI, HSI or RMI) is the same whether the ground station is a C-VOR or a D-VOR.

A D-VOR is used exactly like a C-VOR.



D-VOR

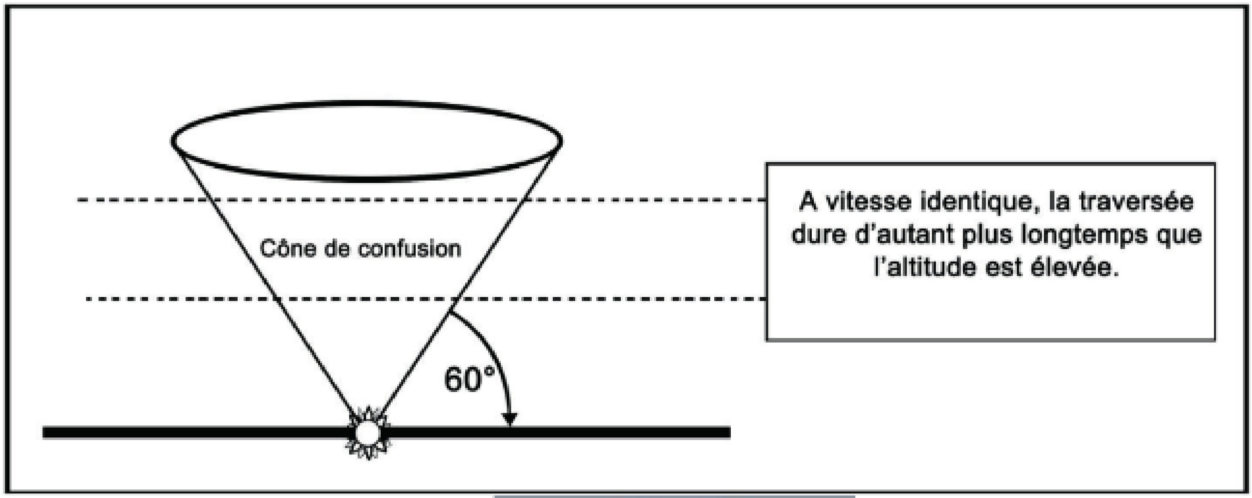
Based on these two technologies (CVOR and DVOR) we have:

- En-route VOR (en-route navigation, high power);
- TVOR (Terminal VOR, used as an approach aid, low power);
- VOT (Test VOR indicating radial 360° or 180° whatever the position of an aircraft);
- ATISVOR (VOR broadcasting an ATIS);
- VORDME or VORTAC (VOR collocated with a DME or a TACAN).

The ground equipment is complex and costly but propagation errors < 1°.

3.1.4 - Cone of confusion

In the vicinity of the ground station there is a cone of confusion, where the received information becomes invalid. The ICAO requires a coverage at least up to 40° above the horizon.
(In practice, modern VORs cover up to 60°, even 80° having a smaller cone of confusion).



While crossing the cone, a NAV flag will indicate on the instrument that the VOR signals are no longer received.

3.1.5 - Monitoring

The ground station is monitored automatically by a receiver located nearby.
The strength and the quality of the emitted signal, and the bearing accuracy are monitored.
Depending on the fault, the signal identifier will be removed or the emission will stop, and maintenance will be automatically informed or emission will shift to an emergency emitter.

3.1.6 – Onboard station

a) Constitution

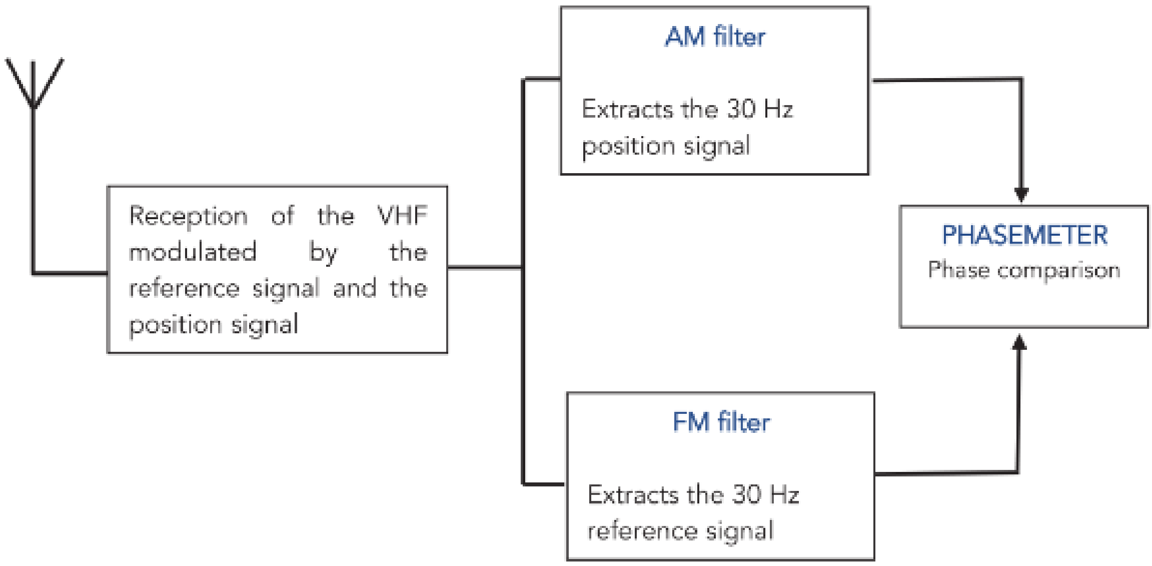
- a VHF receiver. (We will see later that a part is common with the LOCALIZER);
- an omnidirectional antenna (VOR must be received on all azimuth);
- a control box;
- a control system to display the selected route;
- various indicators (CDI, RMI, HSI developed in the Navigation Instruments chapter).

b) Receiver principle

It must retrieve 30 Hertz POSITION and REFERENCE signals contained in the VHF carrier, then measure the phase difference between the signals.
At reception, the complex emitted signal is received, and oriented to two different processing paths.

1) A filter keeps only the 30 Hz component attributable to the rotation of the emission which is seen as an amplitude modulation (position signal).
2) The same signal is sent to a path tasked with retrieving the signal emitted in frequency modulation (reference signal).

The obtained signals are sent to a phasemeter; the phase difference shows the radial on which the receiver is.



The measure of the phase difference of the 30 Hz signals is achieved by two parallel processing chains.

Each of these chains feeds one or more indicators that will be developed later.

1) One chain is called manual chain:
The pilot selects a course to follow. This chain shows the deviation from the selected track, flying is carried following the needle and cancel the deviation.
It feeds the CDI or the HSI seen below.

2) The other chain is called automatic chain: the measure of phase is performed by a servo showing the radial on which the aircraft is at all times.
It feeds the RMI, seen below, which supplies the radial on which the aircraft is automatically at all times, and with no input from the pilot.

3.2 - Presentation and interpretation

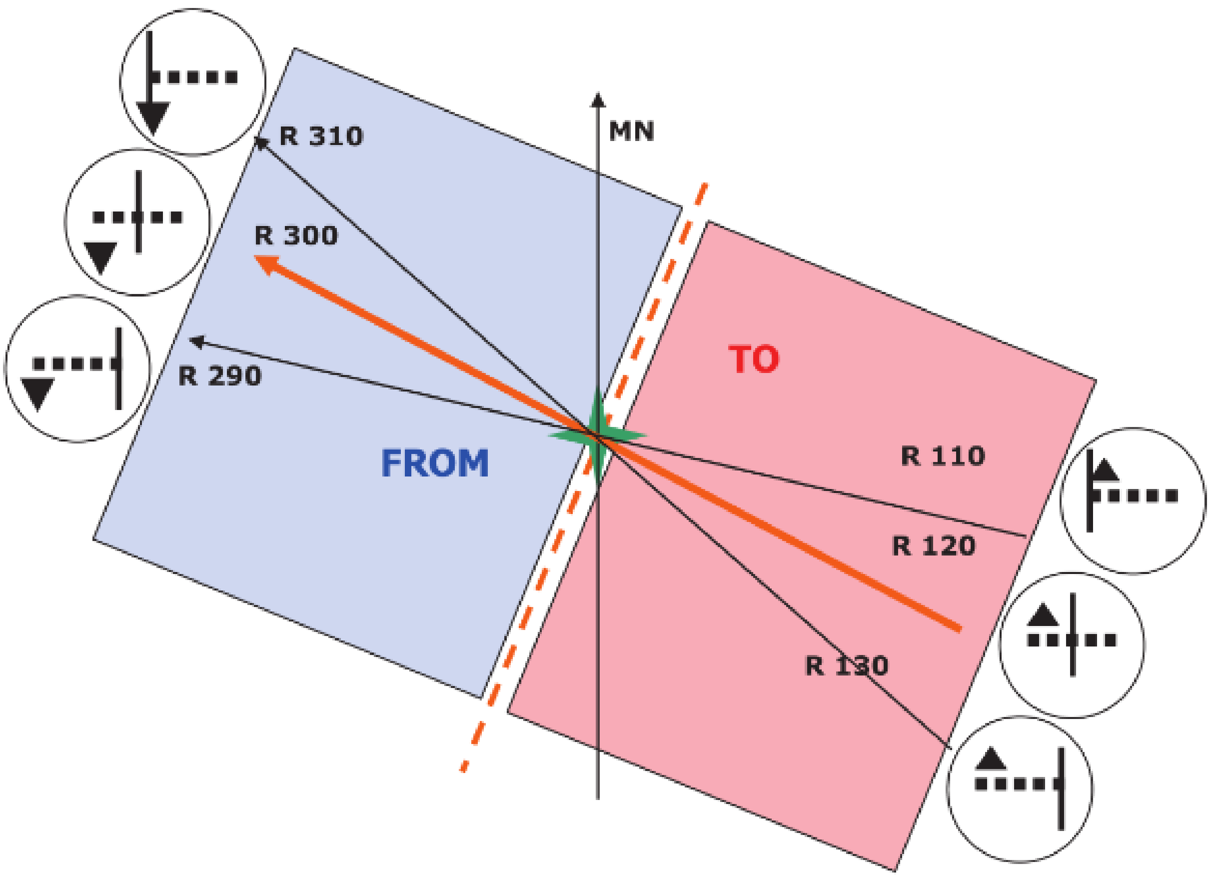
3.2.1 - Course Deviation Indicator (CDI)

The CDI (or OBI – Omni Bearing Indicator) is operated MANUALLY.
This instrument, by means of a vertical bar called course deviation indicator, displays information on deviation from a VOR track selected by an OBS (Omni Bearing Selector) or CRS (Course) knob.

The instrument is calibrated so that the needle reaches maximum deviation on the 2 dot graduation (or 5 dot for some instruments) with a phase deviation of 10°.
We will know if we are on the selected track, or on the right or the left of that track, up to a maximum value of 10°.
The window of the CDI is ±10° and 1 dot = 5°
A TO or FROM flag indicates to the pilot if, facing the direction of the selected track, he is in an area going to the station or moving away from it.
Caution the TO/FROM flag does not tell if the aircraft is moving to the station of away from it.
The flag can display TO while the aircraft is moving away from the station.

The OBS knob enables the pilot to choose a track (here, 300°). Caution, it is a track, this instrument does not display any heading.

The TO flag indicates only that the aircraft is flying in an area, where when looking at the direction of the selected track, goes towards the station, it does not indicate in what direction the aircraft is flying.



Indications according to the aircraft position with 300° selected by the OBS.

Illustration of a CDI with track 300 selected, the FROM flag is displayed, the needle is just at maximum position, 2 points on the right.

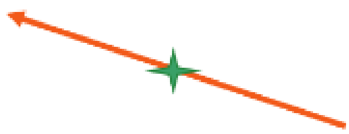
QUESTION:

What is the position of the aircraft in relation to the station?

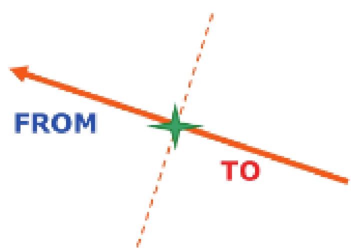
Method:

1) Materialize a station;

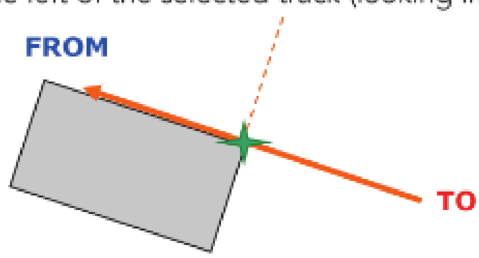
2) Draw track 300°;



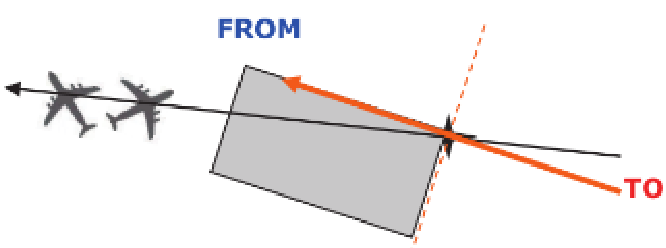
3) Draw a perpendicular track defining a TO area and a FROM area;



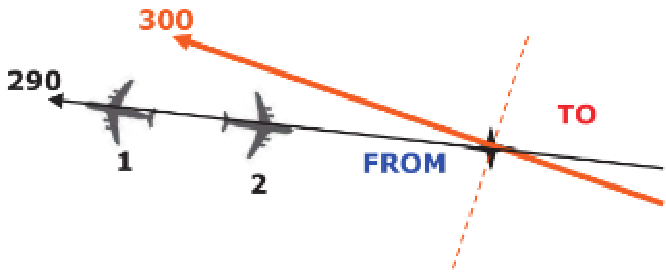
4) The FROM flag is displayed, therefore the aircraft is in a FROM area, the deviation needle is on the right so the aircraft is on the left of the selected track (looking in the direction of the track).



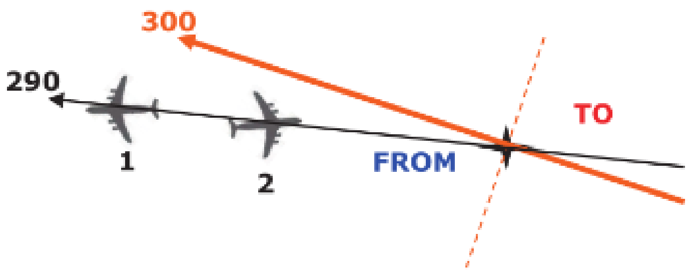
5) The aircraft is in the grey area 10° away from the selected track (300°) because the needle is just at maximum deviation, therefore on radial 290°.



CAUTION: the indication will be the same independently of the HEADING of the aircraft, this can be a source of error when looking only at the CDI.



Therefore, on the figure above, the indication is the same for aircraft 1 and 2.
If aircraft 1 says: « indicator on the right so I turn right in order to reach the selected track 300 », it will happen, if the pilot of aircraft 2, relying only on this indication takes the same action, the aircraft will only fly farther away.
The indication for aircraft 2 is said to be ANTI-DIRECTIONAL.
There is a simple way to figure out the correct heading to reach the track.
The aircraft must be imagined flying on its heading on the CDI.
This principle is the basis of the invention of the HSI explained below.



Aircraft 1 HDG 270°

This method is the origin of the development of the HSI.

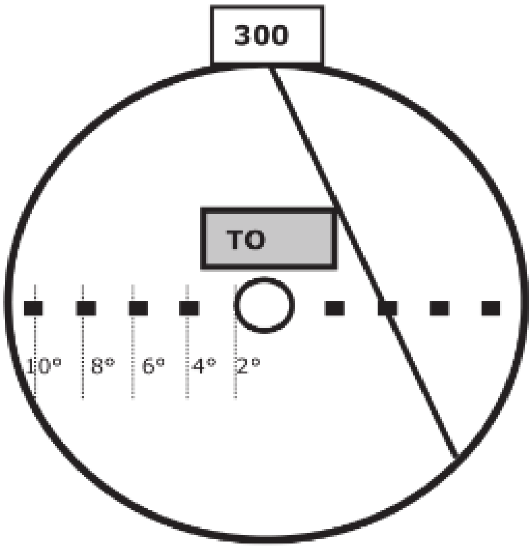
Aircraft 2 HDG 090°

Aircraft 1 visualizes that it must make a RIGHT turn, at a heading of say 340°, to intercept the track and aircraft 2 must make a LEFT turn heading 060°, for example, to intercept the same track.

Note that a needle with a two-dot deviation indicates that we are at least 10° from the selected track but without giving any information on which radial.
This means that the CDI displayed here can be for an aircraft on radial 290° but also 280°, 270°, 260° etc.
The CDIs shown so far are graduated 5° by 5°.

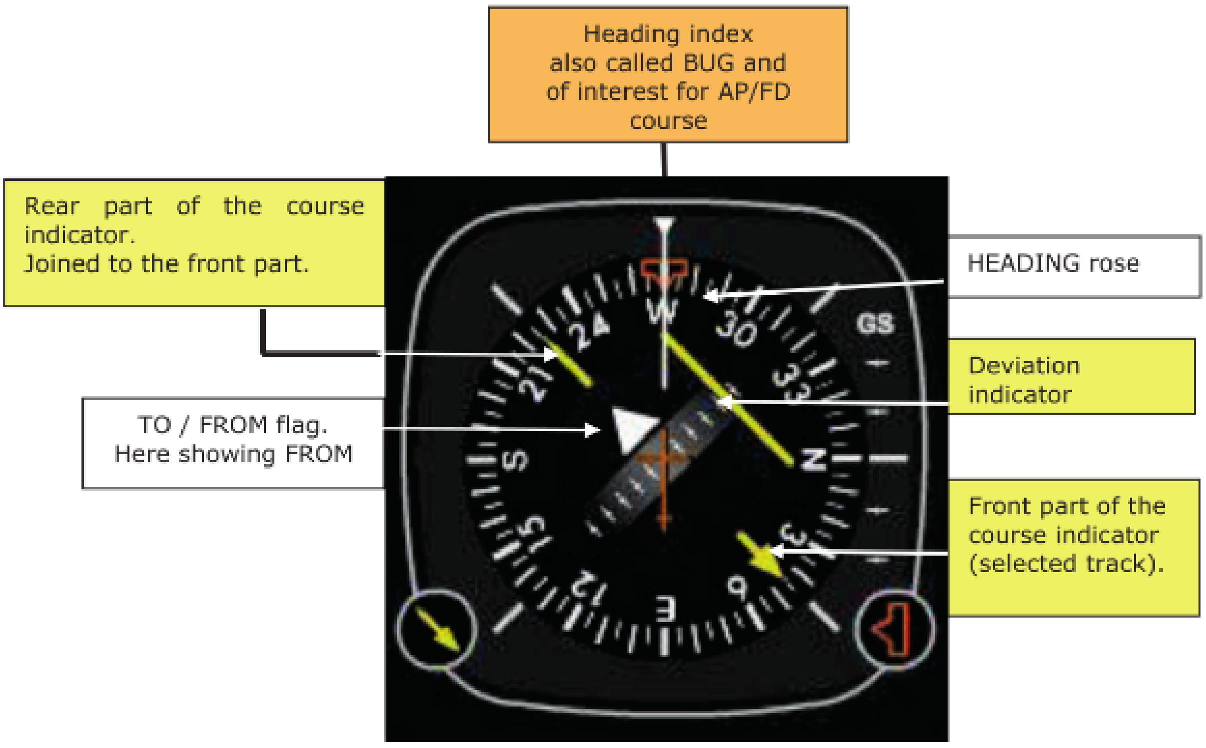
There are also 5-dot CDIs.
(1 dot equals approximately 2°).
They appear as on the opposite figure, with a scale made of a central circle of which the tangent is the first dot, with 4 dots on both sides, for a total of 5 dots.

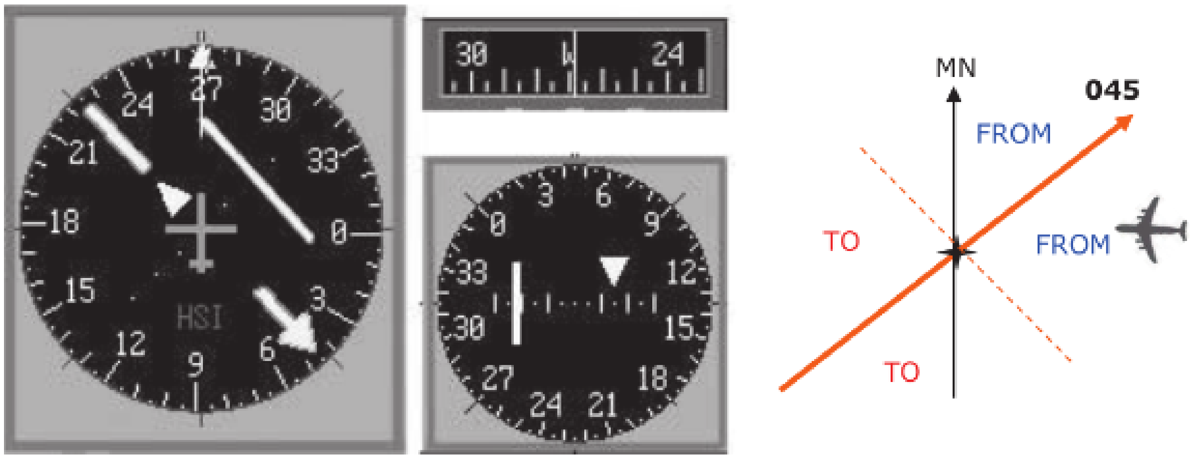
The window of the instrument is still ±10°.



3.2.2 - Horizontal Situation Indicator (HSI)

This instrument is also fed by a manual chain.
It is actually a CDI mounted on a compass rose (magnetic heading provided, either by a directional gyro on light aircraft, or the IRS/INS on an airliner).
It includes an aircraft model in its center which provides at all times a visual representation of the situation.





Example of an aircraft with heading 270°, flying FROM, track 045° selected.

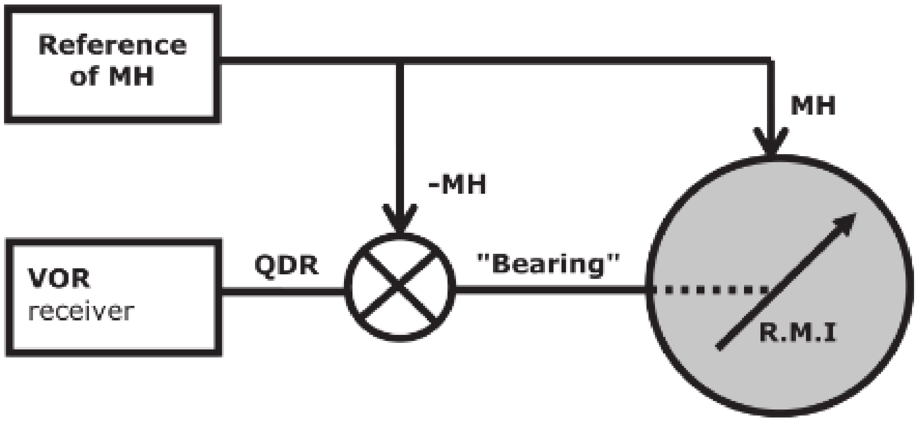
The HSI lets us see clearly that the aircraft has not yet crossed track 045° and will intercept it.
We can also see that (the aircraft model has to be moved towards the selected track), if we want to move away from the station on track 045°, we must fly a heading North, for example, and once the deviation indicator is centered on the selected course, a heading 045° (without wind).
Right of the HSI, there is a CDI showing the same situation but requiring mental efforts to calculate if the heading will allow interception of track 045°.

3.2.3 - Route Magnetic Indicator (RMI)

The RMI is activated by the automatic chain.
While the CDI does not give a usable radial indication if we are more than 10° from the OBS selected track, the RMI, by conception, provides the radial on which the aircraft is on at all times without any manual handling by the pilot.
It is an instrument that was seen in the ADF chapter, that provides a permanent display of the QDM.
With the VOR, the RMI provides with the position of the tail of the needle, the radial at all times.
At short distances and only if the magnetic variation is the same at the aircraft and at the station, we will read at the tail of the needle, the QDM.

The QDR is the magnetic track from the station on which the aircraft is.
To obtain the true track we use the magnetic variation at the station.

The principle of the RMI is to superimpose a needle indicating the radial at the tail of the needle on a moving rose connected to the heading of the aircraft.
As with the RMI ADF, the instrument makes a mechanical addition but the VOR provides directly a QDR contrary to the ADF which provides a bearing.
Yet, QDR = QDM ± 180° = Heading + Bearing and if we send this signal directly to the instrument needle it will perform the addition Heading + Bearing + Heading, that is two times the heading.
The RMI VOR needs a special connection which consists of removing the heading included in the QDR signal before feeding it to the needle to recreate ~~on the~~ instrument Heading + Bearing ± 180° = QDR.



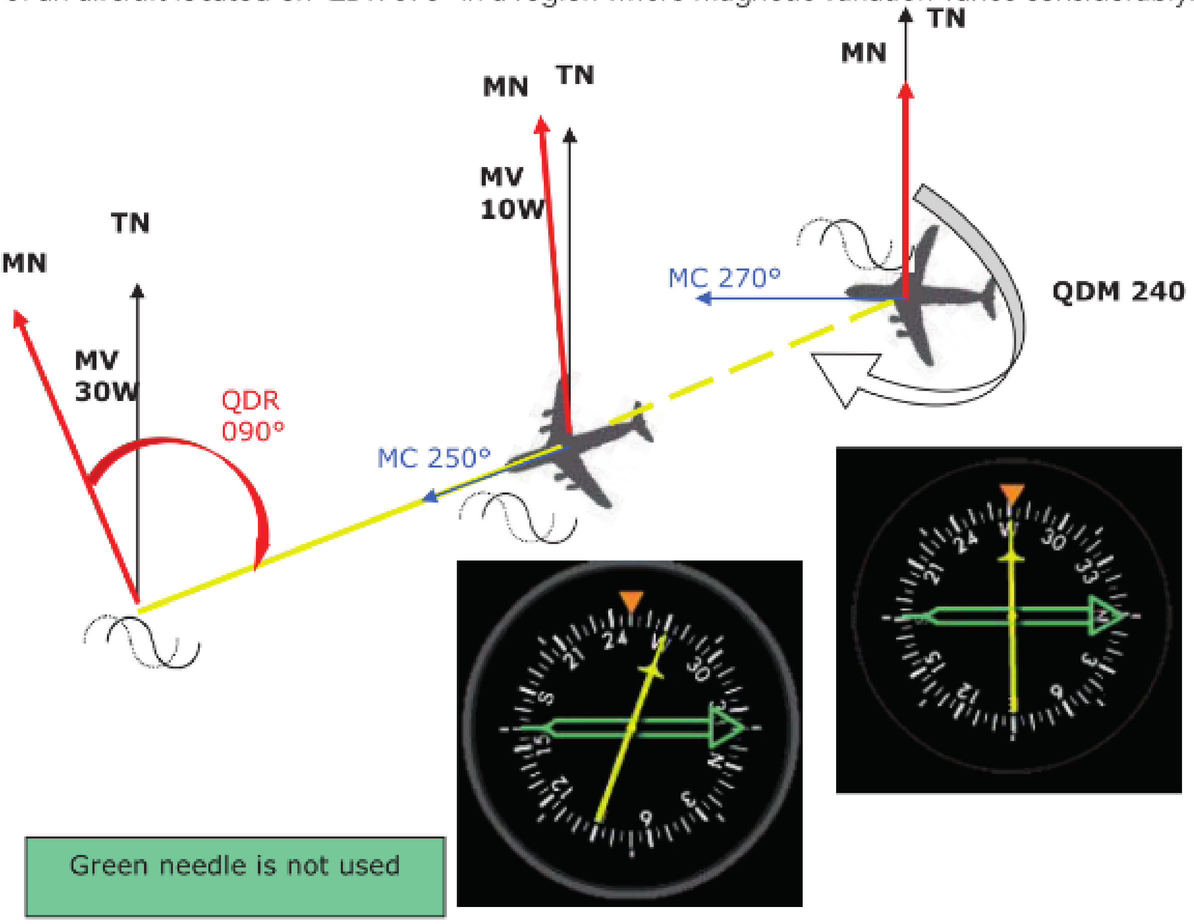
It can happen that the directional gyro of an aircraft is false.
In that case, the false heading is deducted from the QDR provided by the VOR, and it provides a FALSE BEARING, but this false bearing transferred to a rose with the same false heading will produce a CORRECT RADIAL (read at the tail of the needle).

To conclude this subject, we will say that:

On a VOR RMI, if the MH is false, the radial remains correct, but the bearing of the sation is false (the needle does not indicate the direction of the station).
On an ADF RMI it is the opposite, if the MH is false, the bearing remains correct, but the indicated QDM is false.

We note also that we have talked of a magnetic heading displayed on the rose (MH), but strictly speaking it is a compass heading, thus a magnetic heading corrected by a deviation that we have neglected.
If it is clear that the tail of a VOR RMI needle displays the QDR, the tip of the needle does not always indicate a QDM and does not always point to the station.
This will be the case only if, the magnetic variation is the same at the aircraft and at the station.

Case of an aircraft located on QDR 090° in a region where magnetic variation varies considerably:



Note that, for the two positions, the indicated radial 090 is correct but the bearing of the station is false.

This is a textbook case because most often, as a result of the latitudes where the VOR is used, and the fact that it is a medium to short distance navigation aid (VHF propagation), magnetic variations at the VOR and at the aircraft are mostly the same and thus the tip of the needle indicates the direction of the station (QDM).

In regions where our example is found, it is possible to set the VOR to the True North (reference and position signals in phase with True North) and fly with a rose showing to the True North.

We then find ourselves in the case where magnetic variations are equal, and the tip of the needle indicates the direction of the station and the tail indicates a "true" radial.

We find the equivalent of the instruments presented in the preceding chapters on the Navigation Display, called ND.

On the left, picture of an HSI on an EFIS that also has the function of ADF RMI and VOR RMI.
The green needle on the rose is the ADF and the blue needle is the VOR.
On this instrument, in case of disappearance of the VOR signals, there is no flag but the deviation indicator (magenta) will disappear.

On the right, a MAP mode with a VOR (green needle tip) and an ADF (blue needle tail) like on an RMI.

The other information, and the way to select radio aids or tracks to follow, displayed on these screens, will be developed in the RNAV chapter.

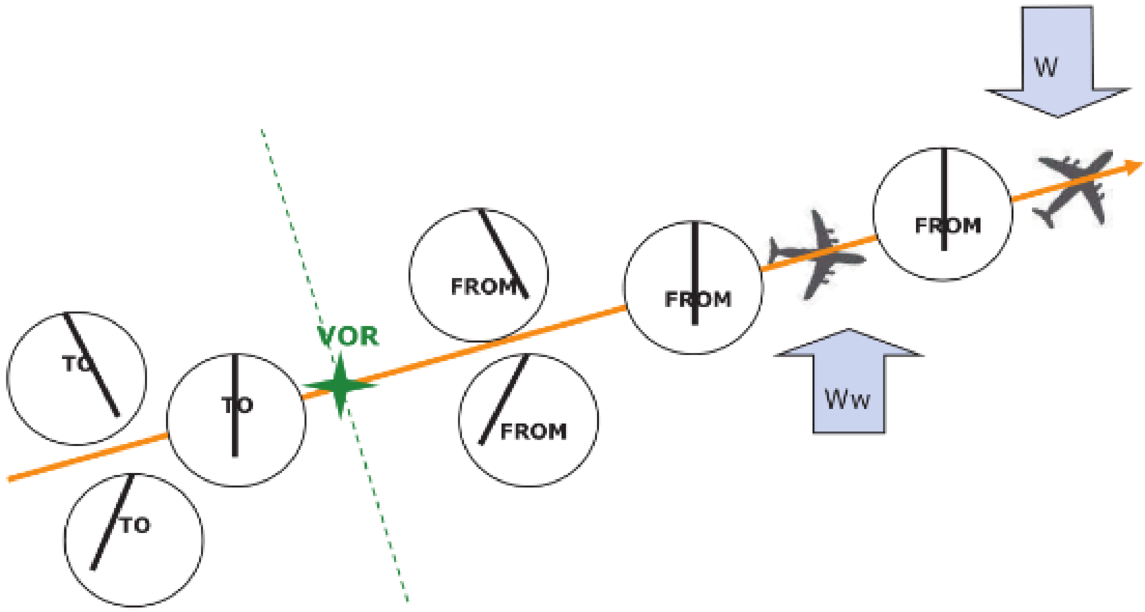
To close this instrument chapter, remember that we will go from magnetic bearing (QDR) given by the instrument to a true bearing by applying the magnetic variation at the station.

3.2.4 - Course Deviation Indicator (CDI) operation

Flying to catch-up the needle.

The pilot, depending on whether the aircraft is in TO or in FROM, must materialize his position in relation to the selected track and fly the appropriate heading to center the needle.
Once the needle is centered, he will fly the heading of the selected track (without wind).
Remember that while on track, the needle remains centered, irrespective of the heading of the aircraft.
If there is wind, (example given here flying away from the station), to stay on track the pilot must fly a heading different than the selected course.
The CDI does not allow materialization of the drift (see aircraft 1 and 2 on the schematic).

2



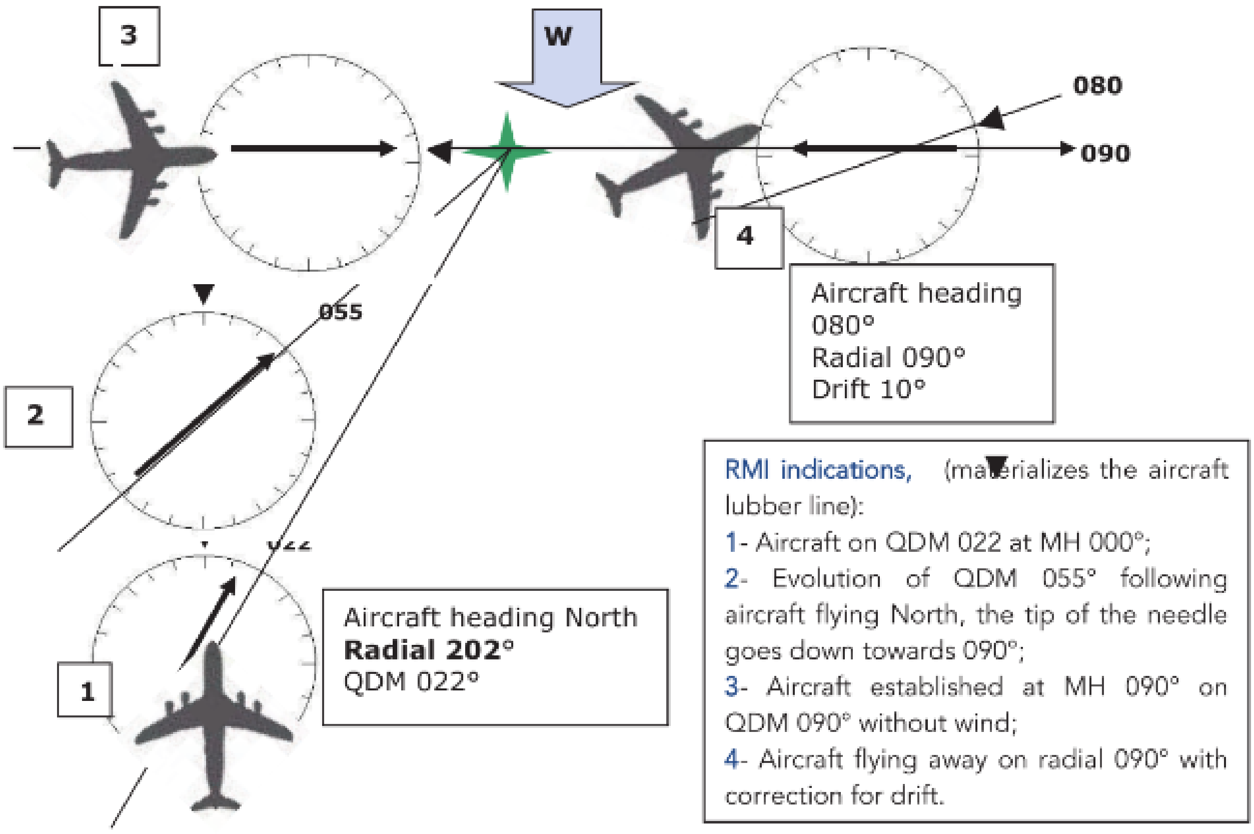
3.2.5 - Horizontal Situation Indicator (HSI) operation

The aircraft model is flown towards the desired track.
As with a CDI, when the course deviation is maximum (±10°), there is no information concerning the radial on which the aircraft is.
When on track (deviation indicator centered), the HSI allows materialization of the drift.

HSI with course 020° selected, TO flag showing.
The aircraft maintains track 020° with MH 045°.
The drift is 25° left.

3.2.6 - Route Magnetic Indicator (RMI) operation

The RMI displays the radial (and, in general, the QDM) on which the aircraft is at all the time.
There is, like for the ADF, flying over the station, an area called cone of confusion where the VOR receiver does not receive the signal properly.
In that area, the CDI needle beats from left to right, the TO-FROM flag toggles from TO to FROM or vice versa numerous times, and the RMI needle acts erratically not knowing if the station is in front or behind.
Approaching that area, the pilot flies a heading, keeping the average magnetic heading observed before entering the area.
Exiting the area, the pilot sees the deviation and makes the necessary heading corrections to return to the desired track.
At constant speed, the higher the aircraft flies and the greater this area is (radius is around 4 NM for an aircraft at 30 000 ft).



At the moment, the terms QDM or QDR are no longer used and are replaced by radial inbound or radial outbound.
For the aircraft in 3, ATC will request to follow radial 270 inbound.

3.3 – Range and coverage

3.3.1 - Range

The frequency band in use only allows quasi optical range.
Remember: If the two antennas, for the aircraft and the ground station, are at height H and h, the range D is:

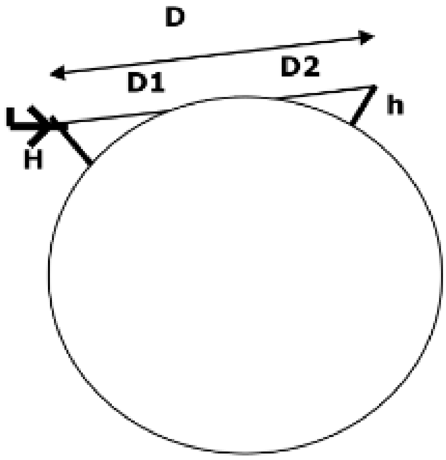
D = D1 + D2

$$D1 = 1{,}23(\sqrt{H})$$

$$D2 = 1{,}23(\sqrt{h})$$

$$D_{NM} = 1{,}23(\sqrt{H_{FT}} + \sqrt{h_{FT}})$$

H and h in Ft



Note:
Courses in English give 1,25 instead of 1,23.
It is possible that the answers to the practice questions vary slightly depending on their origin.

CAUTION $D = 1{,}23(\sqrt{H + h})$ **IS FALSE**

3.3.2 - Coverage

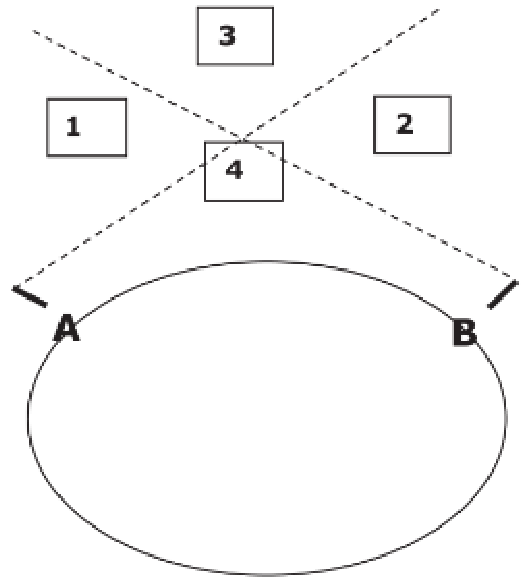
The coverage of the VOR even if it is omnidirectional depends on the terrain geographical conditions which cause signal absorptions, reflections, etc… especially at low altitude.

3.4 - Errors and accuracy

3.4.1 - Errors

Propagation errors:
Because of nearby obstacles, terrain irregularities, or high terrain, errors caused by reflected signals, or loss of signals can occur.
These errors can reach 6° at some places.
They are smaller if the station is a D-VOR.
These errors translate into curvatures of radials.

a) Interferences



An aircraft in area 1 receives VOR A.
An aircraft in area 2 receives VOR B.
An aircraft in area 3 receives A and B, there is a risk of interference and erroneous indications if A and B emit on the same frequency.
An aircraft in area 4 suffers a loss of signals from stations A and B.

VORs have a certified coverage area called DOC (Designated Operational Coverage).
It is the area where it is guaranteed that up to a certain Flight Level (thus for a given quasi optical horizon) there will be no other VOR on the same frequency (and no interference) until a distance of X Nm.

Example: a VOR is rated 80/28 000.
This means that there will be no interference from another VOR in a radius of 80 NM when flying at an altitude below 28 000 ft.

In general, two VORs with the same frequency are separated by approximately 500 Nm.

a) Instrument errors

Mainly caused by the mechanical parts of the servo and connection chains of conventional instruments, it is in the order of ± 3°.

b) Flying error

In manual flying, depending on the stability of the flying, the stability of the signals, and the distance to the station, there is an inaccuracy of around 2°.

3.4.2 - Accuracy

Taking into account all of these errors, the maximum error within the Designated Operational Coverage is ± 5°.
Because of the absence of interfering skywaves accuracy is guaranteed at day time and night time.

According to ICAO DOC 8168, a pilot flying a VOR approach is considered established if the approach path is maintained with a maximum of a half instrumental deviation (5°).

04 DISTANCE MEASURING EQUIPMENT (DME)

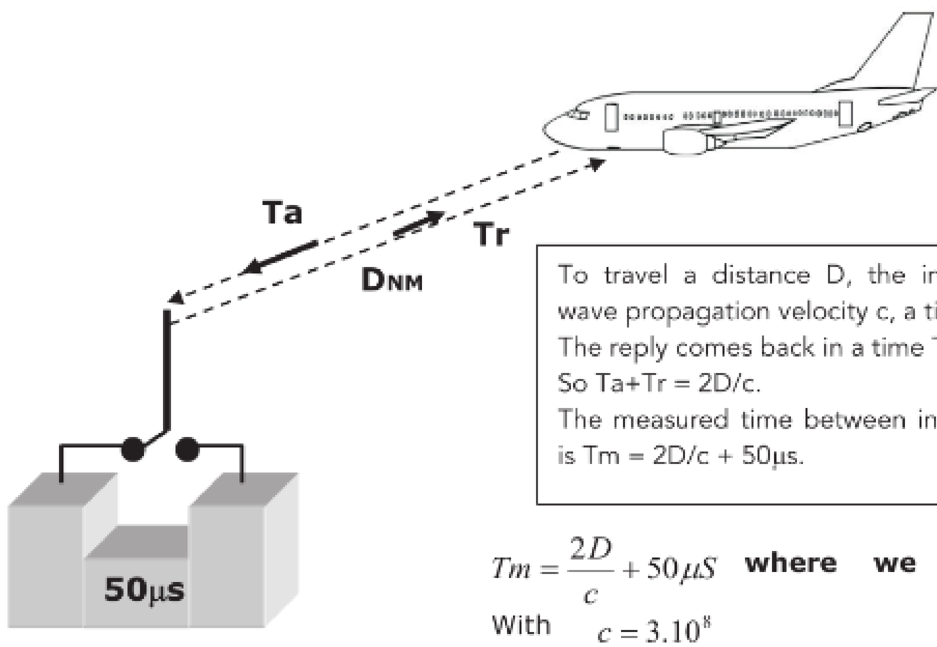
4.1 - Principles

4.1.1 - Generalities

The DME is a navigation aid, for short and medium distance, that supplies an indication of Slant Range to a ground station.
The information is presented onboard the aircraft digitally in Nautical Miles (NM).
The DME uses frequencies in the UHF band (quasi-optical range).
The frequencies are collocated with those of VOR, ILS and MLS systems.
The DME works with pulses.

- Ground station

The working principle is based on the measure of the propagation time of UHF pulse pairs between the aircraft and the ground station.
The aircraft possesses an interrogation transmitter that sends pulse pairs.
The pulses are received by the ground station then transmitted back, after amplification and a FIXED DELAY of 50μs, on a different frequency. This system allows the reception of pulses of a higher level compared to those based on the reception of echoes.
By measuring the elapsed time between the interrogation and the reply, we deduct the distance.
The measurement principle is based on a measure of time.
This technique of interrogation on one frequency, and reply on a different frequency, requires a ground station equipped with a receiver, an interrogation decoder and a trigger circuit transmitting the reply.
This ground **equipment is designated by the term transponder DME.**
Transponder is a system that we will find in the study of the inboard station of the secondary radar.
The DME uses, in terms of radio link, a technique similar to the secondary radar studied in 062 03 04 00.



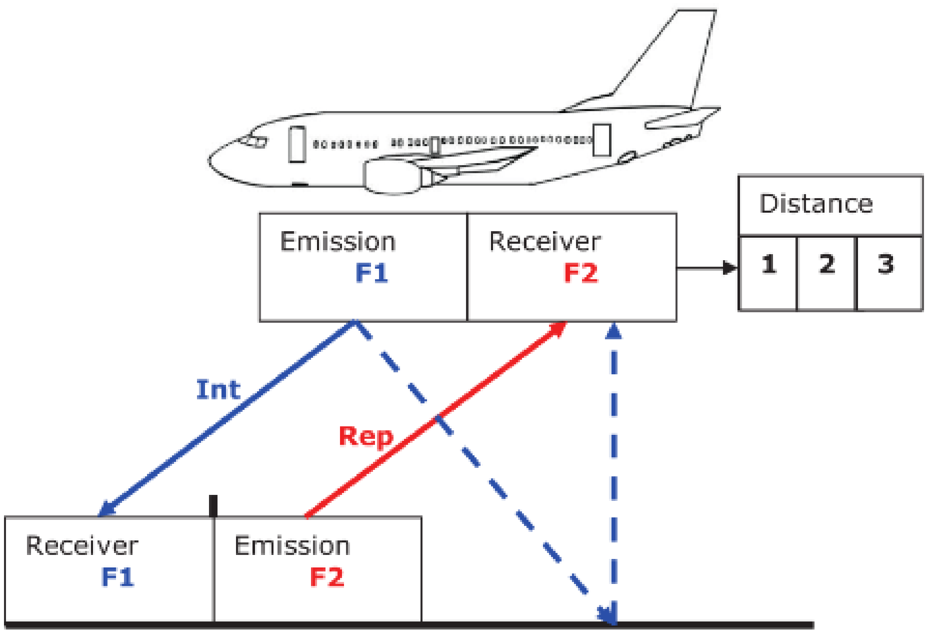
To travel a distance D, the interrogation takes, at wave propagation velocity c, a time Ta.
The reply comes back in a time Tr identical to Ta.
So Ta+Tr = 2D/c.
The measured time between interrogation and reply is Tm = 2D/c + 50μs.

$Tm = \frac{2D}{c} + 50\mu S$ **where we extract D** (in

With $c = 3.10^8$

If the DME must provide a distance in relation to a point that is not its ground location, the fixed delay is modified (1 500 m leads to Rf= -10µs).
Example: the ILS DME.

The DME uses the transponder technique, interrogations emitted on F1 trigger a reply by the ground station on F2.
The onboard receiver is tuned to F2 frequency.
This way, interrogations emitted on F1 and eventually returned to the aircraft as a ground echo cannot be received by the receiver and do not disturb the measurement.
The DME cannot lock itself on its own interrogations reflected by the ground because of the different frequencies used for interrogation and reply.



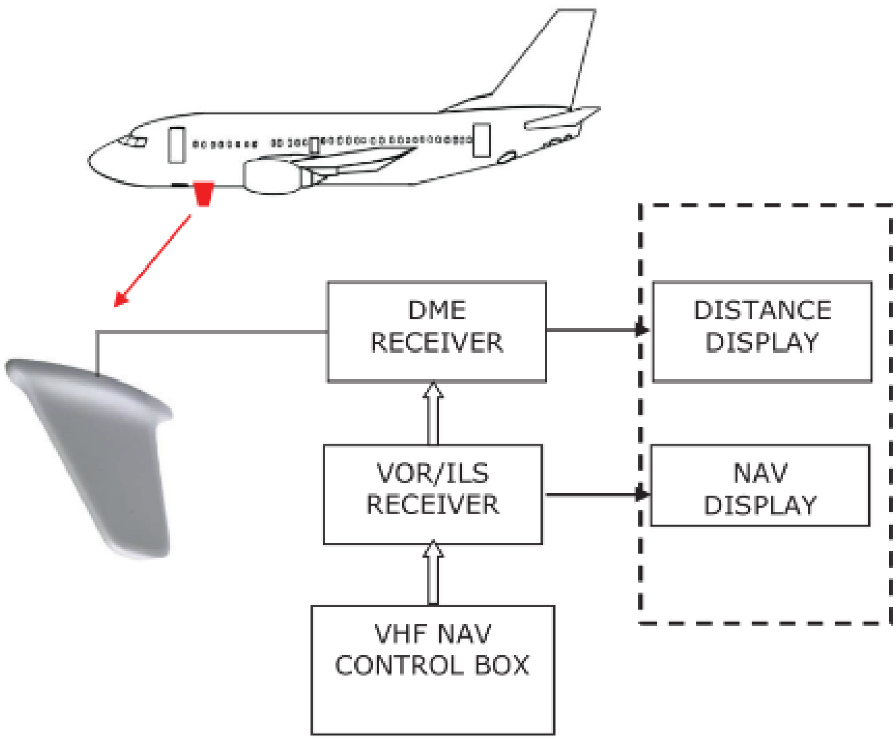
4.1.2 - Identification

The transponder must be able to transmit the identifier of the beacon.
The identifier is transmitted in A2A.
The modulation frequency of the identifier is 1 350 Hz.
When the DME is collocated with a VOR (or ILS) it has the same identifier as the VOR or the ILS.
(They are collocated if they are less than 600 m apart).
In that case the identifier is transmitted once every 40 seconds in synchronization with the associated VOR or ILS identifier, if it exists.
There will be, 3 VOR identifiers at 1 020 Hz followed by 1 DME identifier at 1 350 Hz, so at a higher pitch.

4.1.3 – Onboard station

Measurement principle

This is composed of an antenna (of small dimension in line with the wavelength in use) located under the belly of the aircraft, an UHF receiver and an analogic or digital distance display, isolated or grouped with the associated navigation indication (RMI, ND).
The receiver does not have a control box because the frequencies used by the DME are collocated with VOR or ILS frequencies.
The selection of a VOR or ILS frequency which is a VHF frequency automatically steers the DME receiver on the appropriate UHF frequency.



4.1.4 - TACAN

The TACAN (Tactical Air Navigation) is a system for military use that combines on the same carrier wave the functions of azimuth (VOR) and distance (DME).
Consequently, it provides the position of the aircraft as a radial/distance (Rho -Theta).
It operates in the same UHF band as the DME.
The principle of distance measurement being compatible with the DME, **the distance information provided by the TACAN is available for civilian aircraft.**

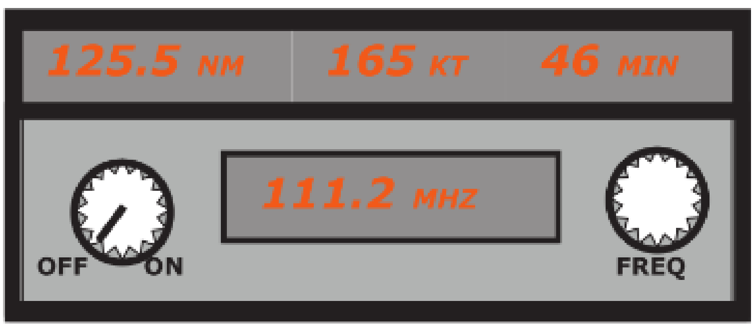
VORTAC symbol

4.2 - Presentation and interpretation

4.2.1- Presentation

Analogic DME meter with mechanical drum, integrated to the VOR RMI.

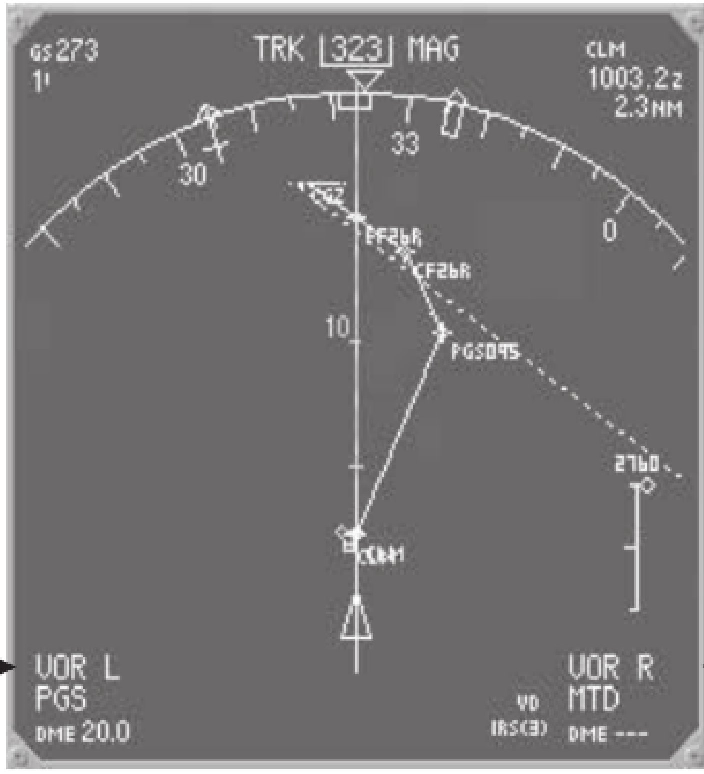
The green needle being associated to an ADF, the DME 1 meter is covered by a flag.



Indications integrated on the onboard system control box.
For light aircraft.
It provides distance and, ground speed based on the evolution of the measured slant range as well as the estimated time.

Display on a ND: the system decodes the identifier, displays it and displays the distance.
If there is no distance (VOR only), dashes appear instead of the distance.

VOR DME LEFT
Identifier PGS
Distance 20.0

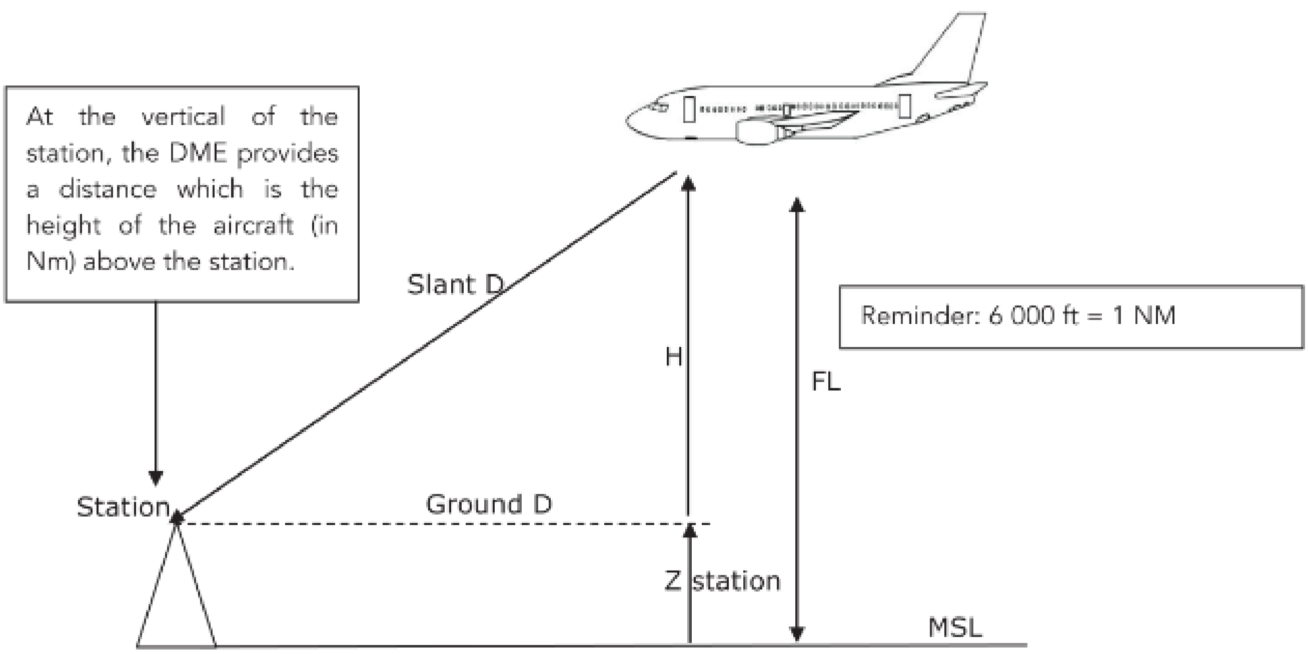


VOR RIGHT
Identifier MTD
No Distance

4.2.2 - Interpretation

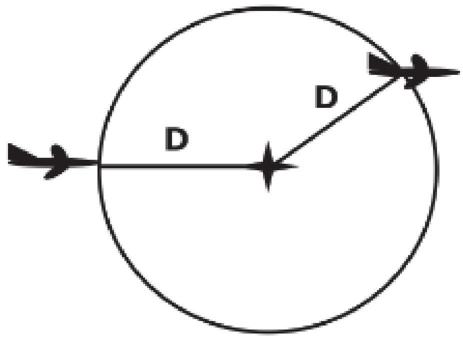
a) Measure of ground distance

The DME provides a slant range from which can be deducted by calculation (by applying the Pythagorean theorem) a ground distance.



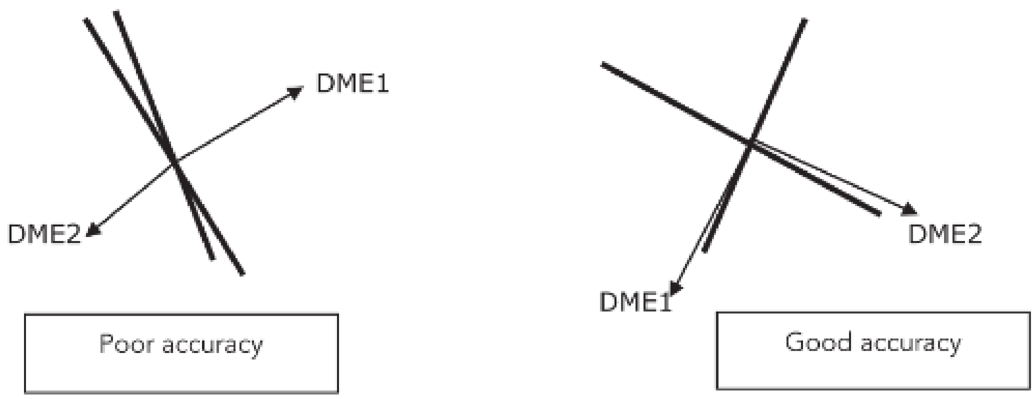
b) Position determination

The DME provides a distance information only, it does not provide a position but an area of position which is a circle of radius D centered on the ground station.



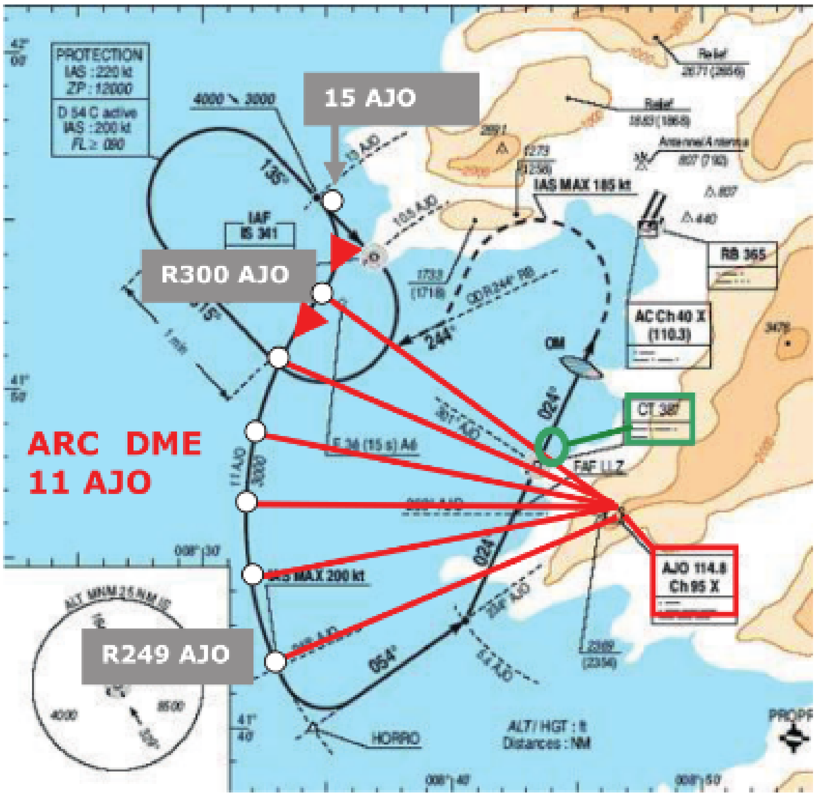
Very often, the DME is collocated with a VOR or an ILS.
In that case, the combination of the two systems provides **a Rho-Theta position** (radial-distance).

To obtain **a position based on DME information only,** it is necessary to use **two DMEs.**
The accuracy of the position will depend on the choice of the DMEs.
The accuracy is best when the areas of position provided by the stations intercept at right angles.
We will see in the RNAV chapter that the system often uses DME/DME to realign its position.



c) DME arc

Some approach procedures use as a transition path a DME arc, a path that is a circle arc of radius D based on a DME, leading to the interception of the final approach course.
While following this arc, obviously, the speed indication based on the evolution of the DME distance will indicate a ground speed equal to zero because the slant range does not vary.



The procedure starts at 15 AJO by a right turn that ends at 11 AJO on R300 at a perpendicular angle heading 210.
Heading is then reduced by 10° every 10° of radial (next radial 290, heading 200, etc.).

If the DME gives a ground speed indication based on the evolution of the DME distance, Vs indication is 0.
If the ground speed is not zero, it means that we are going out of the arc.

The arc ends on radial 249AJO with an interception turn to radial 234 (QDM 054) until intercepting the LOC of the ILS.

The arc is performed at a constant altitude of 3000' and made easier by flying a constant speed.

4.2.3 – DME capacity

The DME principle is to measure the time elapsed between the emission of a pair of interrogation pulses and the return of a pair of reply pulses.
When there are many aircraft interrogating the same ground beacon, a large number of pulses must be treated by the ground DME transponder.
The capacity of a ground DME transponder is limited by the capacity to treat the incoming pulses (station saturation for approximately *2700 impulsions per second).*

Statistical calculation demonstrates that a station can only reply simultaneously to a maximum of a hundred aircraft.

If the number of aircraft is too high, **the ground transponder lowers its sensitivity** in order to limit the reception to the maximum number of aircraft.

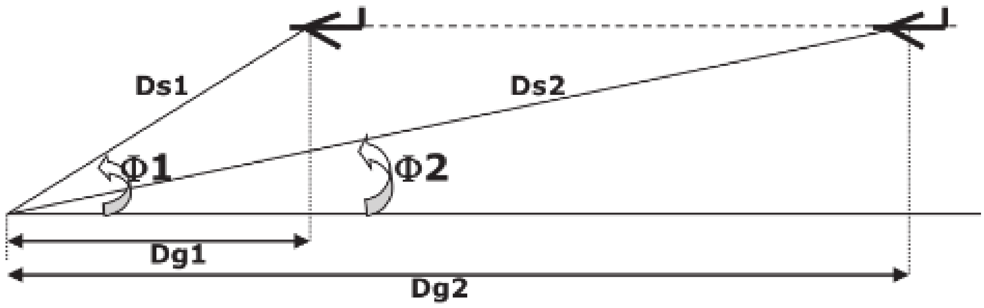
What is sensitivity?
If, in a classroom, all the students interrogate me at the same time, I more or less cover my ears.
I then hear only the questions that come to me loudly.
The transponder does the same thing, namely it increases the minimum level above which the signals become intelligible.
It then hears only those coming in stronger.

This amounts to replying only to the 100 aircraft that the ground station receives more strongly.
Caution this does not necessarily mean the aircraft that are closer.
An aircraft relatively close can find itself in conditions that reduce the signal received by the ground station, whereas a far away aircraft flying at high altitude in good radio connection conditions can deliver a stronger signal to the ground station.
This does not mean either that it replies to aircraft emitting with a greater power because they all emit with the same power.

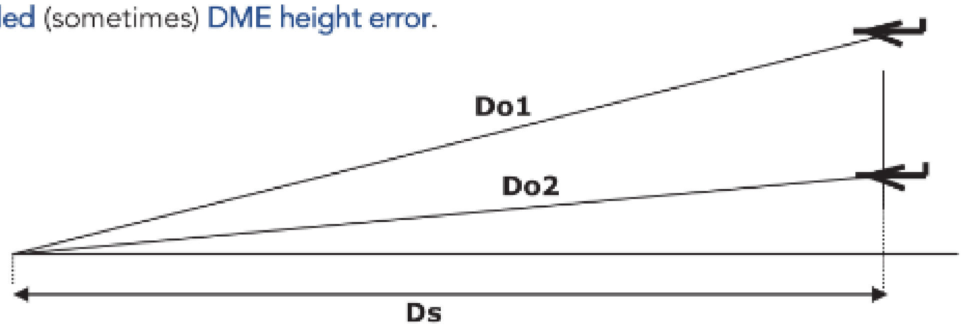
4.3 - Factors affecting range and accuracy

4.3.1 – On distance measurement

The displayed distance (slant distance noted Ds) is closest to the ground distance (Dg) when the aircraft is far from the station (compare Ds2 and Dg2 then Ds1 and Dg1).
The difference between Ds and Dg is negligible if the displayed distance is higher than the height in feet of the aircraft.
An aircraft is at 12 000 ft (2NM), the displayed Ds is representative of Dg when it is more than 12 NM.

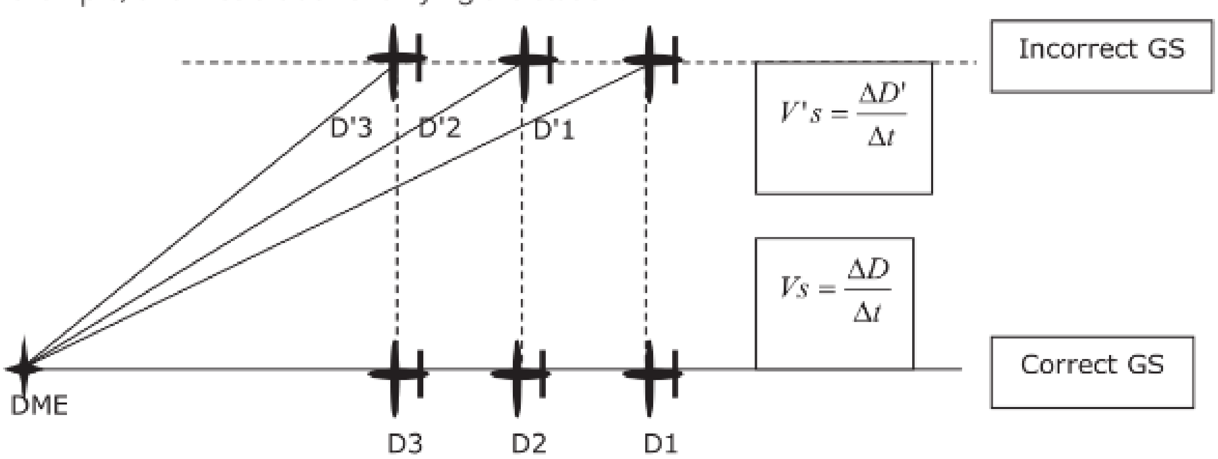


For a same ground position, the higher the aircraft, the greater the difference between Ds and Dg.
This is called (sometimes) DME height error.



4.3.2 – On ground speed measurement

When a Ground Speed indication (or a ground speed calculation) is based on the evolution of a DME distance, the information is closest to reality when the aircraft is far from the station (100 Nm for example) and flies a track overflying the station.



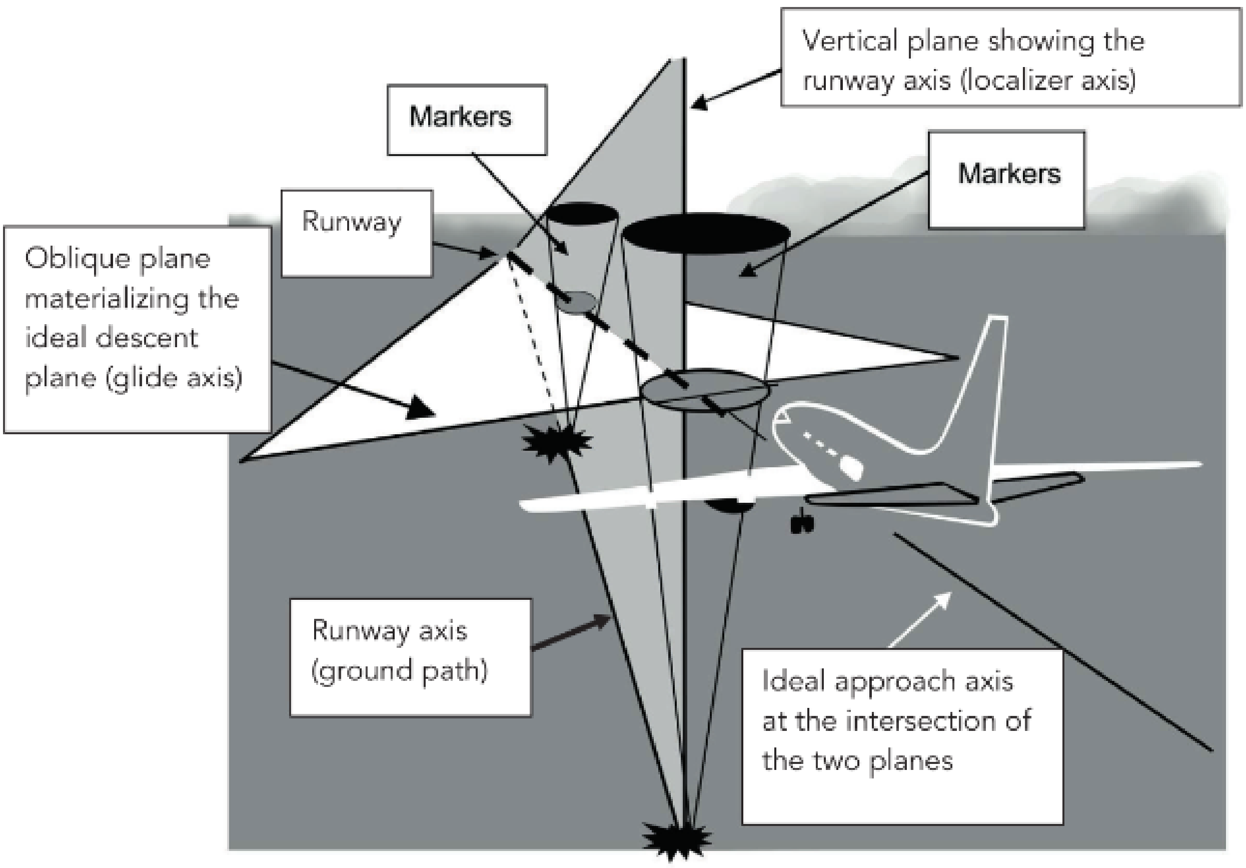
The closer we get to the station the lower the indicated ground speed becomes.
It is close to zero when flying over the station.

05 INSTRUMENT LANDING SYSTEM (ILS) - MARKERS CPL

5.1 - Principles

The ILS is a system based on two perpendicular radio beams that was conceived to provide horizontal and vertical guidance to aircraft during low visibility instrument approaches.
It gives the pilot the necessary information to maintain an ideal approach path.
When it is coupled to the autopilot and it meets certain requirements, it enables automatic landing.
It consists of several sub-systems which are:

- The LOCALIZER (LOC), providing the pilot a continuous information of lateral deviation in relation to a vertical plane passing through the runway center line.
- The GLIDE PATH (GP) providing the pilot a continuous information of vertical deviation in relation to a slant descent plane leading to the runway.
- The MARKERS providing discontinuous information of distance to the runway threshold.



More and more often, the ILS is coupled to a DME (Distance Measuring Equipment), which provides continuous information of distance to the runway threshold.

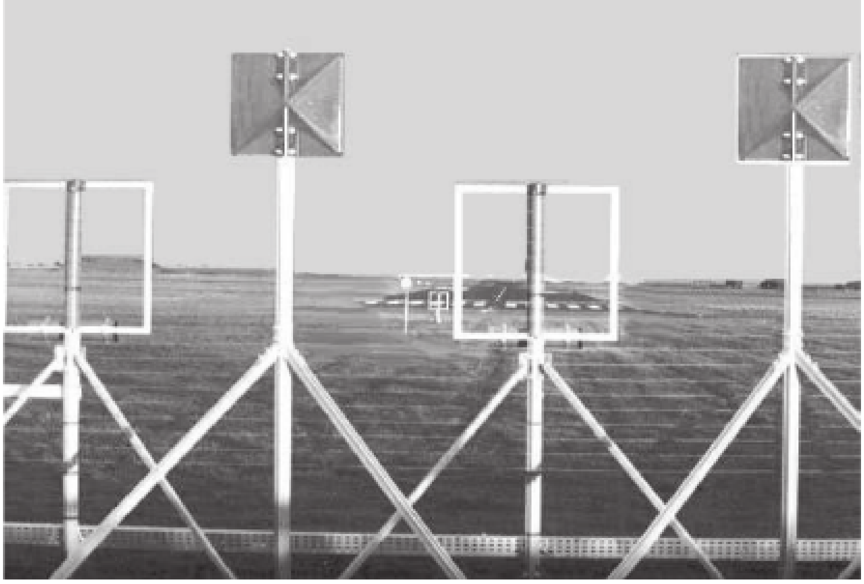
5.1.1 - Localizer

The Localizer (LOC) is a short distance navigation aid.
It provides a continuous deviation indication in relation to a vertical plane passing through the runway center line.
It operates in the metric frequency band.

The VHF frequencies in use go from 108 to 111,975 MHz, that is 40 channels.
To keep it simple, remember from 108 to 112.
This is the shared VOR/ILS band, the Localizer using the frequencies with an ODD number in the first decimal place (Ex: 108,1 – 109,3 MHz).
The information is displayed onboard on a deviation indicator. The indicator needle is centered when the aircraft is on the plane going through the runway center line.
When the aircraft comes in flying the runway QFU, the sense of the deviation indicates the sense of the maneuver to reach the center line.
Reminder: the QFU is the magnetic orientation of the runway.

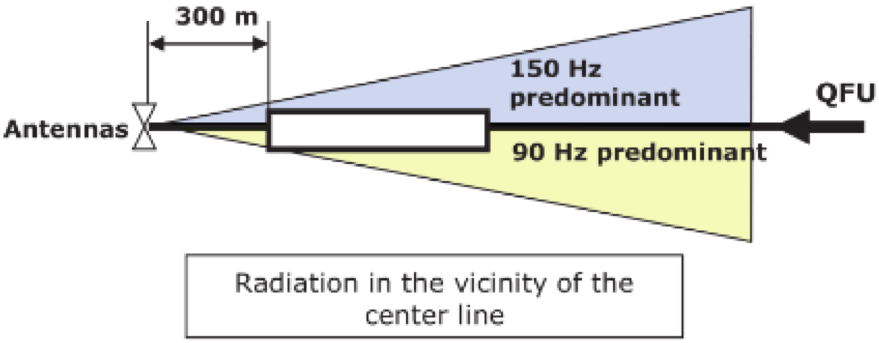
a) Ground station principle

An array of antennas on the ground is located approximately 300 m after the end of the runway, so that it is not an obstacle for aircraft taking off, and also to provide guidance information even after the aircraft has landed along the runway centerline.



Localizer antennas on the runway center line

This array of antennas radiates a VHF carrier modulated in amplitude by two Low Frequency (LF) frequencies of 90 and 150 Hz.
The radiation pattern consists of two lobes overlapping partially in the vicinity of the runway center line.
One of the lobes radiates the carrier modulated at 90 Hertz, and the other one, the carrier modulated at 150 Hertz.
The depth of modulation depends on the emission direction, which allows determination, in the vicinity of the runway center line, of two zones:
On the right of the center line, the depth of modulation of the 150 Hz signal is higher than the depth of modulation of the 90 Hz signal;
On the left of the center line, the depth of modulation of the 90 Hz signal is higher than the depth of modulation of the 150 Hz signal.
On the center line, the received depths of modulation are equal.

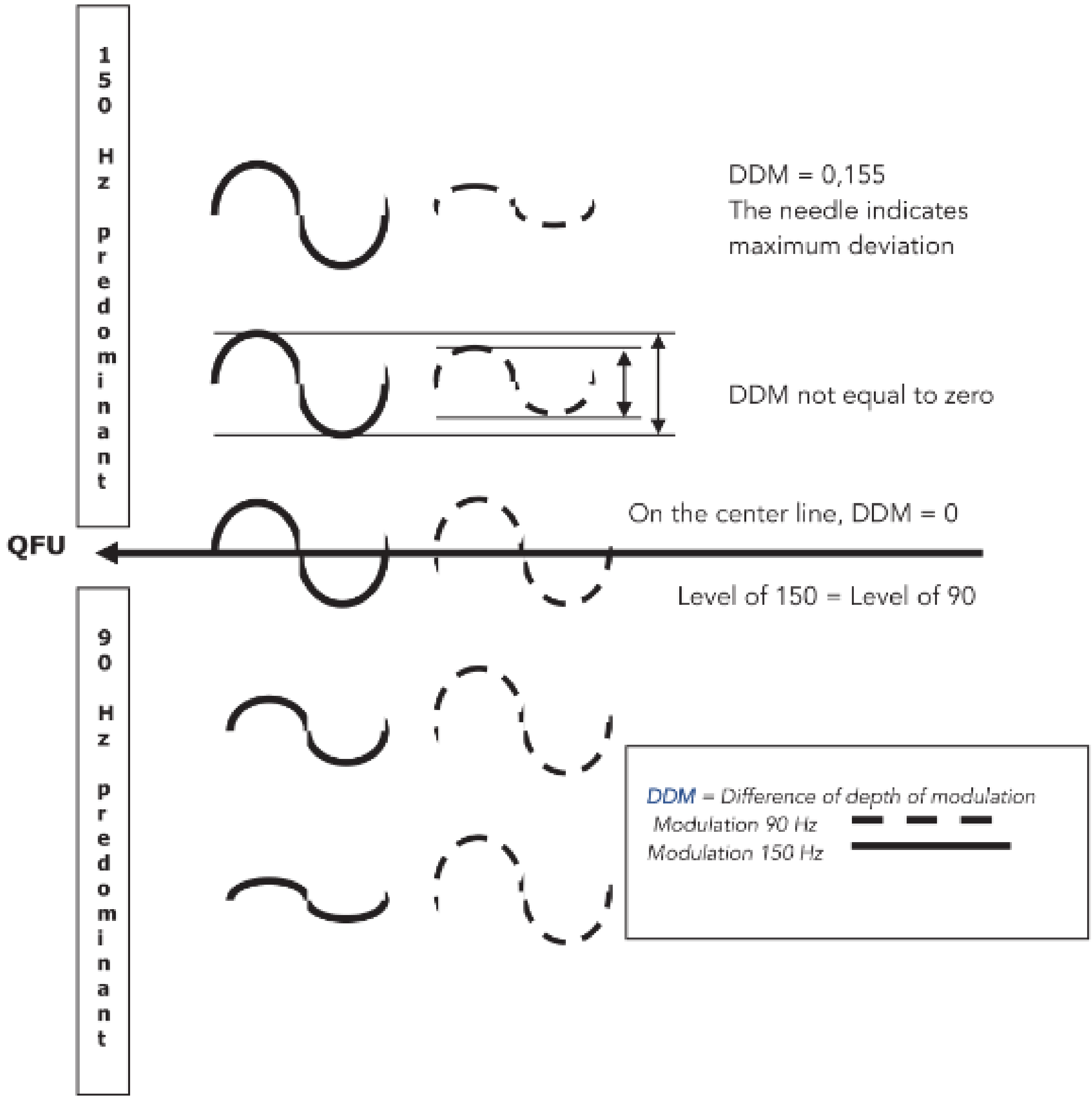


The right side of the center line is sometimes called blue sector, because on old ILS instruments the right part was painted blue.
The left part is called yellow sector for the same reason.

Some ILS radiate a back beam.
This back beam provides LOC guidance only, allowing non-precision approaches.
We must not use the back course if there is no published procedure.

The principle of the Localizer (but also of the glide path, seen below) is based on the measure of the difference of depth of modulation (noted DDM later in this course) of the 90 and 150 Hz signals.
A receiver onboard the aircraft analyses the difference of modulations at 90 and 150 Hz.
If that difference is not equal to zero, it elaborates a signal, which activates on an indicator a needle that gives the direction, but not the value, of the heading correction to make to reach the center line.
When the DDM is equal to zero, the indicator has its needle centered.
Keep in mind that the receiver receives, whatever its angular deviation in relation to the runway center line (in the limit of the alignment area), both 150 Hz and 90 Hz signals.
The (simplified) principle of the measure is based on the measure of the difference of level of the two frequencies (difference of depth of modulation noted DDM).

Indeed, a simple analysis of the received frequency, 150 Hz for example lead us to assume that the onboard receiver, and therefore the aircraft, is right of the runway center line, but it could not deduct information on the value of the deviation in relation to the center line.
However, if the 150 Hz signal is received more strongly, it can be deduced that the aircraft is on the right of the center line, and the difference of level between the 150 Hz and 90 Hz signals in that area, will inform us on the value of the deviation.



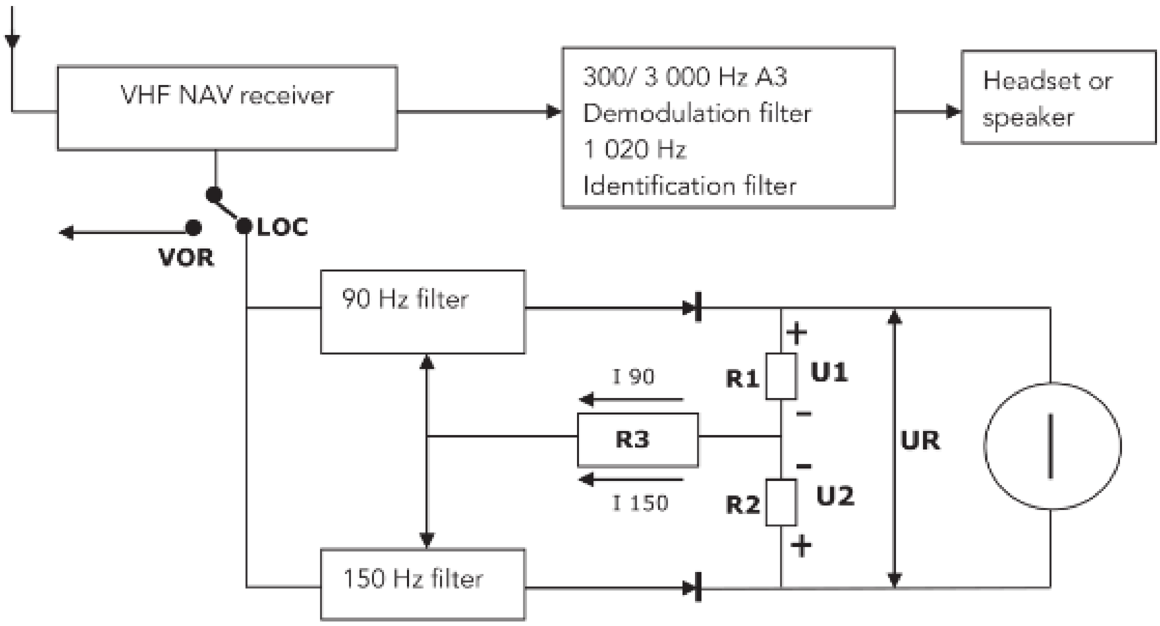
On the figure above, we see in a simplified way, the distribution of the depth of modulation in each area in the vicinity of the center line.
The 150 Hz (solid line) is predominant in the area on the right side of the center line, and we see that the DDM increases.
The needle of the indicator moves away from the central position more and more as we move away from the center line.
For the area on the left of the center line, the 90 Hz signal (dotted line) is predominant.

b) Onboard station

- Composition:
  - A localizer antenna located at the front of the aircraft;
  - A VHF receiver common to the VOR and the Localizer;
  - One or more indicators (CDI or HSI) developed in a chapter common to the localizer and the glide path, because the two information are presented simultaneously on those instruments.

- Principles

The VHF part is common to the VOR and the Localizer.
The VOR/LOC switches automatically according to the frequency selected by the pilot.
A filter common to both systems recovers the A3 modulations and the 1020 Hz used for identification.



The 90 Hz and 150 Hz signals are recovered through appropriate filters.
The 90 HZ and 150 Hz signal voltage rectified by diodes delivers a current (noted I 90 and I 150) which flowing respectively through R1 and R2, develops voltages noted V1 et V2.
Each of these voltages is proportional to the depth of modulation 90 and 150 Hz of the VHF carrier.
The assembly is done so that the voltages oppose each other.
The value of the resulting voltage VR is representative of the depth of modulation at 90 and 150 and will actuate the deviation indicator, displaying the value of the angular deviation.
The sign of VR is the one of the highest of the two voltage (90 or 150 Hz) and is representative of the side (right or left) of the beam where the aircraft is.
The course deviation indicator will drift on the right or on the left depending on the sign of VR.
The indicator is a micro-ampere meter with zero central presenting a full deviation for 150 microamperes equivalent to a DDM of 0,155.

- Integrity monitoring:

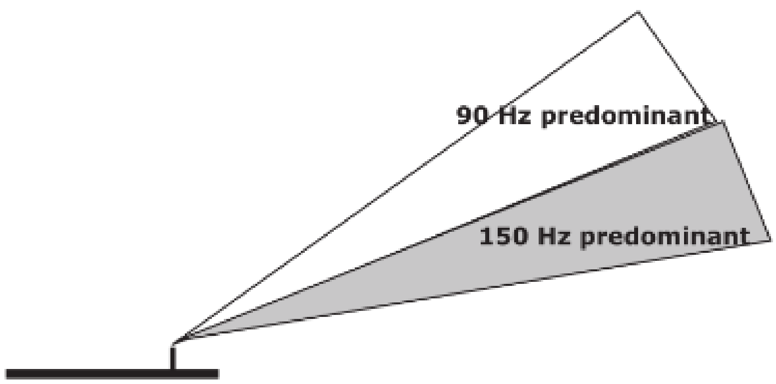
Figure above:
I_{90} et I_{150} currents flow through a common resistance R3 and, by Ohm's law, create a voltage V3 = R3 (I 90 + I 150).
This V3 voltage activates a LOC warning flag if it becomes inferior to a predetermined limit.
(The lower limit is reached if the sum of the modulations becomes inferior to 40%).
Any disappearance of I_{90} or I_{150}, or abnormal weakening of the signals, will make V3 drop to the alarm level, warning the pilot of a problem of information integrity.

5.1.2 - Glide Path

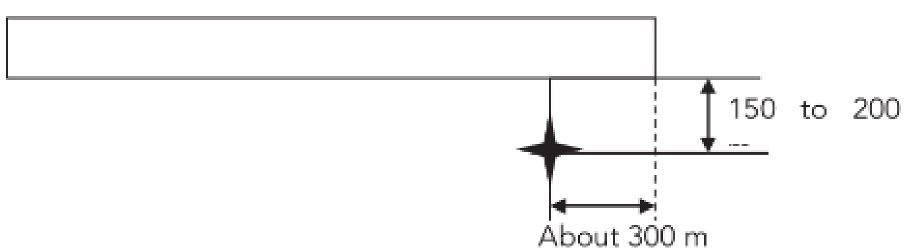
The Glide Path (or GP) is a short range navigation aid.
It provides a continuous deviation indication in relation to a slant plane that materializes the descent path.
It uses UHF frequencies.
The GP frequencies are collocated with the localizer frequencies.
The information is presented onboard on a deviation indicator; the needle of the indicator is centered when the aircraft is on the nominal descent path.

a) Ground station principle

This is identical to the localizer, namely it is based on a measure of differences of depth of modulation of a UHF carrier, modulated by signals at 90 and 150 Hz, and radiated following two beams that overlap partially in the vicinity of the ideal descent path.
The information is presented on the same indicator as the localizer by activation of a second needle.
The difference resides in the radiation of the beams, which is done in the vertical plane, the location of the antennas, and the fact that the 90 Hz modulation is predominant above the path, while the 150 Hz modulation is predominant below the ideal descent path, which is normally around 3°.



The antennas can, of course, create an obstacle on the runway center line.
As a consequence, they are, in general, located at about 300 m from the threshold and at about 150 to 200 meters of one side of the runway.
It is set to have the beam pass over the runway threshold at 50 FT.



In fact, from 240 meters to 480 meters

According to ICAO Annex 10, the glide path is generally set at 3° (5%).
However, in certain cases the path can be raised to guarantee obstacle clearance on final approach (ex: Marseille QFU 31, slope 7%).

c) Onboard station

- Composition:
  - A UHF receiver;
  - An antenna located in the nose of the aircraft.
  - On some aircraft, there are two antennas because the landing gear configuration influences the reception. Depending on the configuration one antenna or the other is automatically operated.
  - Common indicators for the LOC and the GP will be developed later.

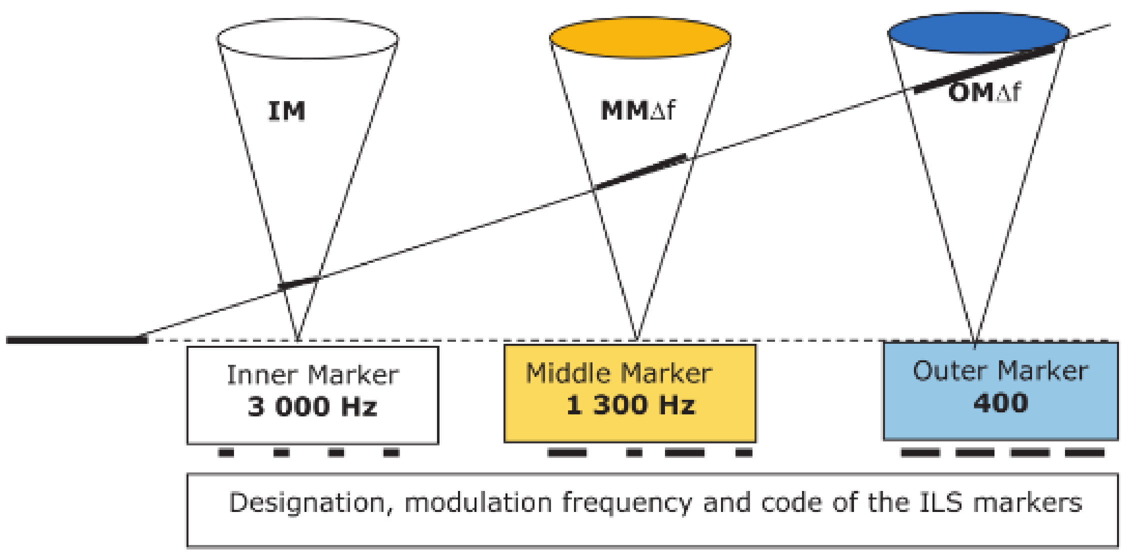
There is no control box specific to the GP.
Displaying the Localizer frequency on the VHF NAV control box is sufficient.
A circuit automatically controls the Glide Path receiver that will tune in to the associated frequency.
Example: for a Localizer frequency of 108.30 MHz there is a corresponding Glide Path frequency of 334.10 MHz.
The principles of the receiver and of the circuits feeding the indicator, as well as the integrity monitoring with potential display of a GS flag, are identical to those studied for the Localizer receiver.

5.1.3 - Markers

They are transmitters located on the ground that radiate vertically on a 75 MHz VHF frequency.
As a result of the vertical radiation, even though they use the same frequency, they do not interfere with each other.
Markers are received only when they are overflown by the aircraft, and thus do not provide any directional indication of navigation.
They are used either together with an ILS to provide a discontinuous distance indication or en-route as airway beacons.
There are two types of markers and their radiation diagram is slightly different depending on their application.

a) ILS markers

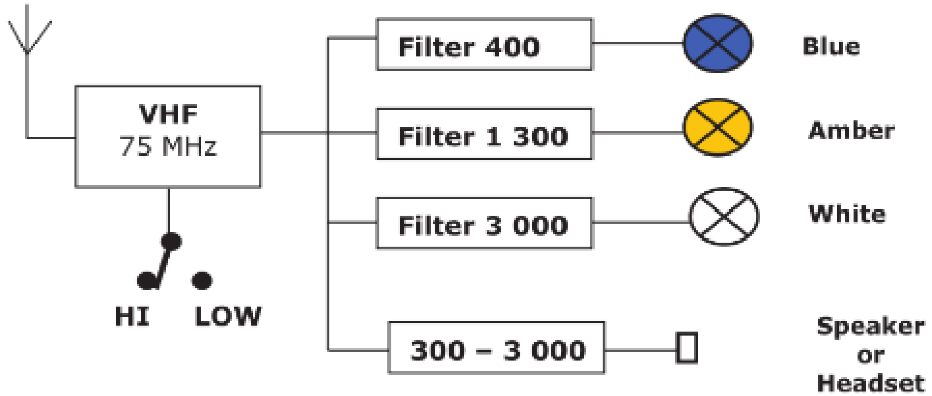
These markers are used with the ILS.
The radiation is, in general, elliptical.
An ILS includes two or three markers transmitting with a power of 3 to 5 Watts.
Each marker transmits on VHF (75 MHz) but is modulated by a different Low Frequency that is cut out to the rhythm of the Morse code (A2A modulation type).
A different coding is attributed to each marker.



- The outer marker (OM) is located about 4 NM from threshold.
- The middle marker (MM) is located about 0.57 NM from threshold.
- The inner marker (IM) is located about 300 meters from threshold (0,16 NM).

b) Markers receiver

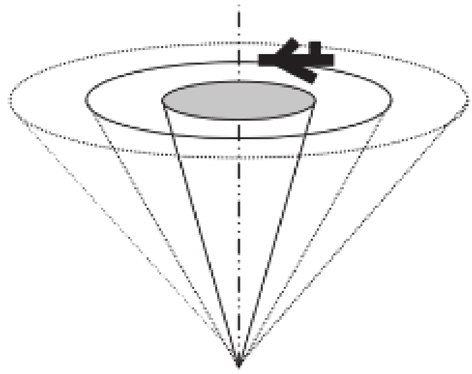
The receiver does not have an ON/OFF switch.
A circuit receives the 75 MHz VHF signal modulated in A2, and after recovery of the LF, filters steer the LF signals (400, 1300 or 3000 Hz) towards lights of different colors that symbolize the OM, MM, or IM if there is one.
The modulations are also sent to the radio listening system.
Markers provide visual and audio indications.



The emission diagram of the markers being a cone, the reception time, and the time that the light is on and the tone is heard, will be longer if the aircraft is high above the marker, or if at a same altitude it is flying more slowly.

c) HI – LOW control

This control allows the pilot to modify the sensitivity of the receiver.
As the name indicates, HI (High) has the higher sensitivity and LOW has the lower one.
As the aircraft moves away from the center of the beacon emitted cone, the field decreases.
By pressing the HIGH key, we increase the sensitivity of the marker receiver.
This will enable us to receive a lower signal, with the consequence of illuminating the marker lights even if we are not in the schematic grey area, which represents the area of high field that is received when on LOW.



In low visibility approach, the HI key will be pressed, at first, to increase the chances of receiving the marker signal.

G1000 –Control available on Audio Panel
MKR-MUTE allows shutting off the audio indication.
HI SENS choice of sensitivity of the receiver.

5.2 - Instrumentation and interpretation

5.2.1 - Instrumentation CDI – HSI

LOC and GP information is available on indicators studied in the VOR chapter.
That is the CDI, the HSI or a Navigation Display (EFIS).
The information is also displayed on the ADI (or the PFD for EFIS aircraft).

Even though there is practically no difference in the physical presentation on instruments such as the CDI or the HSI, there are important differences in the interpretation and the utilization of the provided information.

- The first difference relates to the instrumental window:
    - the LOC needle, (the same one that is used for the VOR), is at maximum deviation for an angular deviation of about 2,5° with respect to the center line;
    - the GP needle, or GP index of the HSI, is at maximum deviation for an angular deviation of 0,7° with respect to the glide path.

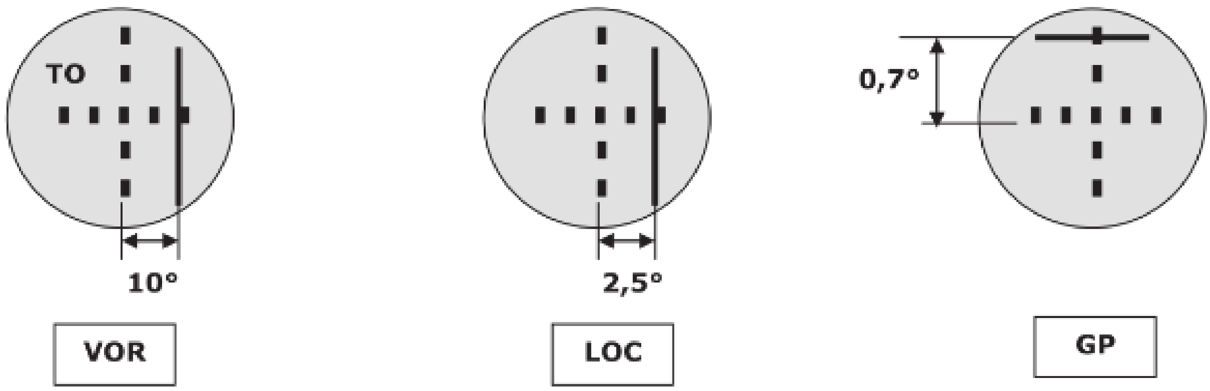
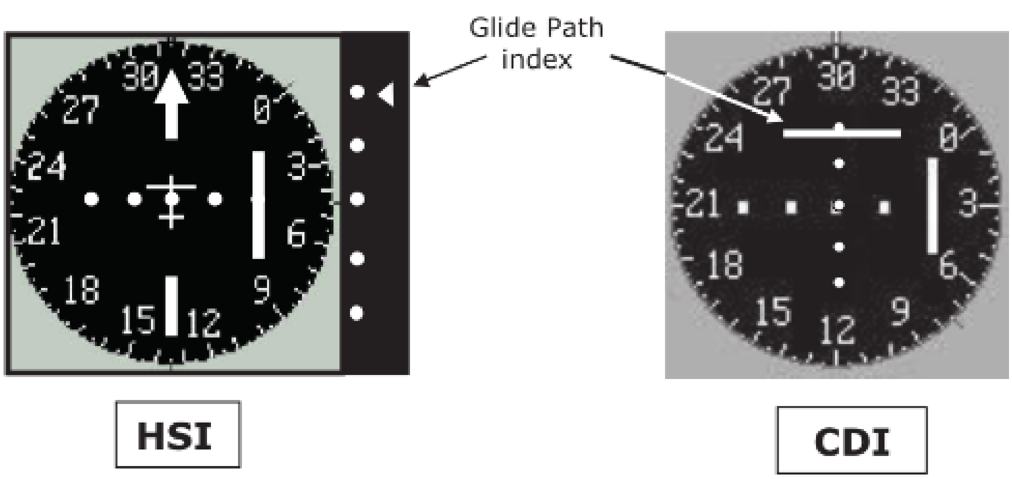
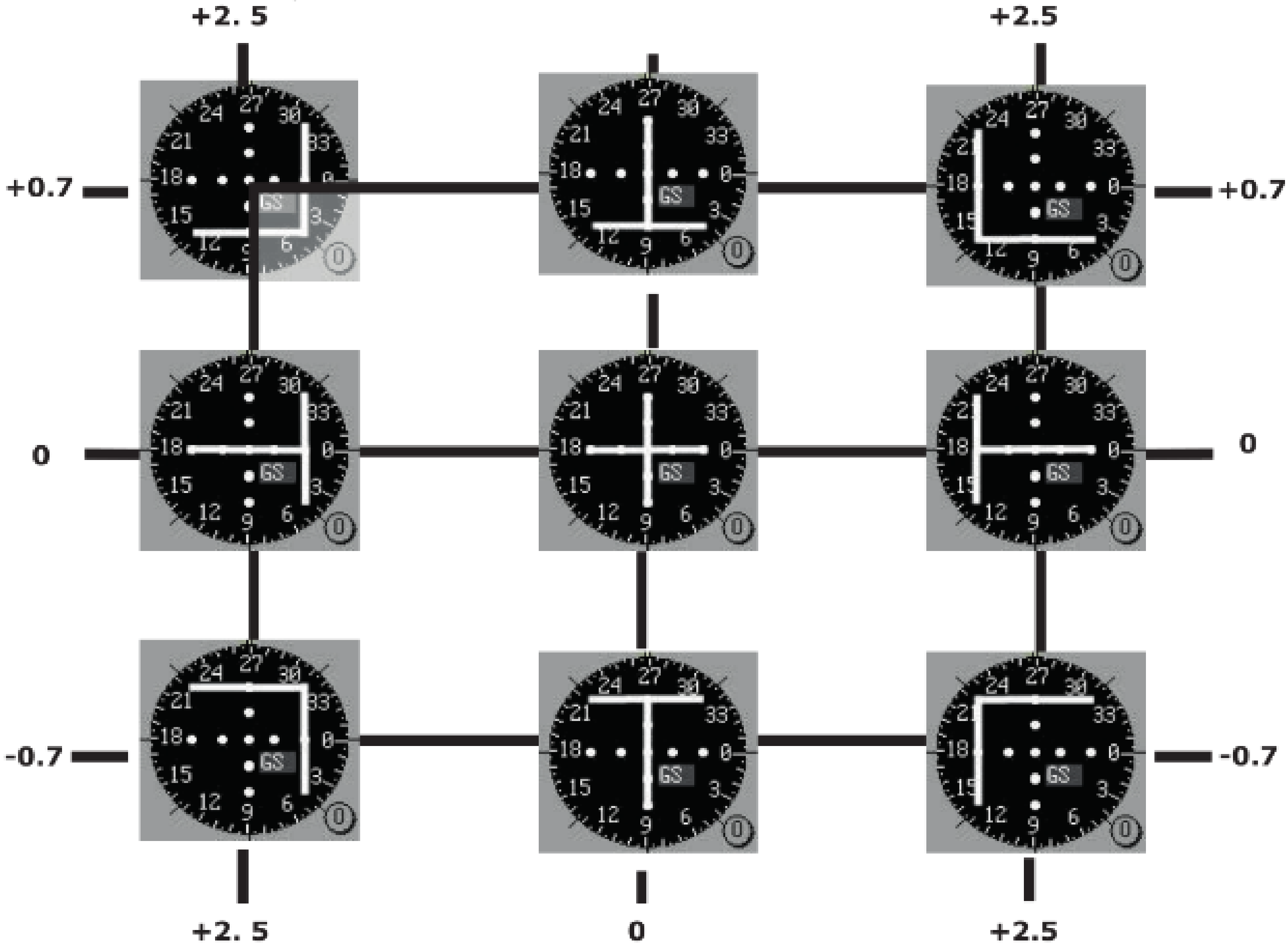


Illustration for an aircraft at 2,5° left of the localizer and 0,7° under the glide path.



Regarding the Glide Path, the HSI index or the course deviation indicator on the CDI, materialize the nominal descent path of the ILS.

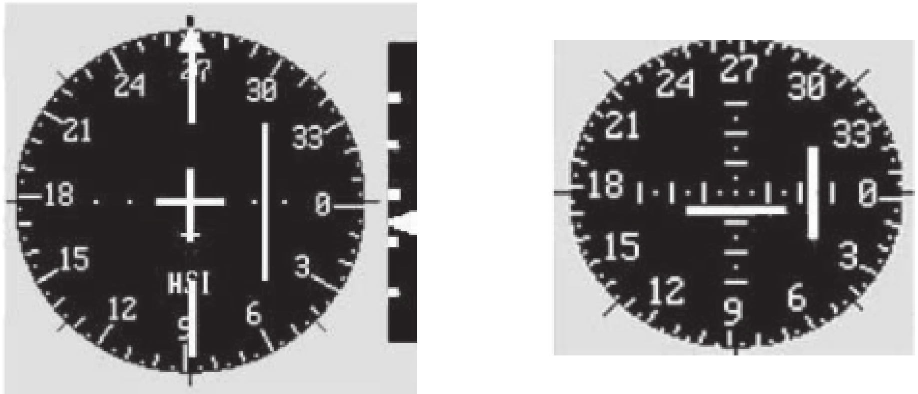


Instrumental window of localizer and glide path

- The second difference relates to the selected course:
  - The course is selected with the OBS knob which orients either the compass card of the CDI or the platter of the HSI;
  - it does not influence, contrary to a VOR utilization, the sense of the course deviation indicator.
  - Whatever the selected course, the aircraft always receives the same depth of modulation at 150 or at 90 Hz and the needle always gives the same indication whether it is on the right or on the left of the center line.

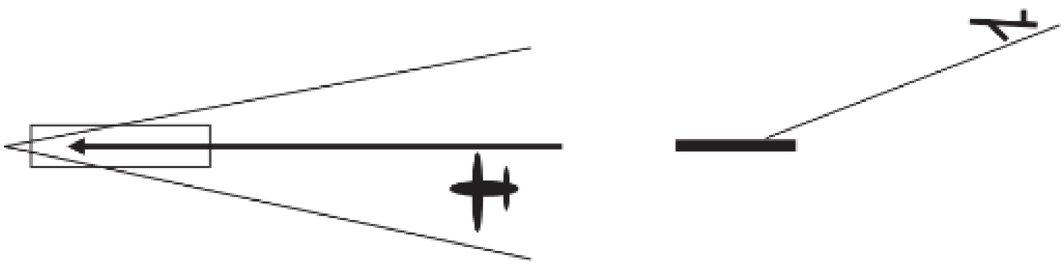
The orientation of the compass card of the CDI or the platter of the HSI in the direction of the QFU of the ILS in service only helps provide a better mental situational awareness.

There is only one way to center the course deviation indicator: MOVE THE AIRCRAFT to bring it to the center line.

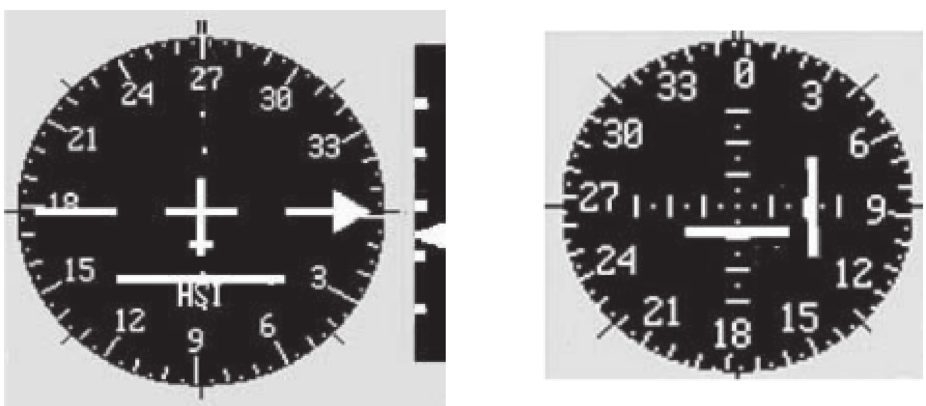


Aircraft approaching ILS 27, flying on the left of the center line, slightly high on the glide path.

The platter, and the compass card of the CDI, oriented at 270, provide a good picture of the situation.
It is easy to materialize that the heading has to be increased to reach the center line.



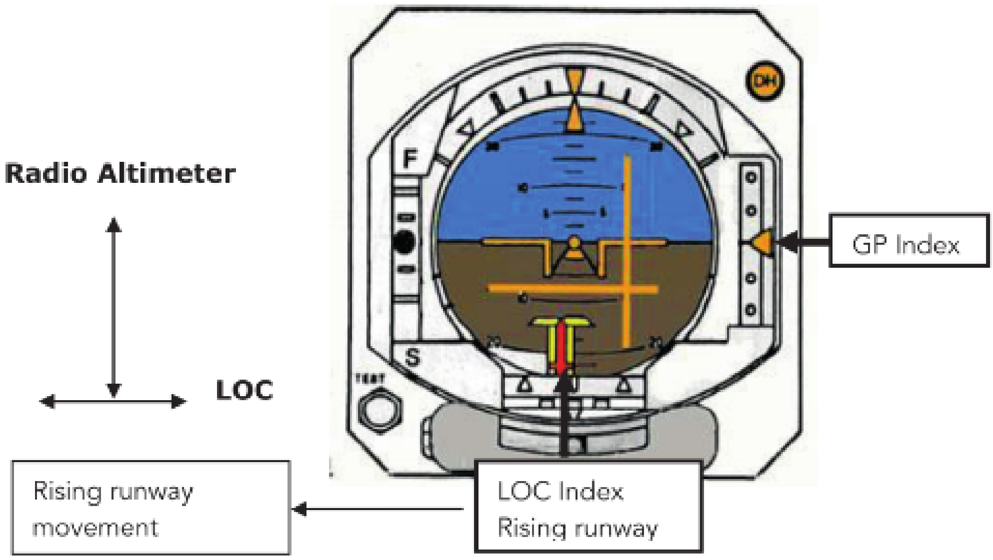
Aircraft positions that correspond to the above instrumental indications



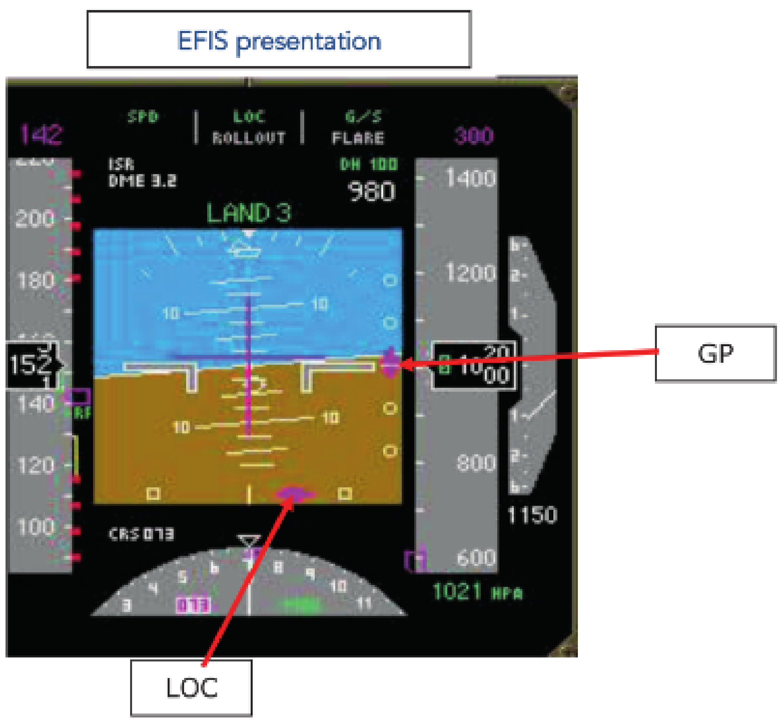
On the above figures, the aircraft is always at the same position, but the platter of the HSI and the compass card of the CDI are oriented North (instrument failure causing a jam for example).
The course deviation indicators have not moved but the mental representation is less easy.
It would be possible to fly the approach in these (difficult) conditions by mentally bringing at all time the platter or the compass card of the CDI in the direction of the QFU in service.

5.2.2 - Instrumentation ADI (or EADI or PFD)

On final approach, in order to reduce the visual scan, all the flying informations are grouped on the ADI, including the LOC and GP information.

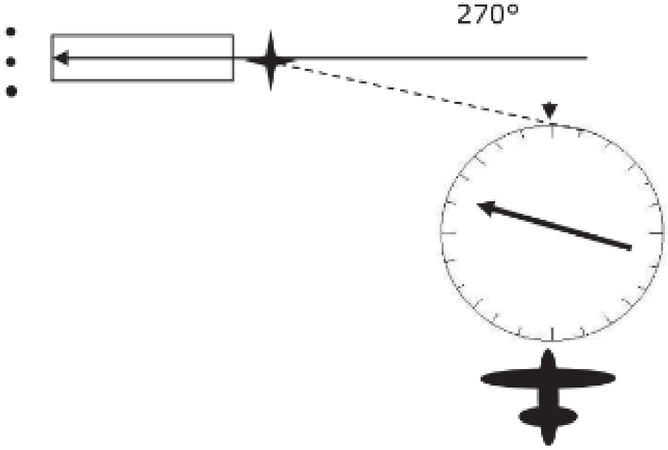


Here the localizer index is a mobile runway.
It shows up below a certain radio altimeter height and is enslaved to the localizer for the left/right position in relation to a central reference point.
It is enslaved to the radio altimeter height.
This was seen in the radio altimeter course.
Caution: do not confuse the Flight Director command bars, displayed on the ADI, with the LOC and GP deviation indicators.



5.2.3 - Utilization

a) Localizer



In order to see the course coming, if a VOR or an NDB is located on the runway center line, an RMI will help anticipate the interception of the course to avoid overshooting it.

The aircraft on the figure on the left, will fly an interception heading that depends on its distance to the threshold and its speed (for example heading 300°).

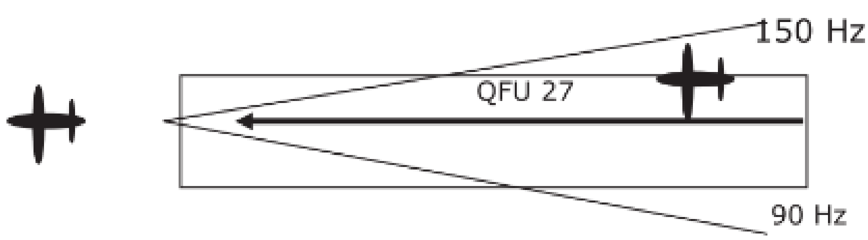
Other way:
If there are two HSIs, one set to LOC and the other to VOR, the LOC needle will be alive when the HIS VOR needle is at ½ point.

Once established on course, flying is carried out catching up the needle.
We must be aware that the closer we get to the runway, for a given angular deviation, the smaller the lateral deviation as a distance to the course will be.
Flying will need to adapt, and corrections will be smaller.
With wind, management of the drift will be the same as with the VOR.

Particular use

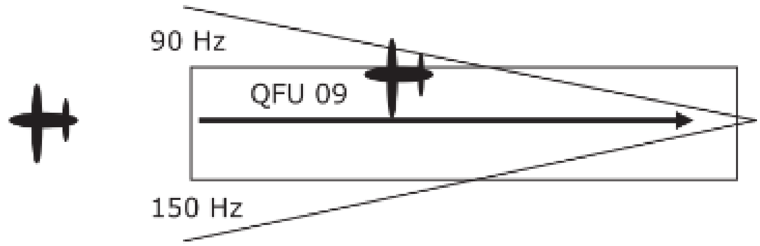
During a low visibility takeoff, we can, for example, use the localizer to maintain the runway center line.

<u>There are two possible cases:</u>
1) The aircraft takes off runway 27 and the pilot has selected the frequency of the ILS runway 27.

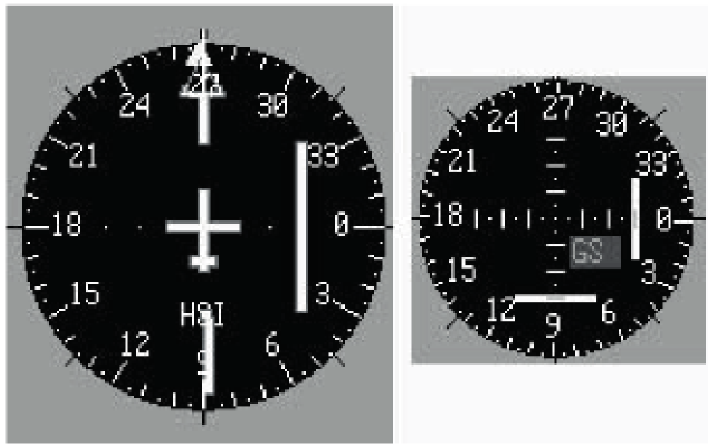


The platter of the HSI and the compass card of the OBI are oriented at 270.
A deviation of the aircraft on the right of the runway center line will translate into a correct deviation on the left of the course deviation indicator, of the localizer index on the ADI, and on the mobile runway.
After takeoff, if the localizer does not radiate a back beam (general case), no signal is available for guidance on the runway center line.

2) The aircraft takes off runway 27 and the pilot has selected the frequency of the ILS runway 09.

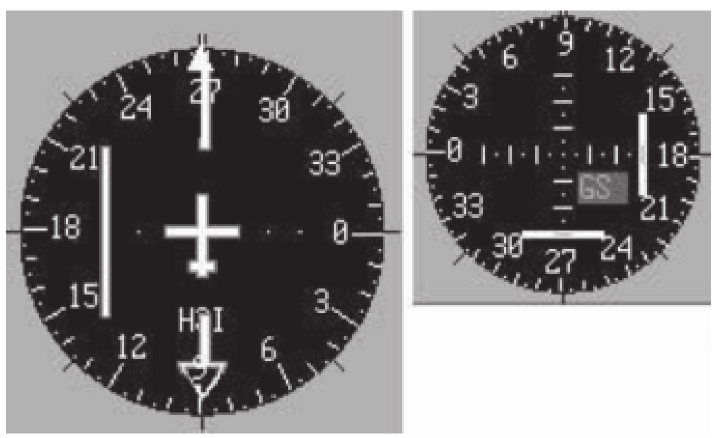


1) <u>The route of the HSI and the compass card of the CDI are oriented at 270°:</u>



A deviation of the aircraft on the right of the runway center line will translate into an incorrect deviation, on the right, of the course deviation indicator.
Indeed, the localizer receiver, receiving more 90 HZ than 150 Hz, will indicate that the center line is on the right.

2) <u>The platter of the HSI and the compass card of the CDI are oriented at 090°:</u>

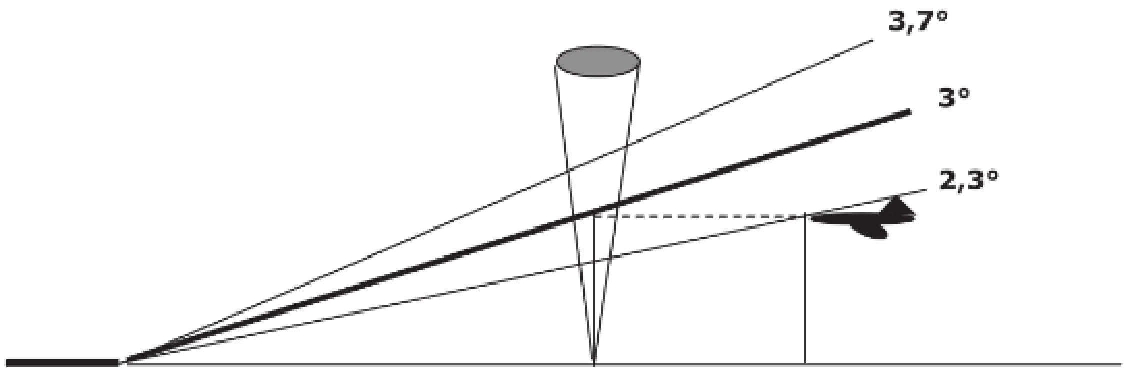


The receiver is still receiving more 90 Hz than 150 Hz, and the course deviation indicator of the CDI shows an incorrect deviation.
If flying is done to center the needle, the aircraft deviates even more from the center line.
The HSI now gives a correct representation of the situation.
Nowadays many aircraft have a Back Course (B/C) function that inverses the signal automatically when the selected course on the HSI platter differs by more than 90° from the aircraft heading.

For the CDI, there is also the possibility on some aircraft to inverse by means of a Back Course (B/C) switch the direction of the deviation indicator of the localizer.
The indications are then correct, and the localizer can be used even after takeoff.

b) Glide path

- Glide path interception:



The glide path is generally intercepted from below, in order to guard against intercepting a false descent path.
These false descent paths come from the parasite beams generated by the directional glide path antenna.
In the same way, there are also false localizer center lines.
When capturing a glide path, it is important to validate that the correct descent path has been captured.
This is done by correlating ~~the~~ information provided by the glide path and the crossing of a marker (or a locator) at a height defined by the procedure.
If the ILS is collocated with a DME, we check the relation distance/ height.
We can calculate at what distance the glide path index will become alive (position 2 on the previous figure).
We start the descent on a glide path only if we are less than one half deviation away on the LOC scale.
Any deviation exceeding one half of the scale for the LOC or the GP on final approach results in a go around (obstacle margin clearance not being guaranteed).

Descent vertical speed:

To follow a descent path materialized by the glide path, which as we know is a ground path, we use a formula like VS = Θ % x GS.
For a ground speed of 120 kt on a 3° ILS, that is a VS of 600 ft/min (remember that 3° = 5 %).

5.2.4 - Identification

The identifier of the LOC is transmitted at least 6 times per minute in Morse code and with a 1 020 Hz tone.
The identifier is made up of sometimes the letter I and 2 or 3 letters.
There is no identifier for the Glide Path.

5.2.5 - Markers

Markers are recognized while flying over the radiation cone by the tone and the Morse code specific to each one.
A cockpit colored light associated to each marker allows visual identification.

75 MHz VHF carrier		
OM	**MM**	**IM**
400 Hz	**1 300 Hz**	**3 000 Hz**
BLUE	**AMBER**	**WHITE**
— per second —	— ▬ per second —	▬ ▬ ▬ per second ▬ ▬ ▬

ILS markers summary

5.2.6 – Integrity monitoring

a) Ground station

Both the localizer and the Glide Path of the ILS ground station are monitored.
The quality and the level of the signal, as well as the precision of the available information are monitored by a monitoring unit.

Depending on the detected problem, either the transmitter is taken out of service, or the category of the ILS (CAT I, II or III) is lowered.
The thresholds are:

LOC center line deviation:
ILS CAT I: ±35 ft;
ILS CAT II: ±25 ft;
ILS CAT III: ±10 ft.

GP deviation:
0,075 θ which is 0.025° for a GP of 3°.

Signal power reduction of 50% or more.

b) Onboard station

Onboard, a monitoring circuit warns the pilot in case of a bad reception or an anomaly of the 90 or 150 Hz radiation.
These faults are reported, depending on the technology of the aircraft, by a flag or the disappearance of the indications (deviation indicator, index).
It is important to remember that an ILS being TESTED (identifier missing or replaced by a continuous tone) can provide false indications without any warning.

The indications will be on the center line and on the path, whatever the position of the aircraft within the coverage diagram.

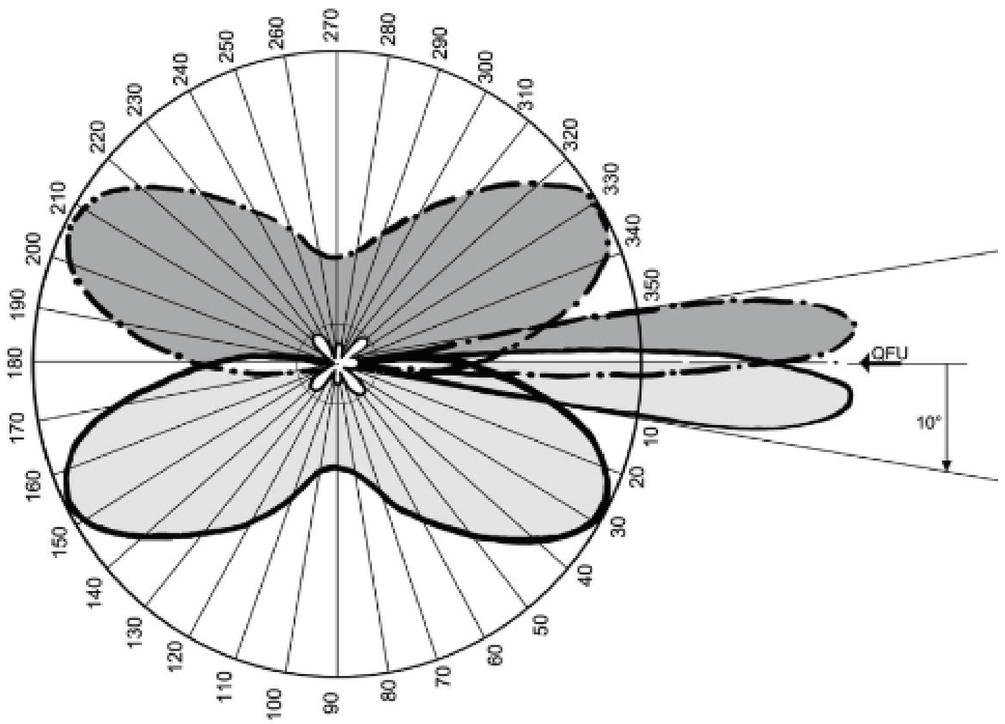
You can find the investigation of in incident on approach by entering in your favorite search engine: SERIOUS INCIDENT REPORT OCCURRENCE 00/2518.

Cross-checking available information coming from different sources is the only way to guarantee safety.

5.3 - Range and coverage

5.3.1 - Localizer

Modern localizers are called directional localizers because they have a double radiation diagram comprising a directional diagram and a coverage (or clearance) diagram.



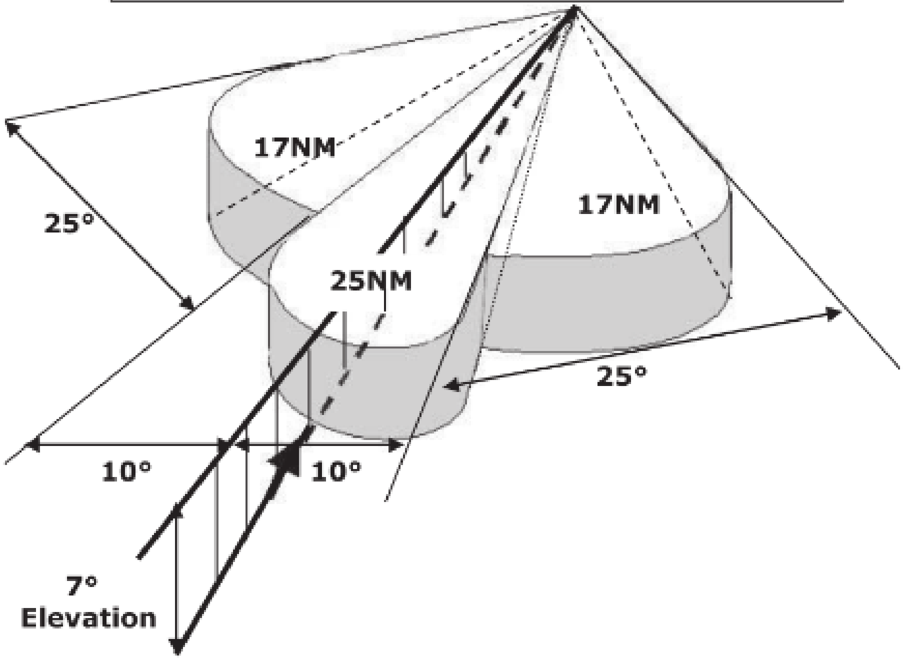
The directional diagram defines the runway center line.

Its narrow aperture angle (10° on both sides of the center line) defines the runway center line by limiting the parasite reflections.

The range is 25 Nm.

The coverage diagram (bean shaped) conceals the false center lines of the directional diagram and allows positioning, on the right or on the left of the center line, outside of the directional diagram.

Azimuth, elevation and distance coverage diagram



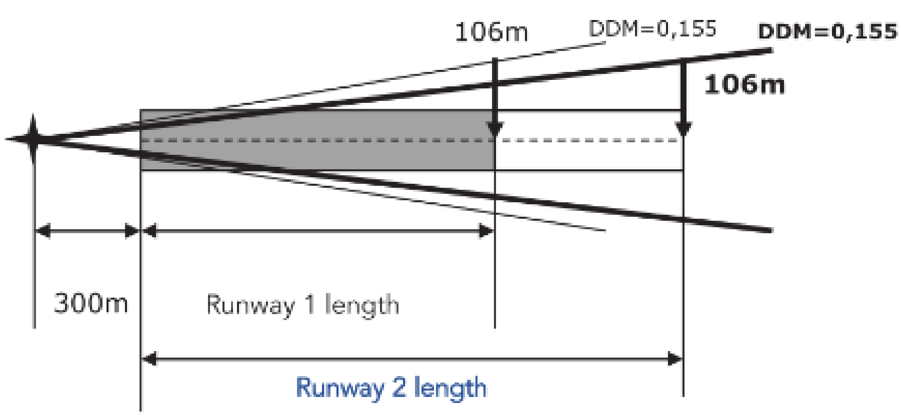
The coverage of the directional diagram of the **Localizer is 10° on both sides of the runway center line in azimuth and 7° in elevation.**

We have seen that within those 10 degrees, the instrument provides a quantified deviation information only in an area of approximately 2,5 degrees on both sides of the center line, which is 5 degrees in total (instrumental window).

This is close approximate value because the area varies from 4 to 6 degrees depending on the runway length.

In fact, the beam is adjusted (ICAO norm) so that 106 meters abeam the runway threshold, the DDM is 0,155, meaning that the indicator displays a needle at maximum deviation.

As a consequence, depending on the runway length, the angle of aperture will be more or less wide.

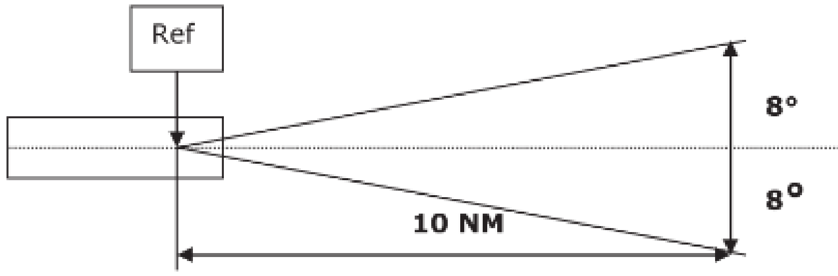


5.3.2 - Glide Path

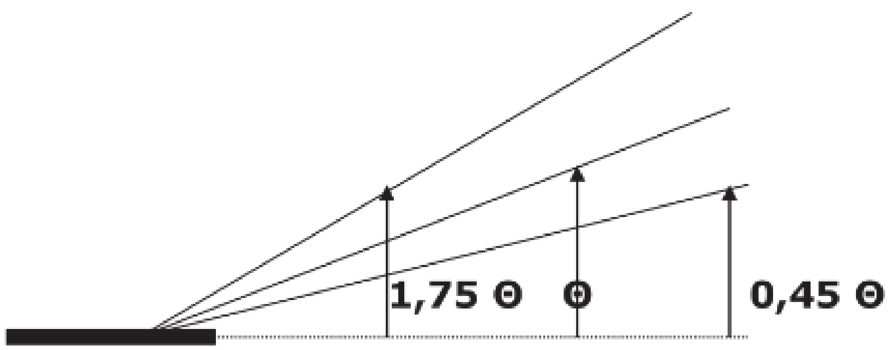
a) In azimuth

The radiation is such that it is possible to have a glide path information before being established on the center line of the localizer.
Coverage is certified up to ± 8°of the center line with a range of 10 NM.

Ref: Point where the straight extension of the downward alignment axis intersects with the runway.



b) In elevation



For a path Θ = 3°, the coverage is guaranteed between 1,35° and 5,25°.
The limit 0,45 θ can be lowered up to 0,3 θ if necessary, to protect the issued procedure for the interception of the descent alignment.

5.4 - Errors and accuracy

5.4.1 – False LOC center lines

Minimized by the installation of the null reference localizer.

5.4.2 – Beam bends

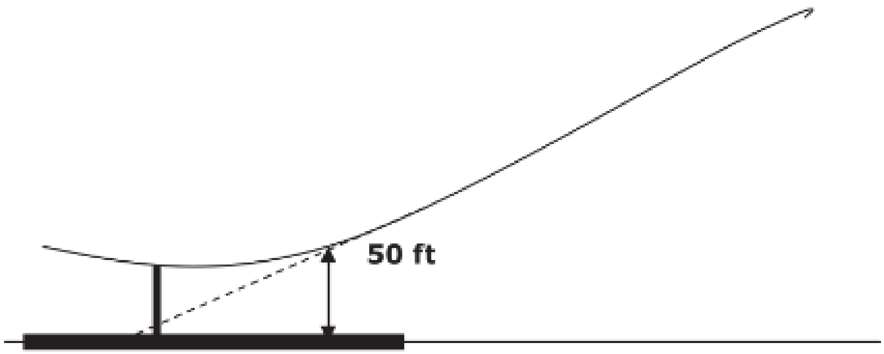
If an obstacle (fixed or mobile) capable of reflecting the signals emitted by the localizer is present in the vicinity of the runway or in the coverage area, the aircraft on approach receives the components of the direct and the reflected fields, producing an error depending on the directivity of the emitting antennas and on the reflecting property of the obstacle.

5.4.3 – Glide path reflection

As for the localizer, obstacles or mobiles located in the clearance areas can disrupt the information.
In contrast, because the antennas are located close to the threshold, aircraft at takeoff do not disturb the information for aircraft on approach, but an aircraft on approach in front of another one can be the cause of error.
In LVP, aircraft separation will have to be increased.

5.4.4 - Glide path parallax

This is due to the fact that the antennas are not on the runway center line.



From a certain height, the descent alignment axis is no longer a straight line but a hyperbole that is not tangent with the runway, and if we continued to follow that beam, we would not reach the runway.
The beam is adjusted so that the extension towards the end of the glide path cuts the runway threshold at 50 ft.
The autopilot of aircraft performing Autoland does not use the glide path information all the way to the ground (see Autopilot course) and this is the reason why the radio altimeter is essential for a CAT III approach.

In summary, the errors are:
- False center lines;
- Beam bends;
- Interferences (potentially with FM radio that goes to 108 MHz);
- Parallax.

5.4.5 - Categories and precision of the ILS

Depending on the performances of the ground station (type of Localizer and Glide Path in use), and the challenges related to the environment, the ILS are grouped in 3 categories.
We must distinguish the category of the ILS, related to the precision provided by the ground station, defining what could be done with such an ILS, from the approach category (CAT I, CAT II, CAT III a, b or c), meaning what we can do, taking into account:
the type of aircraft and its equipment;
the qualification of the pilots;

the quality of the ILS and the measures taken to guarantee its precision, its reliability, and its integrity.

ILS categories:

- Category 1: Guaranteed guidance supplied down to 200 ft above ground level;
- Category 2: Guaranteed guidance supplied down to 50 ft above ground level;
- Category 3: Guaranteed guidance supplied down to the runway and on the runway.

The approach operation categories CAT I, CAT II or CAT III, depend on parameters such as the aircraft category (C or D for example), the decision height (DH) and the runway visual range (RVR), and are part of the Air Law course.
Brief reminder:

Cat I: DH 200 ft, RVR 550 m;
Cat II: DH 100 ft, RVR 350 m;
Cat IIIA: DH 100 ft, RVR 200 m;
Cat IIIB: DH 50 ft, RVR 50 – 200 m;
Cat IIIC: No external visual reference.

The verification of the precision of the ILS and the detection of errors are performed during test approaches by the air navigation service provider and is called ILS calibration.

5.4.6 – Flying accuracy on approach

ICAO DOC 8168 specifies that an aircraft is considered established on approach with a maximum of a half-scale of instrumental deflection for LOC and GP.
The aircraft has to be established within the half-scale deflection of the LOC before starting descent on the GP.
Flying under the glide path with more than a half-scale deviation requires an immediate go-around because obstacle clearance may no longer be guaranteed.

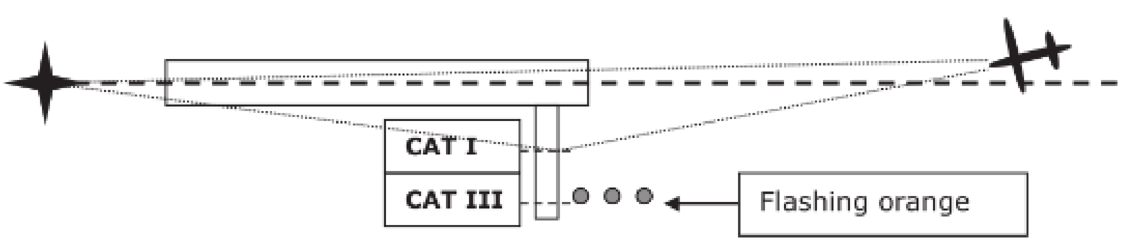
5.5 - Factors affecting range and accuracy

5.5.1 – Critical area

Vehicles or aircraft can generate reflections of signals which interfere with the proper conduct of approaches by inducing deviations larger than the instrumental half-scale (maximum deviation warning).

There is a defined critical area where no aircraft or vehicle can penetrate so that ILS signals are not disturbed.

This area is extended if the airport is in LVP (Low Visibility Procedure) to prevent unacceptable interference to the ILS signal.
A holding point named CAT II – CAT III is placed further away from the threshold than the normal holding point.



On airports where aircraft separation is minimal, when CAVOK, we sometimes witness fluctuations of the LOC indications when flying on approach with an aircraft taking off.

5.5.2 – Sensitive area

It is an area extending beyond the critical area where traffic is controlled to prevent unacceptable interference to the ILS.

5.5.3 - Interference

LOC signals can, on some airports, be disturbed by the emission of free radio stations emitting in wide frequency modulation (WFM) for which the authorized bandwidth (in France) goes up to 108 MHz.

06 MICROWAVE LANDING SYSTEM (MLS) CPL

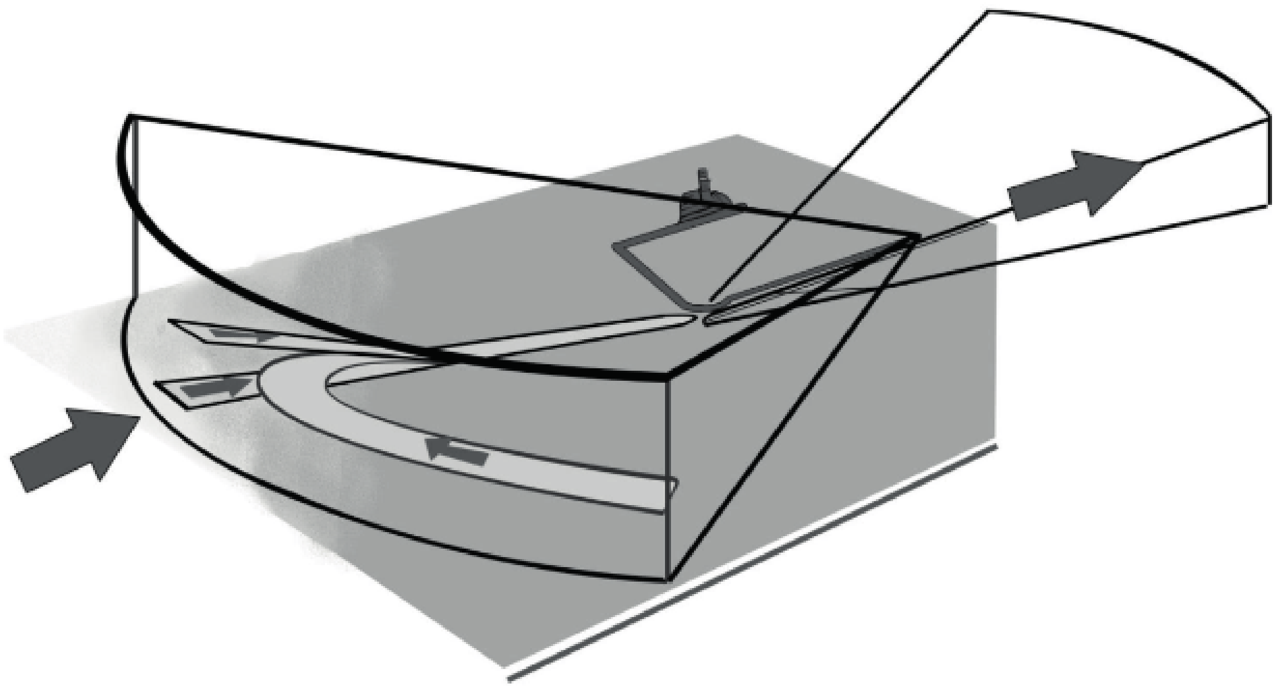
Note from the author:
The new LOs have kept the MLS even though there is no more airport that uses the MLS.
The last approach charts for EGLL that I have (2018) do not mention it anymore.
The author would appreciate any information from a pilot who has flown an MLS procedure.

6.1 - Principles

The MLS was conceived to replace in time the ILS, suffering from certain limitations such as:

- Only one approach course;
- 40 available channels;
- Beam bends;
- Sensitivity to the geographical area.

The MLS has a direct approach volume ensuring guidance for approach and landing, and an opposite volume ensuring guidance for takeoff and missed approach.



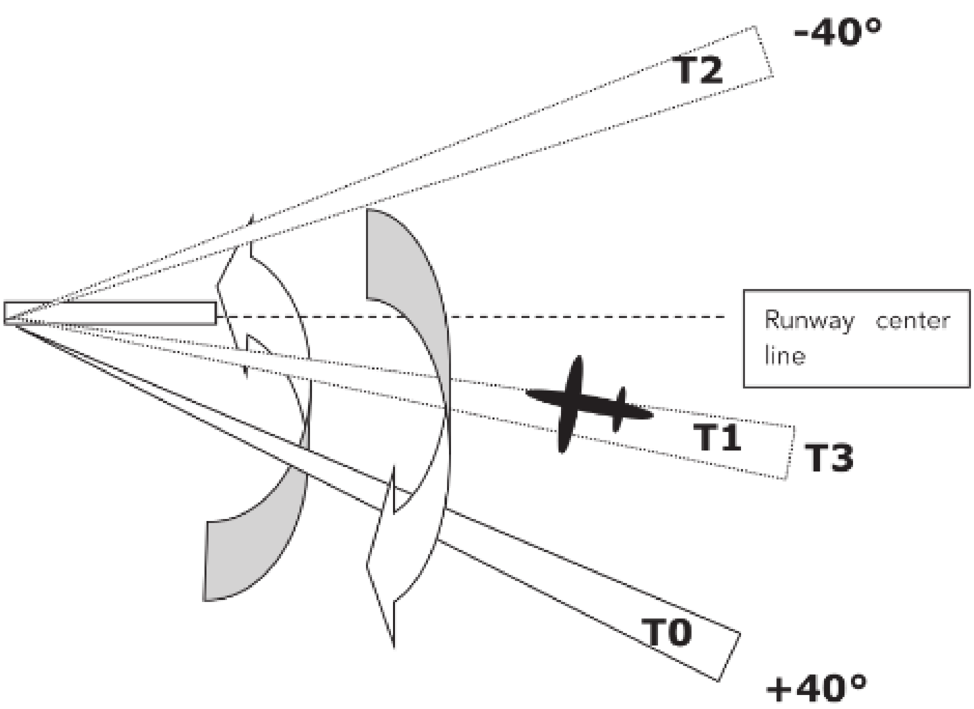
The applied principle allows approaches on any axis and on different paths inside the covered volume.
These approaches can be direct, segmented or curved, and on different paths, which the ILS does not allow.

For now, the MLS operates in the SHF band (5031 à 5090 MHz) divided into 200 channels from channel 500 to 699.
It is associated to a Precision DME, Distance Measuring Equipment, (DME/P).
Its principle, fundamentally different from the one of the ILS, is based on a deviation measurement performed with the time reference scanning beam (TRSB) technique.

It is a measurement of time.

Another technique existed, based on the Doppler effect.

6.1.1 – AZIMUTH tracking



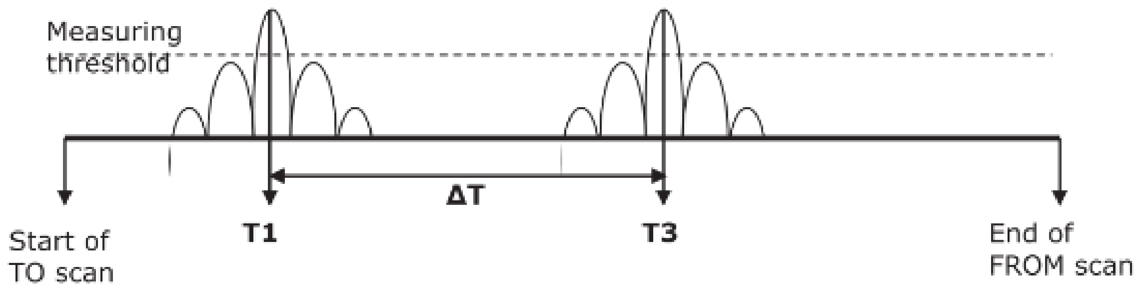
A beam scans at a radial speed ω, a portion of space from +40° to – 40° on either side of the runway center line.
Assuming that the aircraft is on the runway center line, the beam leaving from an origin position at time T0 is received by the aircraft at T1, it then continues scanning.
It scans the space in the opposite direction from a time T2, and it is received in return by the aircraft at time T3.
If, as in our example, the aircraft is centered, the difference Δt1 = T1-T0 is equal to the difference Δt2 = T2-T3.
Any deviation of the aircraft from the center line will have Δt1 not equal to Δt2.
The Δt difference will be considered representative of the angular deviation from the center line.

The sign of the difference of the Δt will be considered representative of the left/right position relative to the center line.

The measuring principle is a measure of time.

Since ω is known and fixed, the scanning time, T0 T2 to the aircraft and T2 T0 from the aircraft, also is known.

The measurement comes down to measuring the difference T1-T3, the time between the reception of the signal from the TO scan and the reception of the signal from the FROM scan.



To perform this kind of measurement, the receiver must be synchronized on the scan of the beams.

6.1.2 – Vertical tracking

The principle is identical to the azimuth tracking, with a scan in the vertical plane, the pilot is able to choose a preprogrammed Δt in the domain, and thus a glide path, with a knob identical to the course selection knob.

6.1.3 - Back Beam

The MLS also radiates information on the opposite track of the QFU, making its use available for go-around or takeoff.

6.1.4 – Auxiliary information

Transmission of special information such as station identification, runway condition, weather information including windshear, or status of the system.

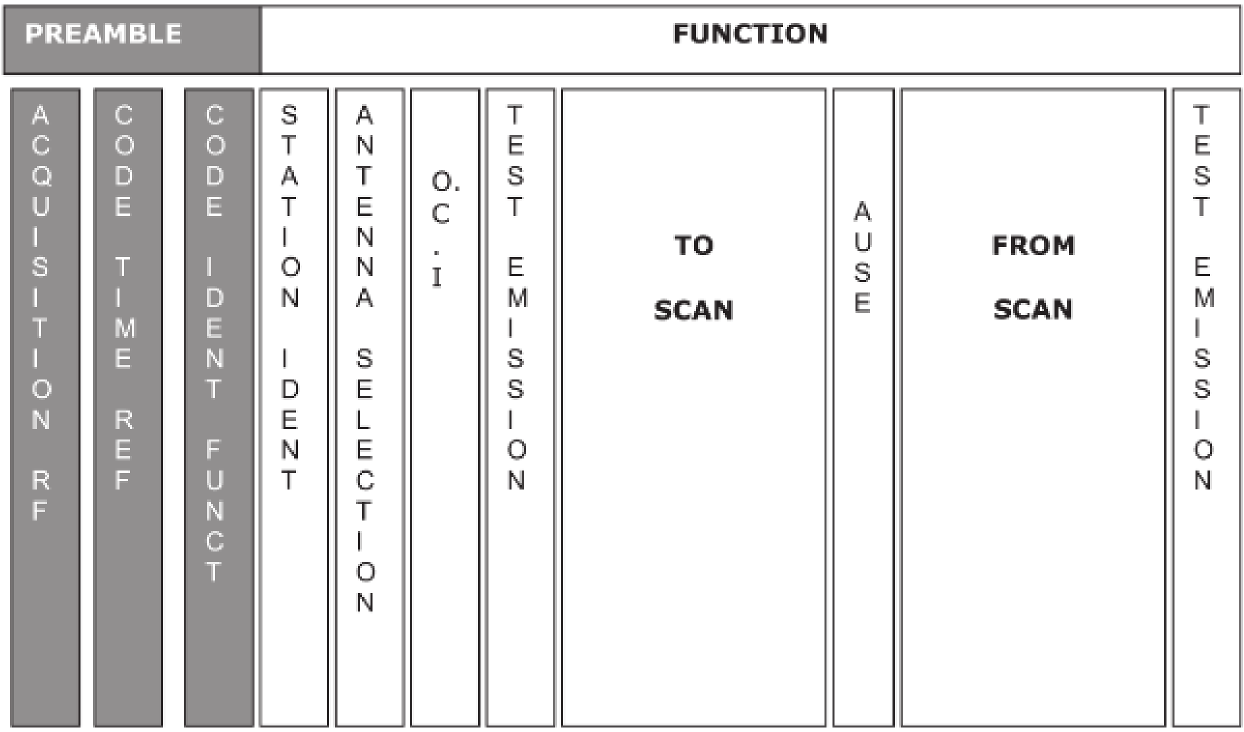
6.1.5 - Functions

The MLS has, excluding DME-P, 4 operational functions:

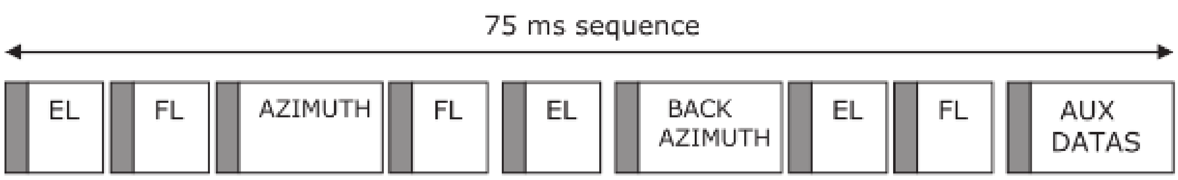
- Azimuth guidance;
- Vertical guidance;
- Back beam;
- Auxiliary information.

All of these functions are transmitted alternately in a 75 ms sequence, on a unique 5 GHz frequency.

Each function starts with a preamble in order to synchronize the receiver, identify the emitted function, and, depending on the function, know the limits of the guidance area and the alignment of the descent area.
The functions are transmitted on multiplex during a 75 ms sequence.



Details of information contained in the preamble and functions.



Function multiplex

▮	:	Preamble of each function;
EL	:	Elevation measure;
FL	:	Flare (not used);
AZIMUTH	:	Azimuth measure;
BACK AZIMUTH	:	Azimuth measure on back beam;
AUX DATAS	:	Auxiliary information.

It is possible to modify the order of the functions in a sequence.

The transmitter selects the proper antennas radiating in azimuth or elevation according to the function.

Note that for one transmission of the azimuth, the elevation function is transmitted three times.

6.1.6 – Better insensitivity to sites

The problems of reflections from obstacles that were encountered with the ILS can be solved with the MLS.
It is indeed possible during the scan to suspend the emission in a given direction without interrupting the scan, and the deviation measurement.

6.2 - Presentation and interpretation

6.2.1 – Onboard station

It is composed of an SHF receiver covering the 200 channels of the MLS (500-699), or a multifunction receiver MLS-ILS-GPS.
In the MLS function, the receiver can be in AUTO mode, and in that case guide the aircraft on the planned course and vertical path, or in MANUAL mode and it is then possible to select an approach course different from the runway center line, and a different vertical path.

6.2.2 - Instrumentation

It is easy, by preprogramming a difference of Δt, to determine a course to follow, other than the runway center line, included in the +40°, -40° volume (see previous figure).
As soon as the system will stop measuring the preprogrammed Δt, it will mean that the aircraft is not on the desired course, and the deviations with respect to the preprogrammed Δt will be representative of the deviations from the selected course.
The navigation instruments developed in the ILS chapter such as the HSI and the CDI are fully compatible.
The deviations presented on these instruments are in respect with the course and glide path selected by the pilot, within the usable volume.
It is then possible if the DME-P is operational to follow curved or segmented paths outside of the runway center line and in the limit of the guidance volume, since we will have at all times a very precise Rho-Theta position

If the DME-P is not operational we can only perform an approach on the runway center line course, like with the ILS.

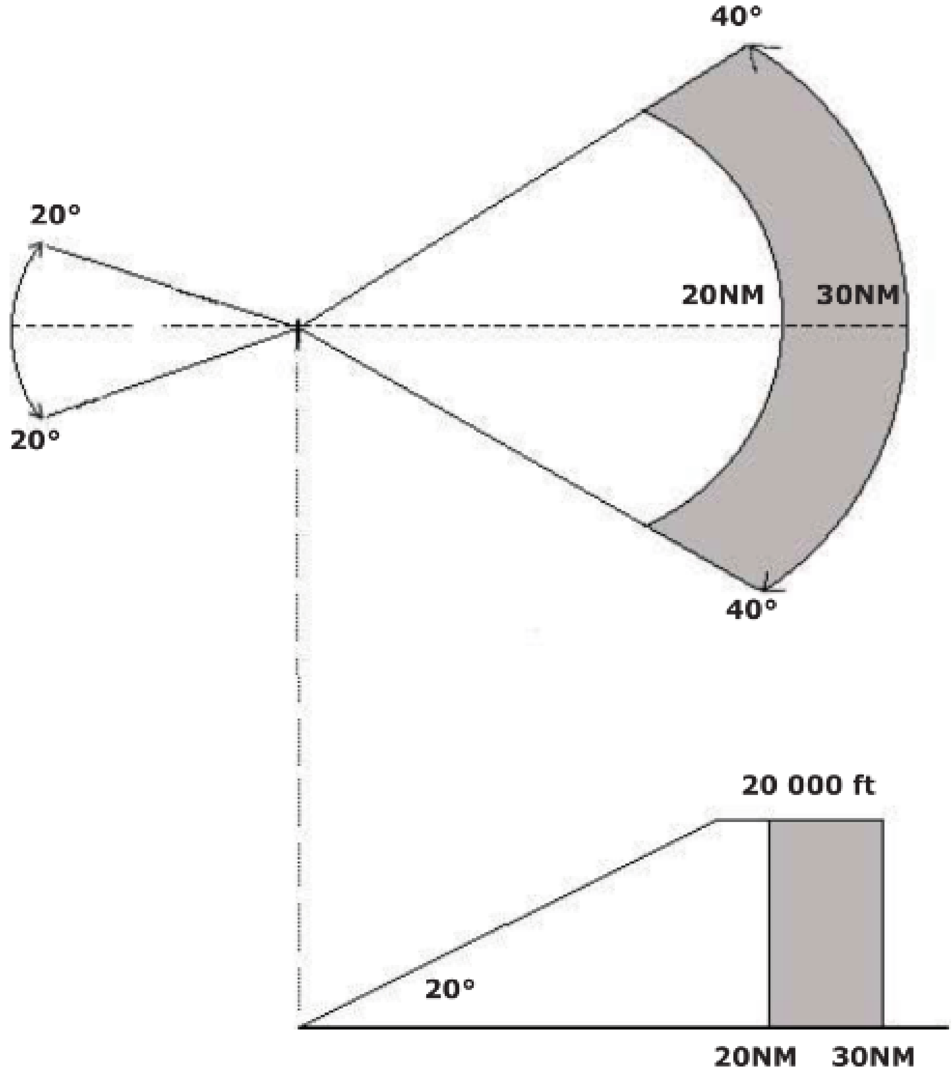
6.3 – Range and coverage

6.3.1 - AZIMUTH function

The azimuth beam covers at minimum 40 degrees on both side of the runway center line and up to 20 degrees in elevation.
Range is guaranteed up to 30 NM from the airport.

6.3.2 – Elevation function

20 degrees in elevation up to 20 000 FT, on an area equal to the one of the azimuth function.



6.4 - Accuracy

±20 ft in azimuth and ±2 ft in elevation in CAT III approach.

03

RADAR

01 PULSE TECHNIQUES

Radar is an acronym for RAdio Detection And Ranging.
As its name indicates, it allows detection of remote objects and location in azimuth and range (sometimes in elevation).
The radar provides a position and, in some cases, a speed.
Developed before World War II, the radar works with pulses.
In aviation, radars are used by air traffic control (ATC), meteorological centers, or aircraft as airborne weather radars (AWR).

There are primary and secondary radars.
This chapter covers only the primary radar using the echo principle (*called here radar*).
An entire chapter is dedicated to the secondary radar which does not use the echo principle.

Radars use UHF or SHF frequencies and, depending on the frequency, are referred to as working in band L, C or X for example.

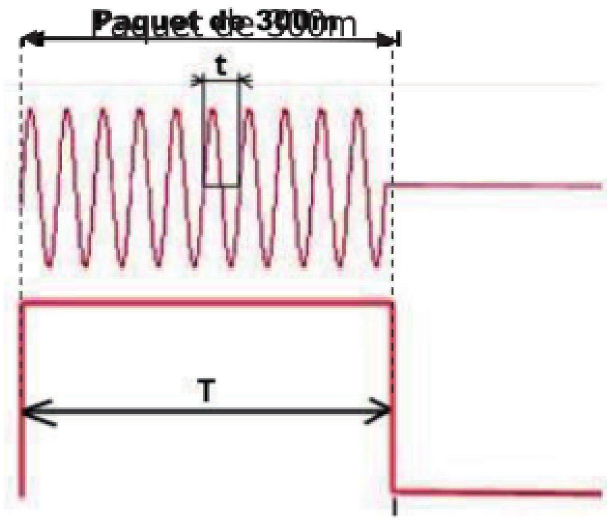
Breakdown of UHF and SHF bands into sub-bands

BAND	SHF				UHF		
Sub-band	Ka	Ku	X	C	S	L	P
Wavelength (cm)	1,00	2,00	3,10	5,60	9,60	23,00	68,00
Frequency (GHz)	35,00	14,00	9,60	5,30	3,00	1,30	0,44

The principle of radar which is based on the reflection of waves (echo), imposes that the wavelength λ be much lower when the objects detected are smaller.
If we talk about frequency, then the frequency must be much higher as the obstacles become smaller.
This is why an en-route ATC radar works at around 1 300 MHz (L band), when a weather radar responsible for detecting hydrometeors works at a much higher frequency (5 400 MHz for the older ones or 9 375 MHz).
Basic principles remain the same, with the difference that some radars illuminate the space around 360° using rotation of their antenna while others only scan a portion of the space.
The radar system illuminates, at a given time, a portion of space with a radio wave and, in return, receives the waves reflected by the obstacles that are in the portion of space that is illuminated.
The primary radar uses the echo principle, which is not the case for the secondary radar.
Analysis of the echoes detects the presence of objects and their position (azimuth, distance, and, sometimes, their nature).

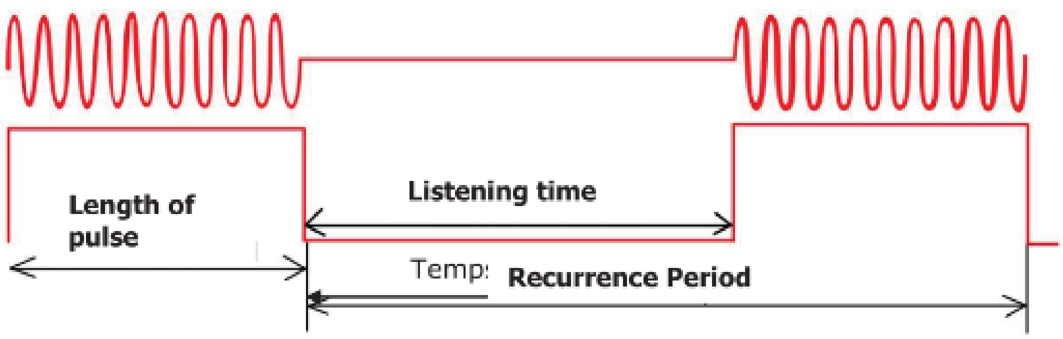
The operation is merely a succession of Transmitting/ Listening cycles.

We call pulse the transmission during a time T of a packet of oscillations at UHF or SHF frequencies.
A pulse of period T = 1 µs will give a packet of oscillations of 300 m.

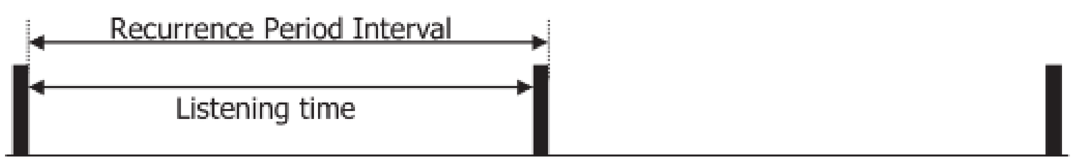


This means that when the last oscillation of period t is transmitted by the antenna the first transmitted oscillation will already have traveled 300 meters in space.

Pulses are transmitted at a rate called recurrence period.
We call recurrence period the time that it takes for a phenomenon to reproduce itself strictly identically and in the same direction.
The listening time on the radar is time between two pulses.



The pulses shown here are quite expanded, but in reality they are very short compared to the recurrence period interval, which is why, to simplify, we will assimilate the listening time to the recurrence period interval (PRI).



This shows pulse length and listening time closer to reality.

Pulse Recurrence Frequency (PRF) is the number of pulses transmitted in one second.
Pulse power is the amount of energy stored and released in the pulse (often compared to pulse release time)
If a radar transmits 300 pulses per second, we say that the pulse recurrence frequency (PRF) of the radar is 300 hertz.
The pulse recurrence interval (PRI) is then equal to 1/ PRF, which is approximately 3333 µS.

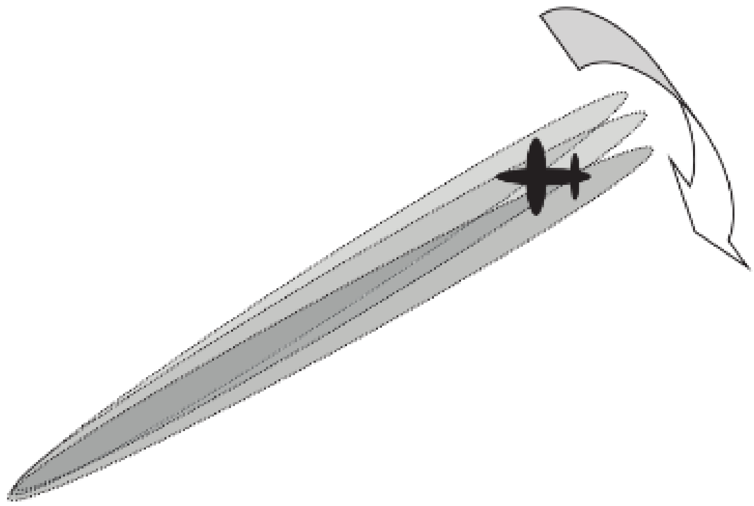
Note:

We must be aware that the operation is not limited to the transmission of one pulse and the wait for one return echo.

The target must be illuminated for a sufficiently long time.

Let's take a radar with an antenna rotation of 15 revolutions per minute (one antenna rotation in 4 seconds) and a recurrence period of 4 ms.

With a beam width of 3°, the target will be "illuminated" during at least 33 ms during the rotation of the beam.



Therefore a succession of pulses hit the same obstacle and come back to the processing chain of the radar in order to give the proper echo.

Equation of power received on return:

$$Pr = G.Pe.\frac{S.\Sigma}{(4\pi D^2)^2}$$

Pr: Power received;
Pt: Power transmitted;
G: Gain of the antenna;
S: Surface of the obstacle;
Σ: Surface of the antenna;
D: Distance antenna / obstacle.

Without dwelling on the mathematical aspect of the equation, it shows that for the return echo to have an exploitable level, it will need:

- A considerable transmission power;
- An antenna with a significant gain.

Antennas are very directive and have a large surface (sometimes 45 m^2 in L band for ground radars).

The gain G is in the order of 1 000 to 10 000 and the width of the beam is 2 to 3°.

For onboard material, a compromise between bulkiness and antenna gain has to be found.

Antennas which used to be parabolic are now slotted planar array and are common used for transmission and reception (referred to as a monostatic radar).

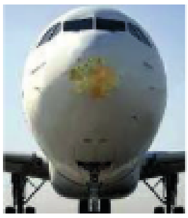
These antennas are developed in chapter 1- Antennas and Radio Propagation

Parabolic antenna on KC135

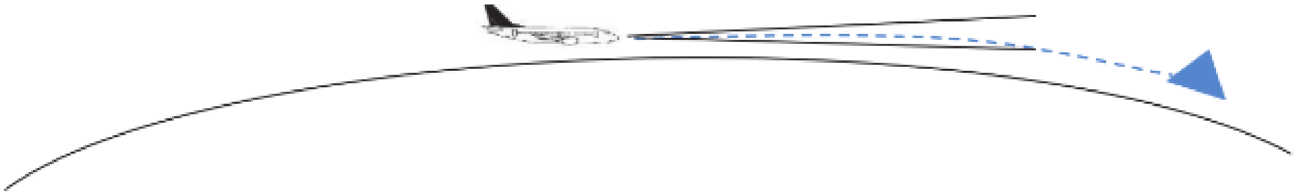
Slotted planar array of a weather radar

Of course, the reflectivity of the obstacle and its size will also play a role.
Metal obstacles (aircraft parts) or hydrometeors (big water droplets) are good reflectors and provide good echoes.

Note:
A very dirty radome, (due to collision with birds) can considerably attenuate the signal.

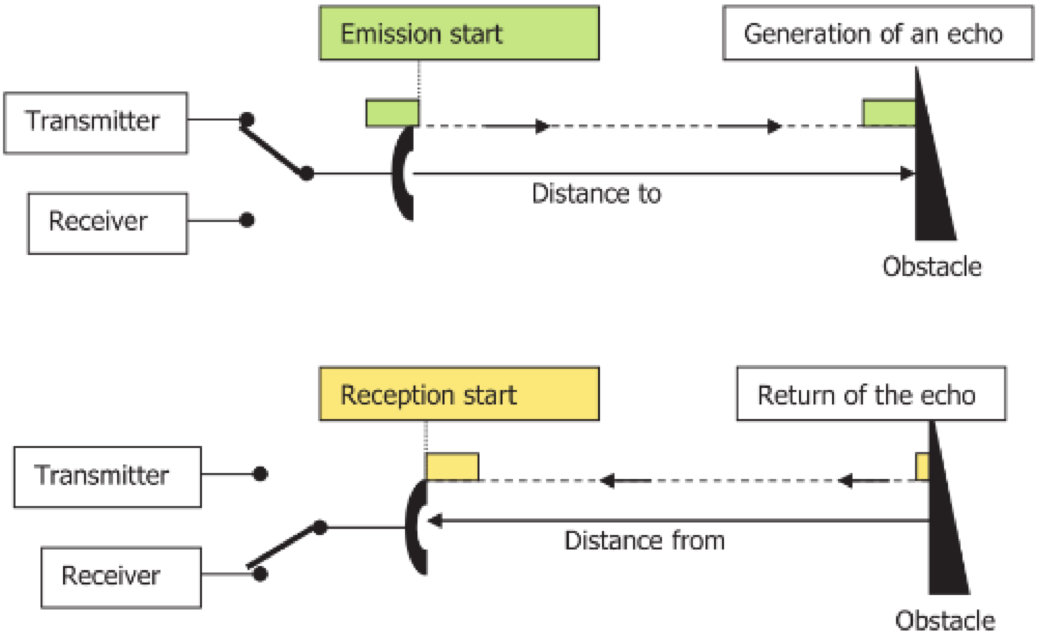


Furthermore, during its travel through the atmosphere the wave is attenuated on its way out and coming back, and the equation mentioned earlier shows that to double the theoretical range of a radar (in the limit of $1,23\sqrt{h}$), the transmission power Pt will have to be multiplied by 16 in order to keep Pr constant.
UHF and SHF normally propagate in a quasi optical range but with some atmospheric conditions (temperature inversion) the range can be significantly increased because of a super-refraction phenomenon already studied in wave propagation.



1.1 – Distance measurement

The measurement used is a measure of time.
The elapsed time between the start of transmission of a pulse and the start of reception of the return echo is, on the basis of a wave propagation velocity C of 3.10^8 m/sec, represents the fly back distance between the antenna and the detected obstacle.



Given Δt (time difference between the emission start and the reception start) equal to 500 µS, the radar/obstacle distance is equal to:

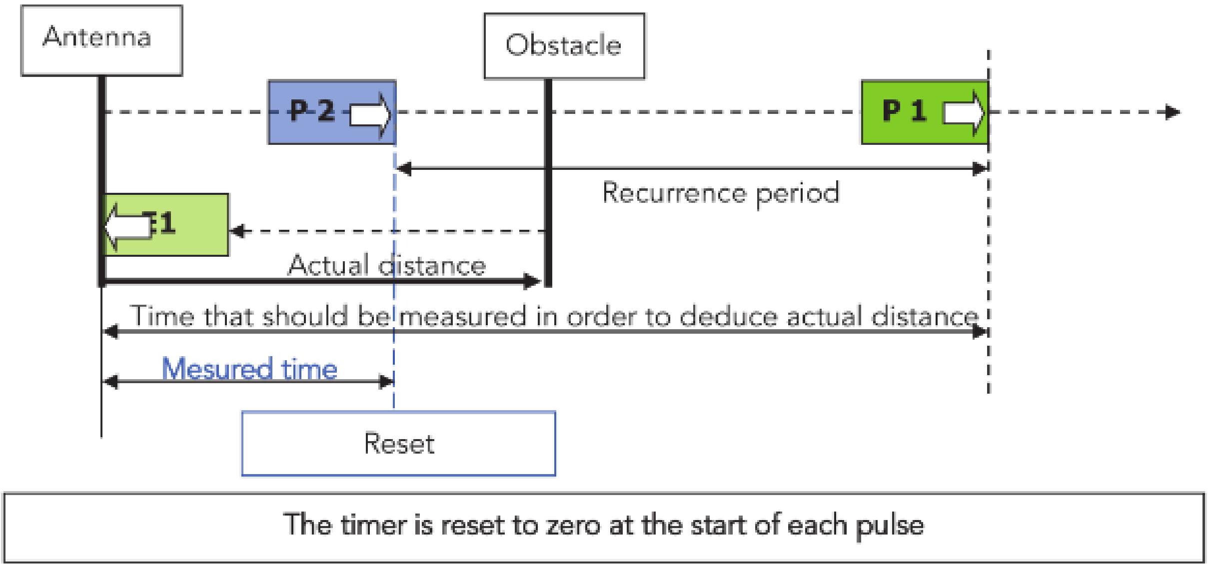
$$\underbrace{3.10^8}_{V} . \underbrace{500.10^{-6}}_{t} = \underbrace{150.10^3 \text{ meters}}_{2D} \quad \text{(it is the fly back distance)}$$

and D = 75 km

This distance is most often expressed in Nautical Miles and thus requires a conversion.

1.2 - Maximum unambiguous range

We recall that the radar principle is based on a succession of transmission-reception cycles and that distance measurement is based on a measure of time from the start of the transmission of a pulse to the start of the reception of the echo by the antenna.
For this process to work as planned, the echo of the obstacle must be received by the antenna before a new pulse is transmitted, otherwise the time measured between the antenna and the obstacle might present an ambiguity.



In the schema above, a new pulse (P2) has been transmitted before the arrival of the E1 echo relative to P1, therefore the measured time is false because it starts with the transmission of P2.
The time limit for the return echo (linked to the recurrence period) taken is just before the transmission of the new pulse.
After this time limit, there is ambiguity on the measurement of time.
In other words, the maximum time allowed for the echo to return is the time separating two pulses, that is the PRI (we ignore the pulse length which is very short).
Maximum unambiguous range is calculated with the formula:

$$D = \frac{C}{2Fr}$$

Also the notion of minimum range (or blind area) for the radar must be taken into account.
If an obstacle is close to the radar (ex: 300m) an echo returns very quickly to the antenna.
This echo arrives before the radar pulse is totally transmitted.
Because the radar is still transmitting, the antenna is not commuted to the receiver which cannot "see" the echo.
Even though this problem is not of great interest to pilots it shows that the pulse radar described here is not capable of measurements on very short distances. This is the reason why a low altitude radio altimeter, which is a radar, uses a different technology called continuous-wave radar.

We have just developed here the radar fundamentals on which old generation of radars operated (transmission of a pulse – reception of an echo).
Modern radars equipped with highly developed algorithms now transmit bursts of pulses of different recurrence periods and frequencies, which then go into receivers.
A virtual picture of the situation is reconstituted from a succession of echoes and using memory capacity. (see compression of pulses and 3D radar at the end of this chapter).

1.3 - Speed measurement

If the objective of a radar is to provide the position of an obstacle, using the Doppler effect it can also measure the speed of a moving object (extrapolation of paths, military interception radars) or the speed of hydrometeors in a cloud (we will come back to this point with the new airborne weather radars).

02 GROUND RADAR (CPL)

2.1 - Principles

Primary radars based on the echo principle provide distance and azimuth information for all detected aircraft, whether or not they are equipped with a secondary radar transponder.
These radars can be used for:

- Route surveillance (range 250-300 Nm);
- Terminal area surveillance (range 80 Nm);
- Approach (around 25 Nm);
- Airport surveillance (runway, taxiway, apron traffic surveillance).

Note:
A new technique called multilateration, based on the use of the secondary radar, allows improved quality of the ground surveillance and accurate traffic identification.
This requires the onboard transponder (See SSR) to be constantly in use up to the parking point.

2.2 - Presentation and interpretation

ATC centers are now equipped with computer processing centers in charge of generating radar pictures drawn from the combination of many sources, allowing an increase in radar coverage, better accuracy and integrity of information.

ATC provide these services:
- Information;
- Surveillance;
- Guidance.


ATC radar antenna

Examples of used sources:

Primary radar
Mode S secondary radar
ADS-B (Automatic Dependent Surveillance)
Multilateration system

03 AIRBORNE WEATHER RADAR (CPL)

3.1 - Principles

The weather radar detects surrounding weather conditions (thunderstorms) and
indicates areas to avoid and which would present a hazard for flights using information on the azimuth, distance and shape of the disturbed area.

It can also be used as an autonomous navigation aid.

The weather radar operates using the primary radar principle (pulses transmission and echoes reception).
We remember that the frequencies used must be in accordance with the size of the objects to be detected, therefore the frequencies used by the modern weather radars are in the X band at 9 375 MHz giving λ = 3 cm, relating to the biggest hydrometeors.
The weather radar works in SHF frequencies.

The radar in the X band (9375 MHz) has finer detection with the disadvantage that this frequency is better absorbed (see the wave propagation course, Attenuation).
A first cell well detected can hide another more dangerous cell located behind. (Creation of a shadow area behind the first cell).

3.1.1 - Composition of the radar system

a) Transmitter

A transmitter receiver working in SHF using the echo principle.

b) Antenna

A common antenna for the transmitter and the receiver.
It is located in the nose of the aircraft (radome) and scans a sector in front of the aircraft.
Depending on the radar, the scan will be 45° on both sides (90° coverage) or 90° (180° coverage), or even 120°.

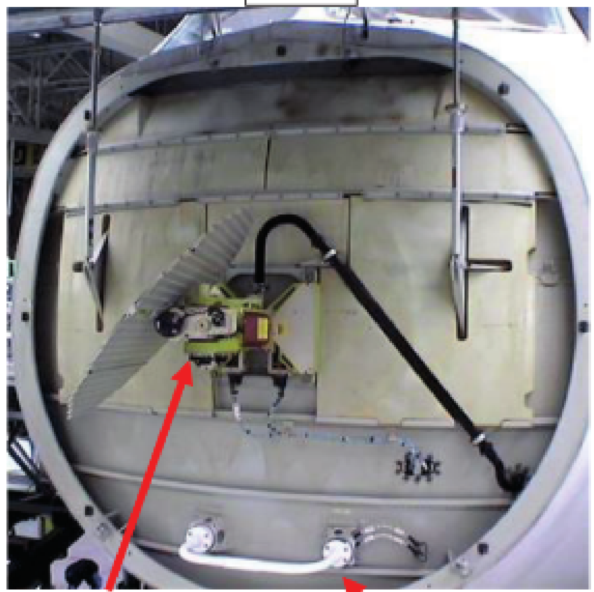
The antenna must be stabilized in roll and pitch so as to not be subjected to movements of the aircraft and not distort the resultant picture or momentarily scan the ground unintentionally.
Stabilization is carried out using signals from the aircraft IRS (or inertial measurement unit).
Any precession of the sensor delivering roll and pitch information will affect the area scanned by the radar and distort the delivered picture.
(This Phenomenon which, depending on the precession speed, is greater with a gyroscope than with an HARS or an IRS).

A340

Slotted planar antenna

Wave guide

A340



Antenna enslavement mechanism (stabilization)

For info: ILS Antenna

The beam width in weather mode is 3 to 5°.

A formula allows us to determine approximately this width:

$$\text{Beam width in } ° = 70 \times \frac{\lambda}{Diameter}$$

λ is in cm
Diameter is in cm

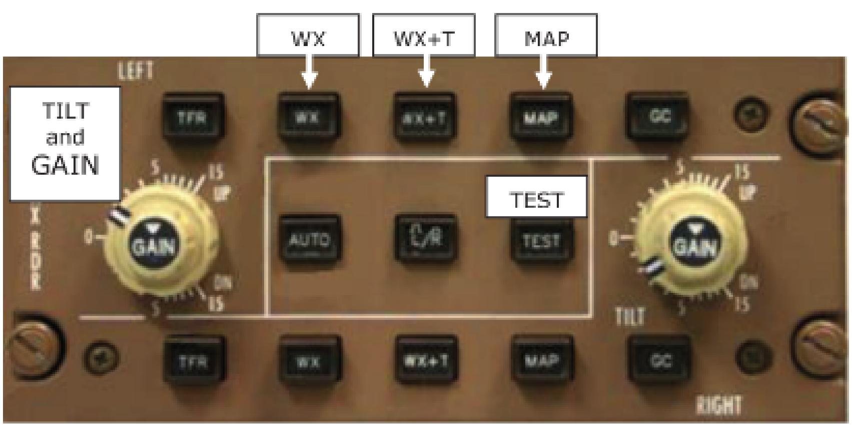
For a wavelength of 3 cm and an antenna with a diameter of 45 cm, the beam width is 4,7°.

A slotted planar array has a narrower beam, less side lobes and thus a higher concentration of energy in the main lobe.

The same range requires a lower power.

3.2 - Presentation and utilization

3.2.1 – Control panel



Collins Multiscan Radar WXR-2100 on Boeing

Only the labeled functions will be covered in this course, the others will be seen during the type rating.

Collins Multiscan Radar WXR-2100 on Airbus

We find the same functions under different forms and names.

The PWS function will be described later.

These two control panels provide control for one or two multiscan radar systems.
The principle of the multiscan radar is to scan the space between different altitudes, to memorize all the echoes after software processing in a 3D memory, and to restitute a picture on display screen upon request.

WX (Weather): Displays the echoes from the disturbed weather according to their intensity in green, amber, red.

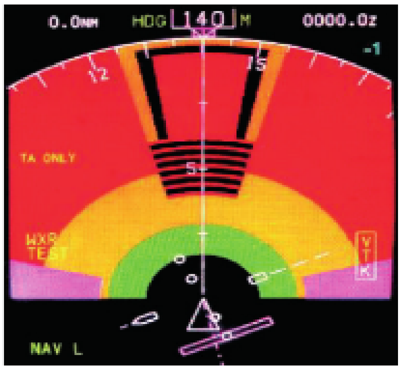
WX +T (Weather + Turbulence): Highlights the dangers from disturbed weather.
Turbulence areas are displayed in magenta, and detection by Doppler effect allows detection within a 40 Nm limit.

MAP: Display of ground echoes as a navigation aid.
Note:
Ground covered with snow will return weak echoes.
The picture can be erroneous due to the masking effect of mountains or hills.

TILT: Allows manual control of the antenna axis between -15° and +15° (deactivated in AUTO mode).
When set to 0° the axis of the antenna is aligned with the longitudinal axis of the aircraft irrespective of the pitch variations (stabilized antenna).
In AUTO mode, the tilt is automatically adjusted according to the altitude and distance of the aircraft to avoid ground echoes while optimizing detection of disturbed weather.
On the RDR-4000 installed on the A380 this control has disappeared.

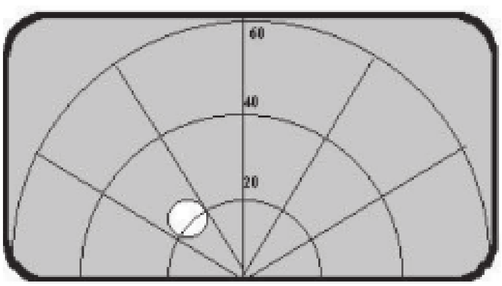
GAIN: Allows amplification of received echoes.
It is recommended to leave this control on CAL (Calibration, sometimes labelled AUTO) because playing with gain can largely inflate or underestimate the dangerous nature of the disturbed weather.

TEST: Allows testing of the radar system, the display screen shows all the associated symbology. During the test the radar is not actually transmitting.

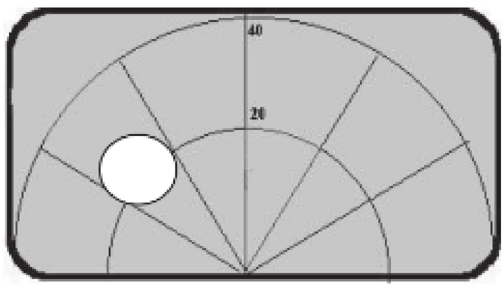


3.2.2 - Range selector

The screen displays the position of the echoes relative to the aircraft (distance and azimuth).
The pilot can select the area to be monitored with the RANGE control.
This control, depending on the type of radar, can be found on the radar control panel or for aircraft equipped with EFIS on the EFIS control panel, so that the radar display is in accordance with the scale of the navigation display (ND), because the radar echoes overlap the navigation display.
The display of the echoes on the radar is a relative bearing.
This relative bearing will be used to avoid weather or navigate towards the ground location returning an echo.
The distance at which an echo is seen is a slant range, and everything that was said in the DME chapter about indicated distance and ground distance applies also to the weather radar.



Range set to 60 NM

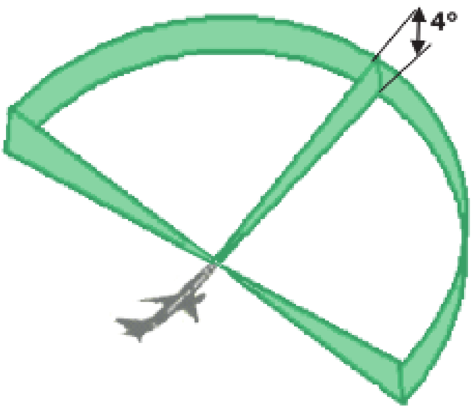


Range set to 40 NM

This control acts like a zoom.

3.3 - Coverage and range

3.3.1 – WXR mode



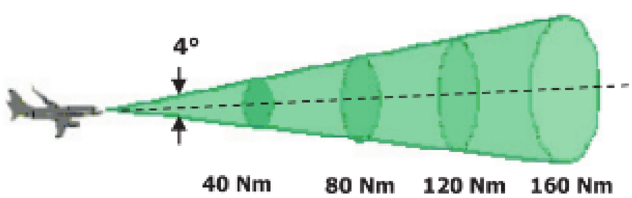
Azimuth coverage varies depending on the radar and can be 90,120 or 180°.
The radar scans a slice of altitude that corresponds to the opening of the radar beam.

The slice of altitude covered vertically by the radar will be greater when the range is increased. By using the rule of 1 for 60, with a beam of 4° (on the left) a range of 40 NM will cover 16 000 ft.

The greater the distance, the more the density of energy in the scanned area decreases. As a consequence, a dangerous cell displayed in red at 40 NM could eventually be seen in green at 160 NM.

To fix this problem, gain positioned on AUTO will increase with time in order to strengthen the distant echoes and display them with their real intensity.
In this way we have a heard echo level that does not depend on the distance but on the nature of the target.

It is therefore recommended to leave the gain on AUTO mode.



A radar with a range of 320 Nm and a correctly set tilt covers an area from the ground up to 60.000 ft.

The intensity of the returned echoes depends on the intensity and the nature of the weather conditions.

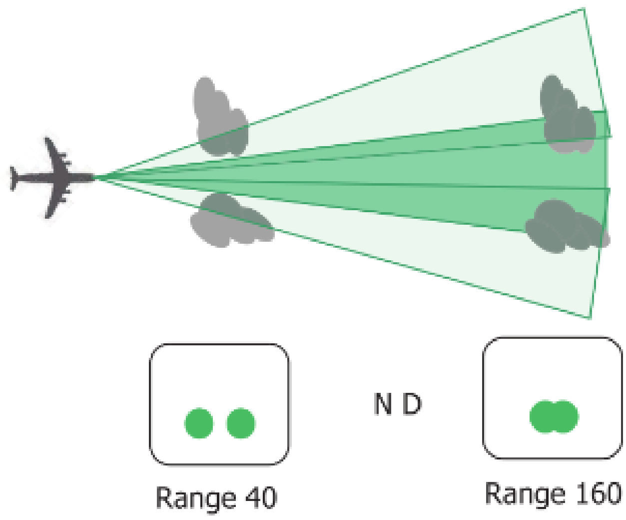
Reflective levels

Hail
Rain
Ice crystals
Wet snow
Dry hail
Dry snow

Black	None
Green	Light
Yellow	Medium
Red	Strong
Magenta	Turbulence

Fog and clear air turbulence (CAT) are undetectable.

For a modern radar like the Collins WXR2100 used as an example for this course, there is also a function Long Range Color Enhancement.
The picture of a faraway cell presented on the display is not just the reflection of the analog level received, but the construction of a picture by an algorithm based on the reception of many echoes that will fine-tune as we get closer to the cell.



Taking into account the beam width, the adjacent cells at close range are each in turn scanned by the beam, providing a good discrimination in azimuth (see display with range 40 NM).

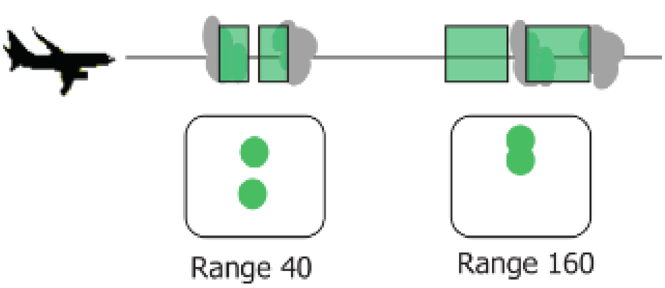
At greater distance the very enlarged conical beam scans simultaneously both cells without providing a separation (range 160 Nm).

The discrimination in azimuth depends on the width of the beam or in other words, on the directivity of the antenna.

In order to have a detectable echo level for faraway cells, they need to be illuminated with more power
To do that the radar varies the length of the pulses.
With narrow pulses to detect nearby cells, cells are scanned each in turn providing clear discrimination in distance.

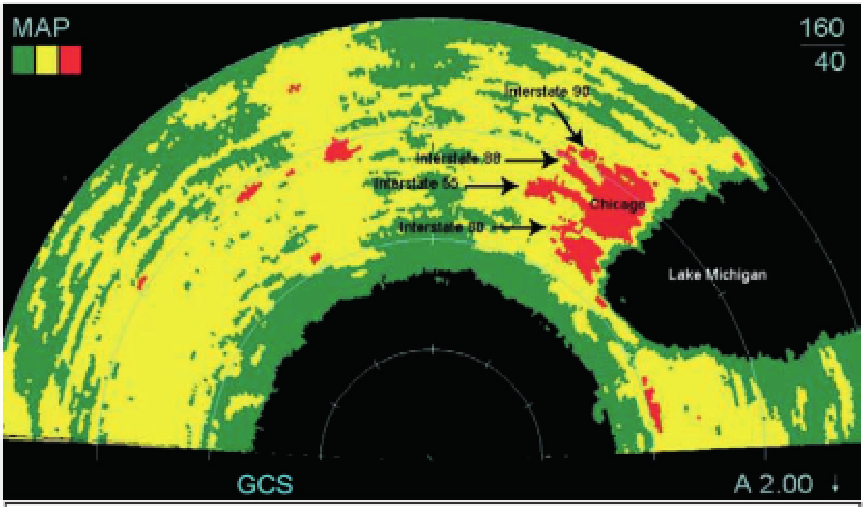
With large pulses, one pulse is still scanning the first cell when it hits the second one.
Both echoes become mixed up.

The discrimination in distance depends on the width of the pulses.



3.3.2 – Mapping mode

In mapping mode, the tilt must be adjusted to scan the ground.
The ground returns strong echoes, the gain must be reduced to avoid saturation of the screen.
The picture displayed is not intended for navigation but provides information on the surrounding environment (location of cities, coast, rivers, islands, mountainous areas).



MAP mode nearby Chicago
Collins WXR-2100 Multiscan ™ Radar operator's guide

Based on these considerations, the correct way to use the radar in manual mode during cruise is to select a range of 80 NM with a gain on MIN (to be adjusted later) and a tilt that provides ground echoes in the upper limit of the screen.

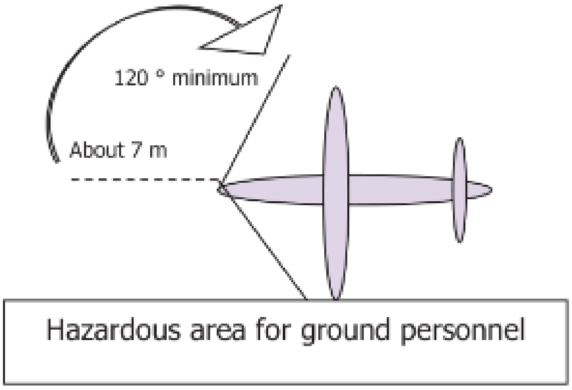
Memory:

Old radar control panel with manual selection of the beam shape (Pencil/Fan), for detection in Weather or MAP mode, and pulse width selection (Long/Short) for long range or short range detection.
On FAN the radiation law of the parabolic antenna is modified in cosecant squared, to widen the beam and therefore scan larger ground areas in MAP mode.

For information:
During taxiing the antenna stabilization could be selected to OFF.

3.4 - Errors, accuracy, limitations

Microwave frequencies at high power are dangerous for ground personnel scanned by the beam. The radar cannot be used in areas where there is ground personnel.



A metal obstacle which is highly reflective, for example a hangar, will return very powerful echoes, which can damage the entry circuits of the radar.
Radars should not transmit less than 10 meters from an obstacle, depending on the radars.
Modern radars have protection systems and their power for an equal range is smaller to that of the older generation.
Ex: P-708A radar for the A320, peak power: 150 Watts compared to the 75 KW for a radar of older generation.

It is also forbidden to operate the radar if an aircraft is refueling within a certain distance (30 to 60 meters according to the radar's manual) in the radar coverage area.

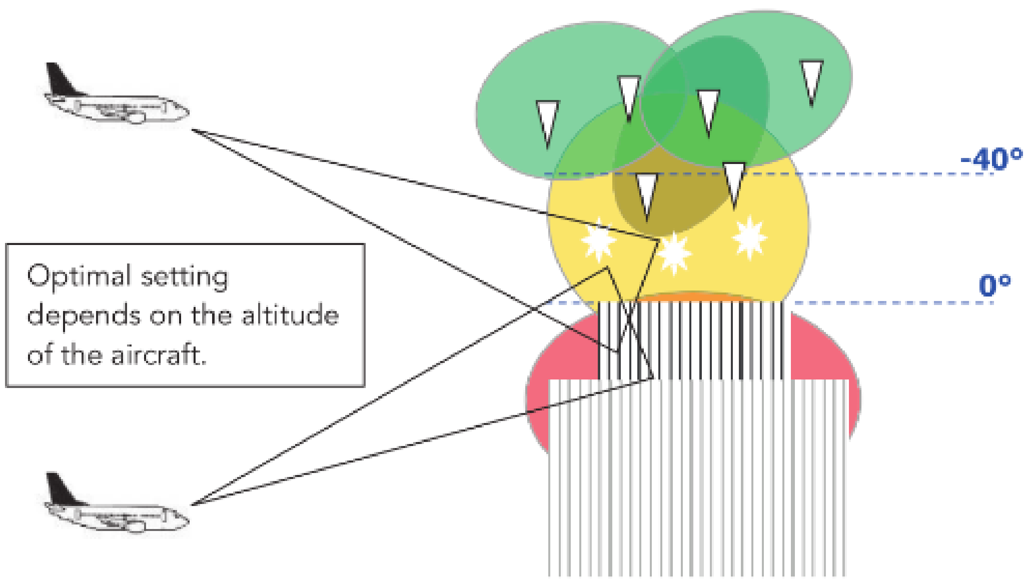
In general, the radar starts operating once the aircraft is lined up on the runway.

3.5 - Factors affecting range and accuracy

3.5.1 - TILT

In manual mode, setting the antenna TILT is critically important.
When set too low, the radar range is limited for detection of weather conditions and the display is polluted by ground echoes (*new radars have a Ground Clutters Suppression (GCS) function that eliminates unwanted echoes).*
When set too high, there is a risk of illuminating only the low reflective areas (areas far below the isotherm zero made up of dry hail) providing a green echo while very dangerous.

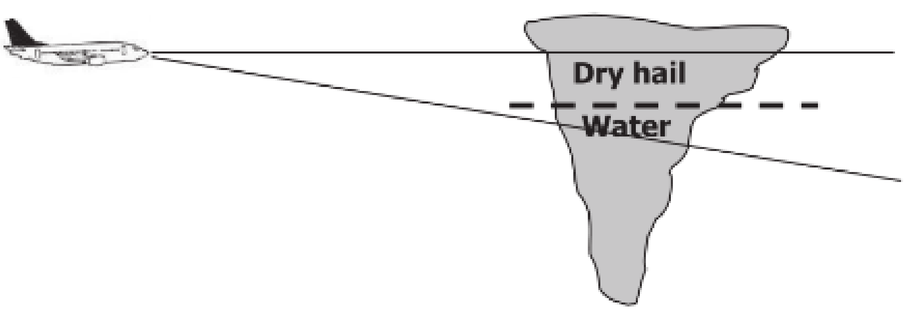
The setting must allow the visualization of the most reflective areas, taking into account the altitude of the aircraft and the selected range.



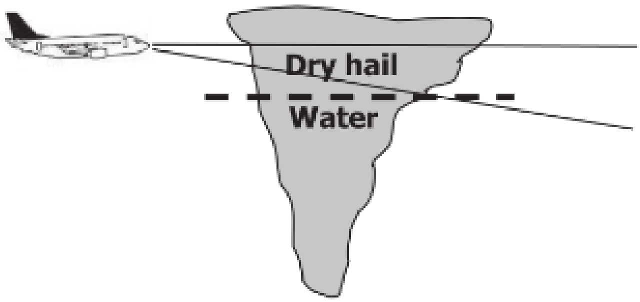
The TILT setting must be frequently adjusted while getting closer to the cell.

<u>Example of two situations:</u>

In the first situation, the radar returns an echo because the beam scans a rainy area.

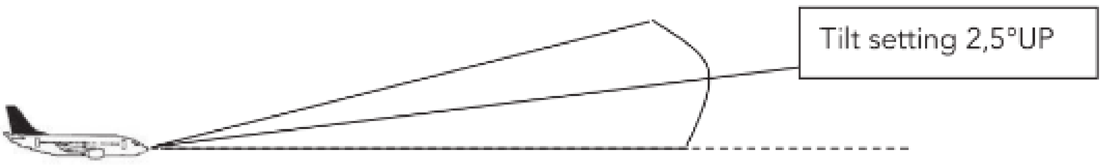


In the second situation, the aircraft gets closer, only the dry hail area is scanned and the absence of echoes gives a wrong interpretation of the situation (danger).

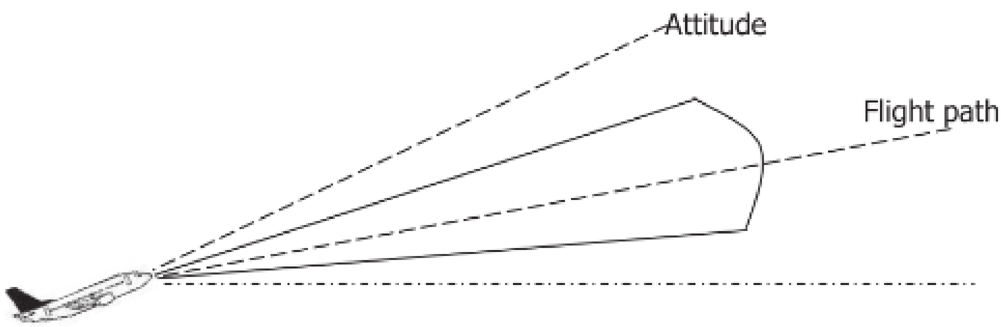


If there is no automatic tilt function it is necessary to readjust the tilt.

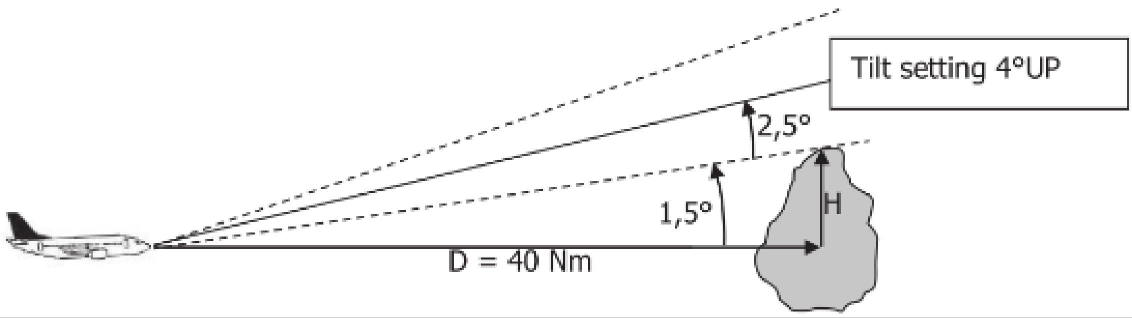
With TILT 0, the antenna is stabilized horizontally.
It is possible to "see", above or below the aircraft, by shifting the stabilization reference in pitch by more or less 15 degrees with the help of the TILT control.



Aircraft with a radar beam of 5°.
To see at its level and above the TILT is set 2,5° UP.



Aircraft with a flight path of 5° (aircraft attitude is 12°) wanting to see straight ahead and 2,5° above its flight path.
Beam width of 5° and TILT set up at +5° (flight path).



Calculation of the height of a cloud above the FL of an aircraft.
Given data:
A cell appears at 40NM;
Radar beam width is 5°;
With a tilt 4° UP if the cloud disappears, by what height does it surpass the aircraft FL?
The top of the cloud is seen under an angle of 1,5°.
Rule of the 1 for 60:
At 60 NM, for 1° H = 1 NM;
At 60 NM, for 1,5° H = 1,5 NM;
And so at 40 NM, and for 1,5°, H = 1,5 x (40/60) that is 1 NM or 6000 ft.

For 1° at 10 NM, 1000 ft at 60 NM, 6000 ft at 120 NM, 12000 ft, etc…

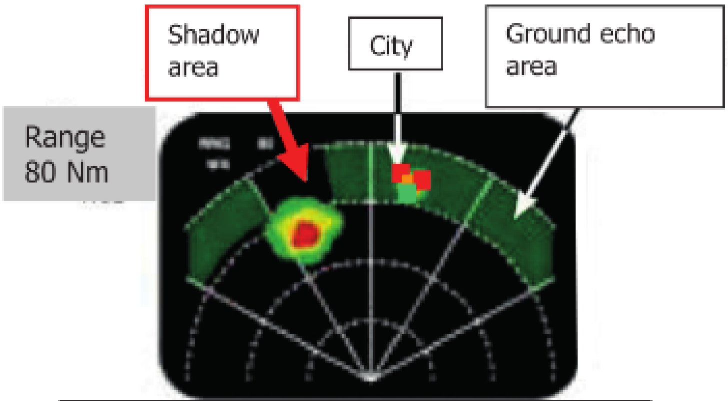
For use during cruise, the manufacturer recommends to select a range of 160 Nm for the PNF and 80 Nm for the PF.
This table gives TILT setting according to the range.

FLIGHT PHASE	DETECTION AND MONITORING PROCEDURES	COMMENTS
CRUISE	Use TILT slightly NEGATIVE to maintain ground returns on top of ND: Range 320 TILT≅1 DN Range 160 TILT≅1,5 DN Range 80 TILT≅3,5 DN Range 40 TILT≅6 DN In higher altitudes, closing weather: - Decrease ND - TILT down Use TURB to ISOLATE Turbulence - GAIN to AUTO.	No ground returns beyond line of view. Dnm = 1,23√ALT ft FL 370 D 240nm Poor ground returns over calm sea / even ground.

Extract from FCTM A330/A340

3.5.2 – Shadow area

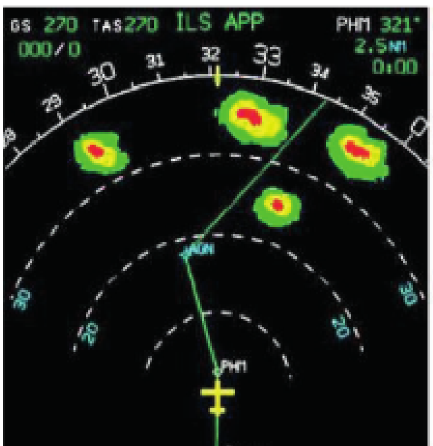
When the radar beam hits a very reflective target, all of the energy is reflected and there is not enough energy left to generate the echo of the targets located behind.
This is how a very dangerous cell can be masked by a closer cell that is less dangerous.
We have to pay special attention to those shadow areas when crossing disturbed weather. (This is the case in the summertime when there is a large thunderstorm area).
Flying through a cell followed by a shadow area cannot be considered.



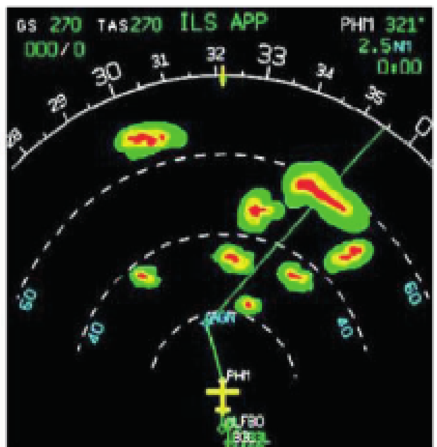
The shadow area shows in black.
No cell echo or ground echo.

3.6 – Navigation application

The radar is a tool that can detect areas of potential danger.
It is used to plan the route that will allow avoidance of dangerous area.
Anticipation is crucial.
160 Nm selected range
Vertical exploration of the disturbed weather with TILT in manual mode
Range 80 then 40 Nm getting closer (Gain AUTO, AUTO TILT)
Anticipation must allow to avoid flying into a dead-end situation.



Range that does not allow to see that the route leads to a dead end (blind alley).

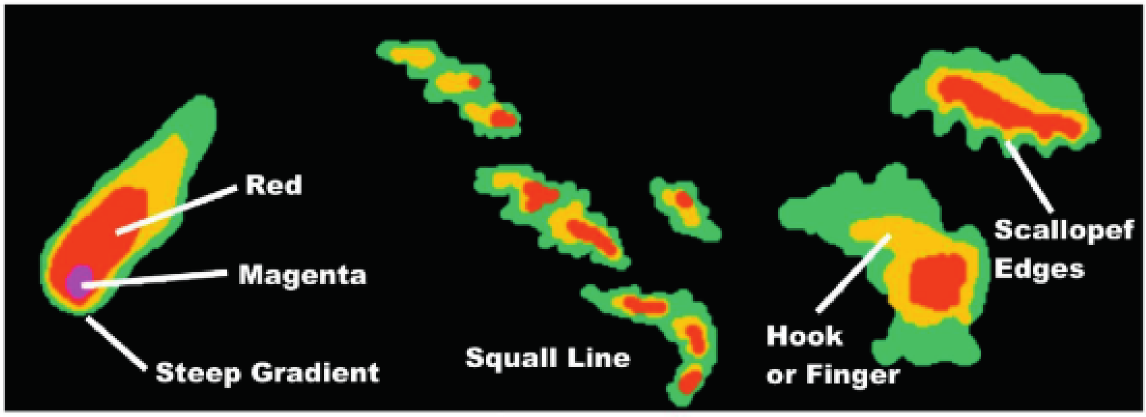


Range that allows to see that the route leads to a dead end (blind alley).

3.6.1 - Analysis of typologies of detected cells

Echoes are allocated different colors depending on their intensity.
The most dangerous areas are those having the most rapid change of color. (on the picture below where we go very quickly from green to yellow and then to red) or cells with very particular shapes (fingers, hooks, etc.) indicating hail.

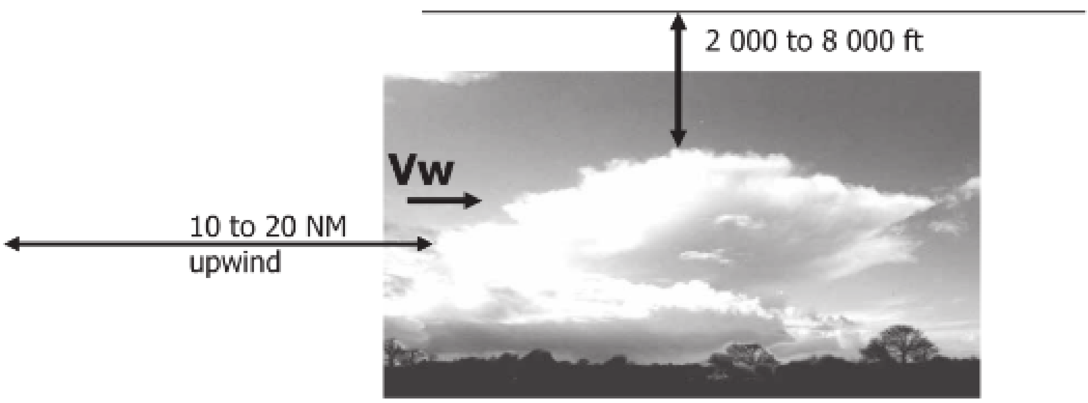
The turbulence is highlighted (on most equipment) by magenta areas.



Turbulence detection on new generation radars is achieved measuring by Doppler effect the velocity of hydrometeors in the cloud, the most turbulent areas being those where movements are most rapid.
This technique only allows close detection (40 Nm) hence the importance of paying attention to the color gradient.

3.6.2 – Avoidance of thunderstorms

Among the cloud formations, those giving the best echoes are those loaded with rain, hail or very wet snow (for example cumulonimbus).
Cloud cells returning strong echoes should be broadly avoided whenever possible.
At equal distance it is preferable to fly upwind.

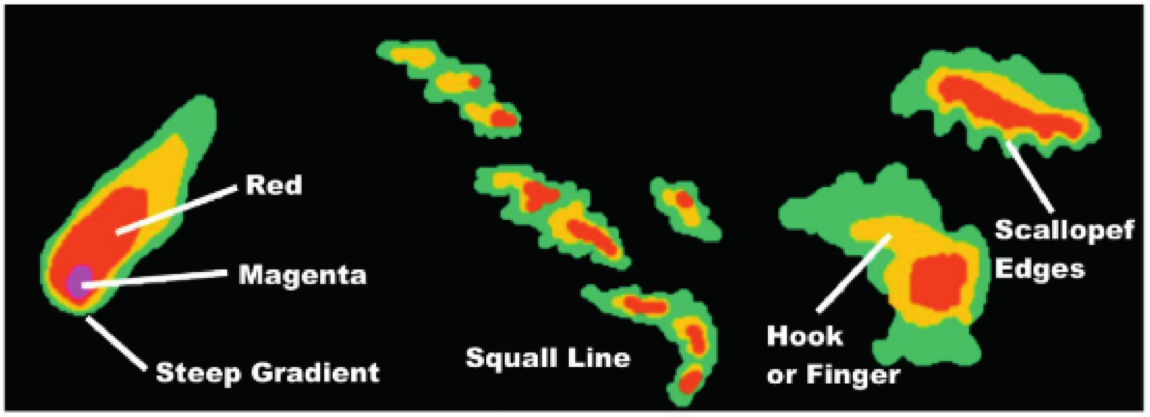


Because of kinetic energy, turbulence is present well above the cloud mass (up to 8000 ft).

3.6.3 – Mapping mode

Mapping mode based on a modification of the radiation law into squared cosecant is now obsolete.
Modern radars use the weather beam also in MAP mode.
The weather beam (3 to 5°) scans the ground, the echoes are memorized, and an algorithm creates a picture of the ground.
The use of ground echoes for navigation holds many traps such as the masking effect or certain mountains.

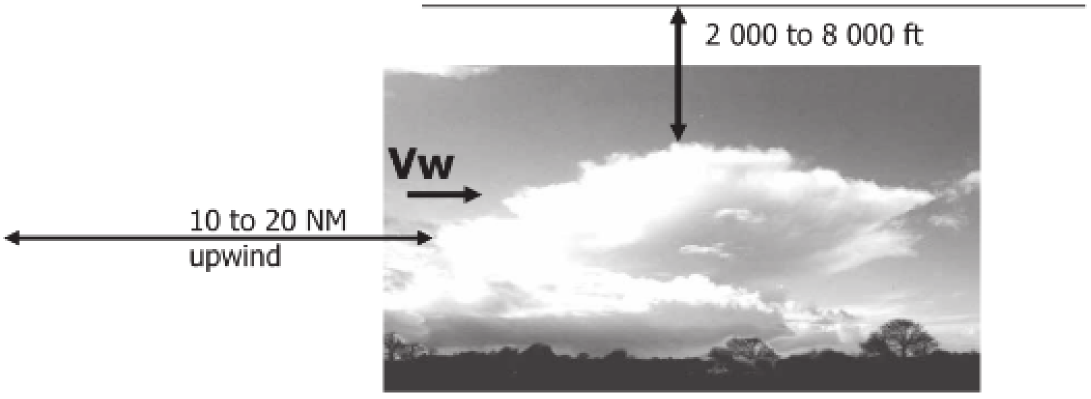
The turbulence is highlighted (on most equipment) by magenta areas.



Turbulence detection on new generation radars is achieved measuring by Doppler effect the velocity of hydrometeors in the cloud, the most turbulent areas being those where movements are most rapid.
This technique only allows close detection (40 Nm) hence the importance of paying attention to the color gradient.

3.6.2 – Avoidance of thunderstorms

Among the cloud formations, those giving the best echoes are those loaded with rain, hail or very wet snow (for example cumulonimbus).
Cloud cells returning strong echoes should be broadly avoided whenever possible.
At equal distance it is preferable to fly upwind.

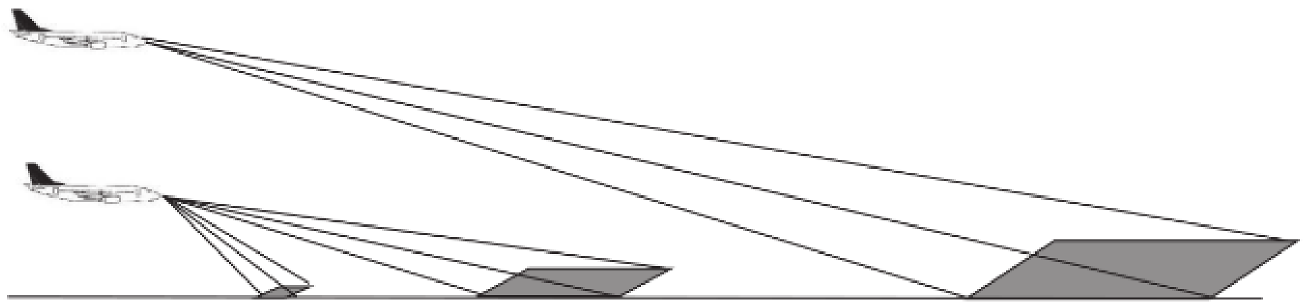


Because of kinetic energy, turbulence is present well above the cloud mass (up to 8000 ft).

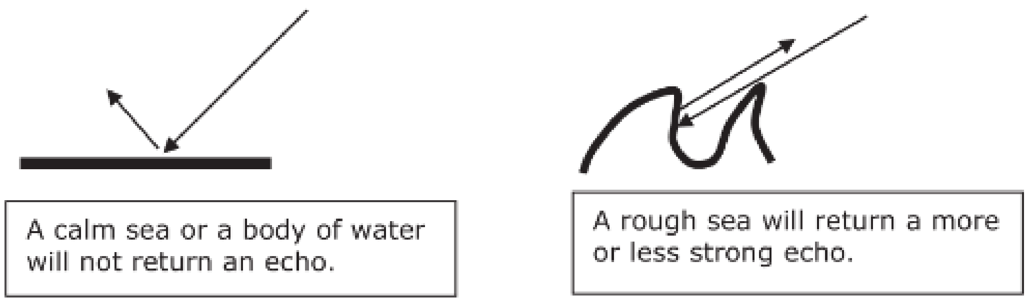
3.6.3 – Mapping mode

Mapping mode based on a modification of the radiation law into squared cosecant is now obsolete.
Modern radars use the weather beam also in MAP mode.
The weather beam (3 to 5°) scans the ground, the echoes are memorized, and an algorithm creates a picture of the ground.
The use of ground echoes for navigation holds many traps such as the masking effect or certain mountains.

With the beam in weather mode, the detected ground area is weaker when flying at low altitude with a tilt selected low.
Gain is controlled manually in MAP mode.



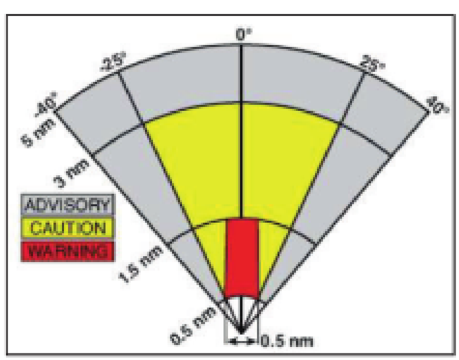
A common feature used is the display of the cutout of a coast because the contrast between ground and water is very clear.



3.6.4 – WINDSHEAR detection

Modern radars can now detect windshears on the condition that they have reflective moving particles.
By Doppler effect the speed of the particles is measured and memorized at different altitude levels and the path of the particles is recreated with the use of their speed vectors.
The variations in direction of these speed vectors and the speed gradient allow the prediction of windshear.
The predictive windshear function (PWS) is automatically activated during takeoff and on approach below a certain altitude (about 2 300 ft), even if the radar is off.

The predictive windshear function will deliver ADVISORY, CAUTION (monitor radar display) or WARNING (wind shear ahead) messages depending on the location of the phenomenon and not on its intensity.



ATTENTION: we are here in the presence of a predictive wind shear (PWS) and not a reactive wind shear (RWS) as studied with the GPWS / EGPWS.

To illustrate this: with the PWS you are warned that you will encounter a problem, with the reactive wind shear, you already have the problem.

04 SECONDARY SURVEILLANCE RADAR (SSR)

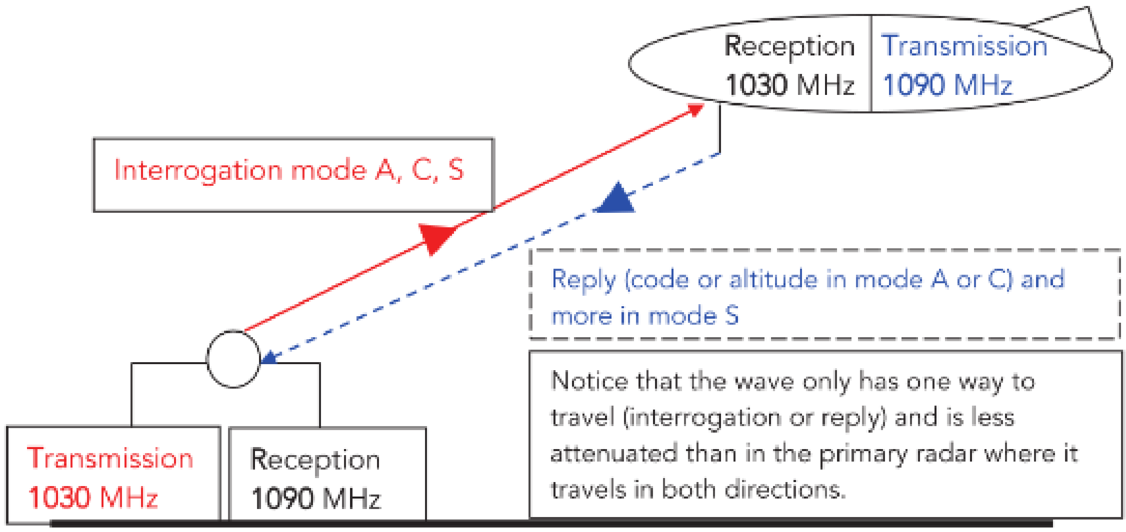
4.1 - Principles

Further in this course we will use terms SSR or secondary radar (implying surveillance).
There are basic secondary radars, also known as mode A or C based on analog technology, and mode S secondary radars, based on digital technology.
Supplying a picture that represents the air situation, in a given area, allows ~~the~~ air traffic control to ensure the safe conduct of aircraft in the area under its responsibility.
Since 1960 two means are used to generate this picture:

- the primary surveillance radar;
- the secondary surveillance radar.

The function of both is to detect and position the aircraft but if the primary radar, based on the echo principle, provides a great accuracy on the azimuth and distance information, it does not provide any information allowing identification of the aircraft.
The secondary radar, originally conceived for identification information through an embryonic dialogue between the ground and the aircraft, provides more information.
It suffers, in its basic form, of deficiencies such as garbling studied later and must be associated to the primary radar.
The secondary radar is an aid to air traffic. It complements the primary radar.
The technique in use enhances the echoes by providing information that cannot be acquired by the primary radar and provides safer identification on the display screen.
In its more evolved form (mode S), it is part (coupled to a TCAS computer) of the anti-collision system.
The mode S allows ground-air or air-ground datalink communication with ATC, or the transmission of information to other aircraft as part of the ADS (Automatic Dependent Surveillance) system.

The SSR (Secondary Surveillance Radar) works in the UHF band.
Unlike the primary radar, it transmits and receives on different frequencies and does not use the echo principle.
A ground station **transmits interrogation pulses,** according to a certain mode, in a well defined direction, on one frequency, and **listens to the answer on another frequency.**
A **transponder, located onboard** an aircraft in that direction, **catches the interrogations and replies by transmitting a coded information** that contains certain data (a code assigned by ATC or an altitude, among others).



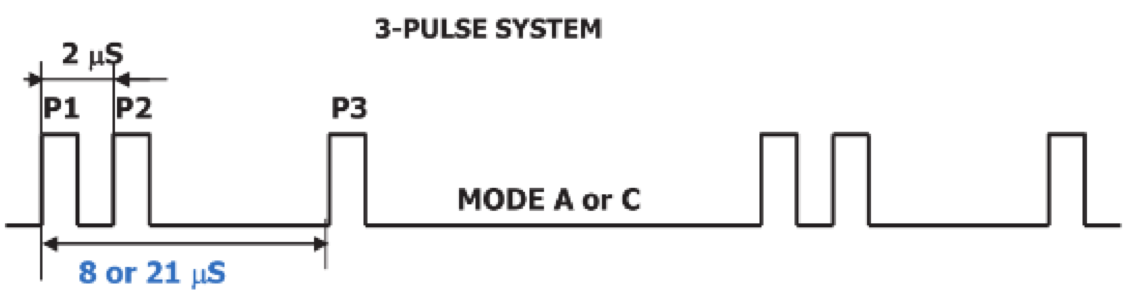
4.2 - Modes and codes

A transmitter equipped with a rotating directional antenna transmits series of pulses that interrogates surrounding aircraft.
Each aircraft is interrogated sequentially during one rotation of the antenna.
The transmission frequency is 1 030 MHz.
The interrogation mode defines the content of the requested reply which will be, either an identification, the transmission of a code in the form of 4 digits (mode A) that the pilot has setted upon request from ATC, or an altitude (in mode C) and other parameters (in mode S).

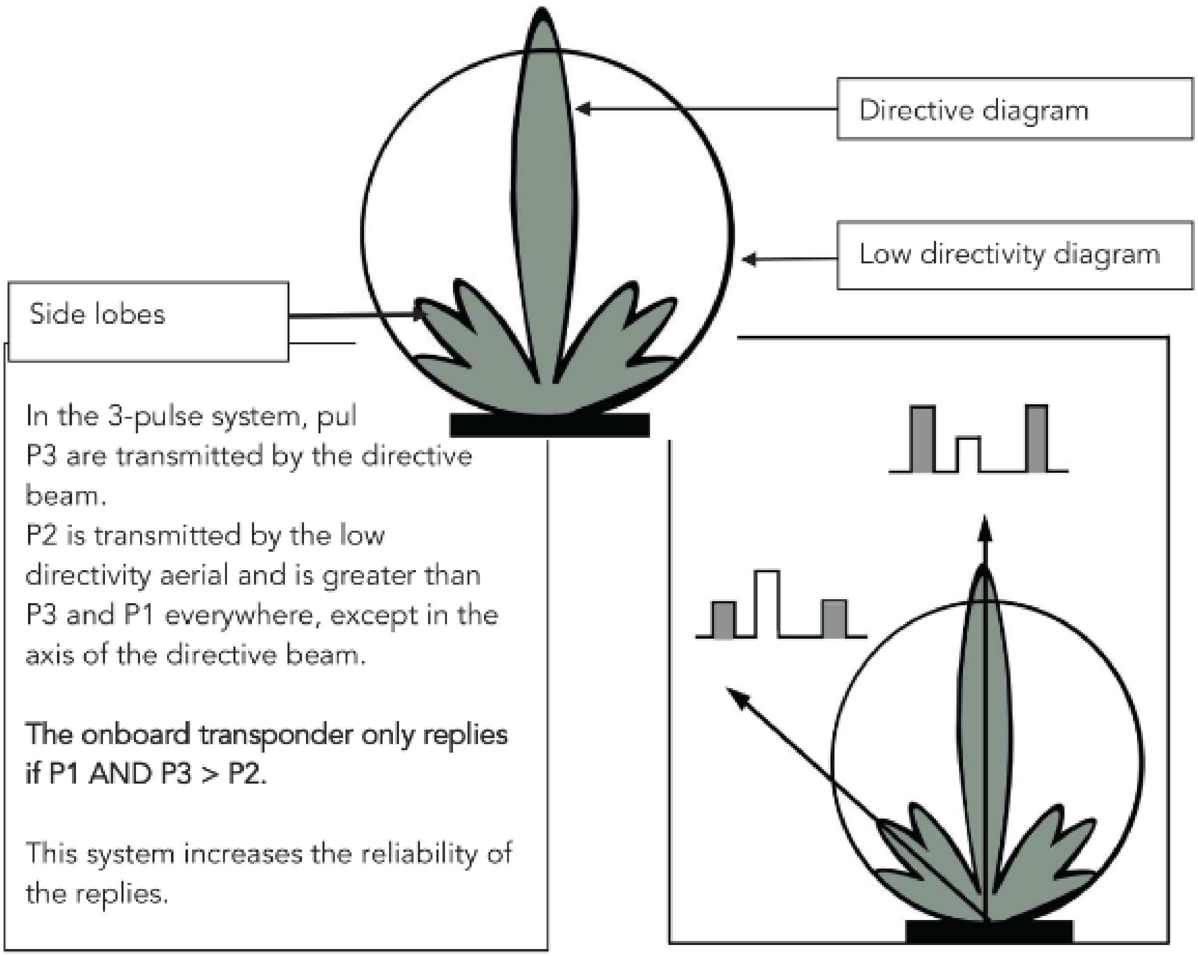
4.2.1 - Mode A or C interrogation

The transmission of the interrogation is achieved by transmitting 3 pulses.

The pulse spacing, 8 µS for mode A, 21 µS for mode C, defines the interrogation mode.
There are other modes which we will not discuss here because they are not used in civil aviation at present (they were planned for future implementation).



The directivity that is needed for the transmission antenna generates side lobes that can unintentionally trigger the inboard transponder upon interrogation by a side lobe which does not correspond to the pointing azimuth of the antenna.
To overcome this, a diagram with low directivity is generated by a control track to mask the side lobes of the directive transmission.



3-PULSE SYSTEM

Transmitted pulses P1 and P3 contained in the directive diagram are called interrogation pulses.
Pulse P2 contained in the low directivity diagram is called **control pulse.**

4.2.2 - Mode A reply

Onboard the aircraft, a secondary radar responder commonly called airborne transponder produces the reply.
The transponder is a transmitter-receiver set, associated to an omnidirectional antenna because it has to be able to pick up the interrogation and send a reply irrespective of the ground station azimuth.
The receiver listens on 1 030 MHz and the transmitter transmits on 1 090 MHz.

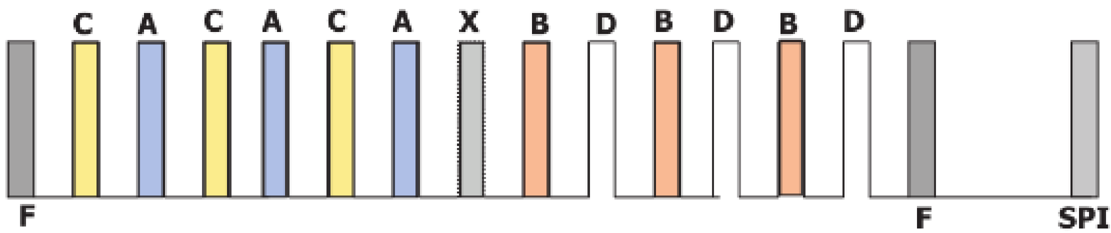
The reply consists of a frame of several pulses.
It is conditioned by the interrogation mode transmitted by the ground station.
As with all communication protocols it is necessary to know where the information starts and where it ends.
For this, the pulses that constitute the reply are transmitted between two framing pulses.

The mode A reply is made up of 12 pulses, transmitted between framing pulses.
The pulses have only two statuses, present or absent.
This allows for 2^{12} = 4 096 different combinations.

Therefore, ATC can assign to the aircraft 4 096 transponder codes ranging from 0000 to 7777.

The code assigned by ATC is a 4-digit number that we will rate "A", "B", "C", "D".
Each pulse corresponds to a coding element of these 4 digits "A", "B", "C"," D", and the position of the pulse in the thread determines the weight of that pulse.
There are 3 "A" pulses with a weight 1, 2, 4.
These pulses are rated respectively A1, A2, A4.
The sum of the weights of the A pulses is the "A" digit.

Therefore, if:
Only A1 is present, A digit is 1;
A1 and A2 are present, then A=3, (1+2);
A1, A2, A4 are present, then A=7, (1+2+4);
This way we can make all the combinations of A, from 0 to 7.
The same will be done with the pulses of digits "B", "C", "D".

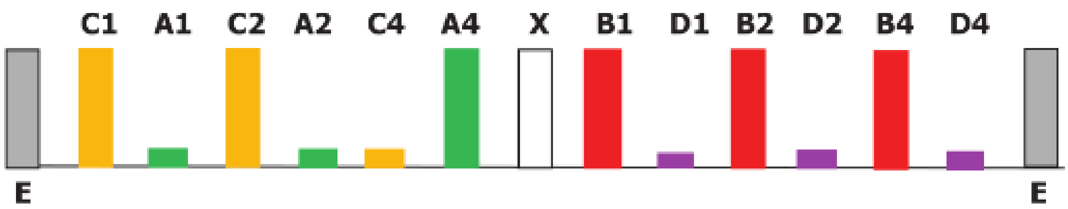


Format of a mode A reply

If all pulses are "present", the transponder code is then 7777.
The X pulse is not used.
It was intended for a system implementation.
(For example: X present, gear down and locked, X absent, gear up).

Upon request from ATC, the pilot can press a button, sending an additional pulse (Special Pulse Identification) after the last framing pulse.
This allows ATC to visualize a flashing target on the radar display, showing a particular aircraft.

Without specific request from ATC, this special pulse is sent every 20 seconds.



A1	A2	A4		A
0	0	4	=	4

B1	B2	B4		B
1	2	4	=	7

C1	C2	C4		C
1	2	0	=	3

D1	D2	D4		D
0	0	0	=	0

ASSIGNED CODE = 4730

We use here OCTAL coding (3 binary pulses represent 1 decimal digit).

The highest possible digit is 7.

There is no code such as 4718 or 4719.
We will have 4716, 4717, and then 4720.

Example of a mode A reply

4.2.3 - Mode C reply

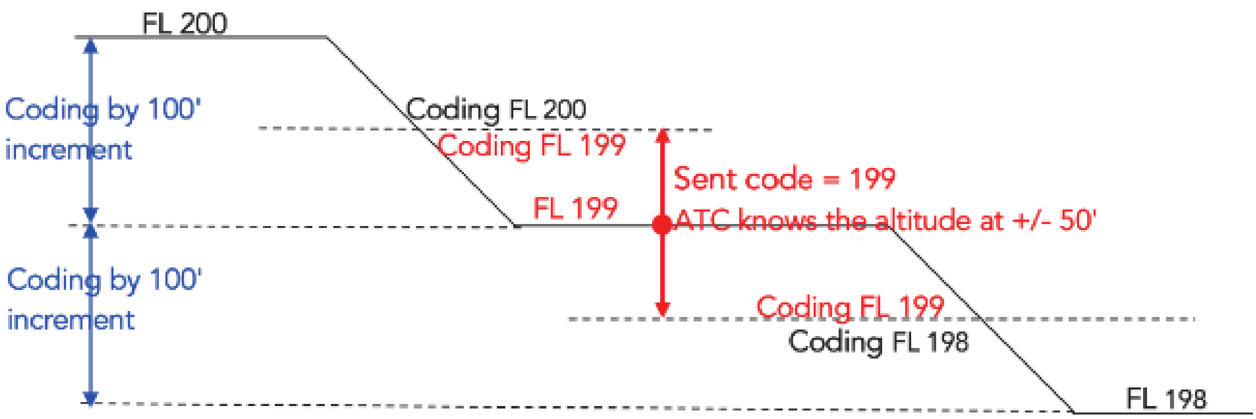
When the ground station, by spacing pulses by 21 µS transmits a mode C interrogation, the onboard transponder produces a reply that provides the altitude of the aircraft based on 1 013,25 hPa.
The reply is independent from the setting of the onboard altimeters.

The protocol in mode C is identical to mode A, with the difference that only 11 pulses are used (D1 and X are always set at zero).
(Source: EUROCAE Document ED-26: Table 13: Altitude Encoding Transition Points).

Here we use coding which is not octal, but a modified Gray code called Gillham code (which codes with 11 bits).
A question that I am frequently asked is "how does the ground station know that it receives a mode C rather than a mode A reply, and does not confuse octal coding with Gillham code (defined by ICAO Annex 10 volume IV) that allows coding digits from 0 to 9 (ex: FL189) as opposed to the octal coding"?
In fact, the ground station ASSUMES that after a mode C interrogation it receives a mode C reply, so it is quite basic.

The combination of 11 pulses allows the use of 2 048 possible altitudes.
By 100 feet increment this represents a range of over 128 000 FT which is sufficient to transmit pressure altitudes from -1000 to +126 750 FT.
The system is arranged like this:

An aircraft leaving FL 200 will send a FL 200 until 19 950 FT and then from 19 950 FT, a FL 199.
ATC knows the altitude of the aircraft based on 1 013,25 hPa within ± 50 FT.



Note:
An exam question states that there are 4 096 codes in mode A and mode C.
This answer is yes, BUT since only 11 pulses vary in mode C, we use 2 048 different codes among 4 096 available in mode C.

4.2.4 - Mode S interrogation

Mode S (for Selective) is the last improvement of the secondary radar, it allows general (All-Call) or selective (Roll-Call or Selective Call) interrogations (for one out of numerous aircraft).
It also allows additional data exchange, other than the code and the altitude (mode A or C), such as speed, vertical speed, heading, flight ID, etc…
The transmission of altitude in mode S can be done with 25 ft intervals if the altimeter equipment permits it.
We recall here that the mode S transponder, when coupled to a TCAS computer transmitter broadcasting mode S interrogations, is part of the collision avoidance system because of its capacity to transmit this data.
Data is coded numerically in specific formats with strings of 56 or 112 bits.

Differential phase shift keying (DPSK) is used to code 0 and 1 in the interrogation.
In this kind of modulation, 1 is defined by a synchronized phase inversion at the start of the data and 0 maintains the phase (see GPS modulation).
The mode S reply uses another type of modulation called pulse-position modulation (PPM).

In mode S each transponder has a unique address (mode S address) coded in 24 bits stored in the hardware (ROM).
In theory, the mode A limitation of 4 096 codes does not exist anymore.
Exceeding the address capacity (16 777 214 possible addresses) would require 1 transponder per 500 world inhabitants.

Mode S selective address format

X bits determine the state of registration of the aircraft.	A combination of (24 minus X) bits is allocated to each aircraft, civilian or military.

24 bits

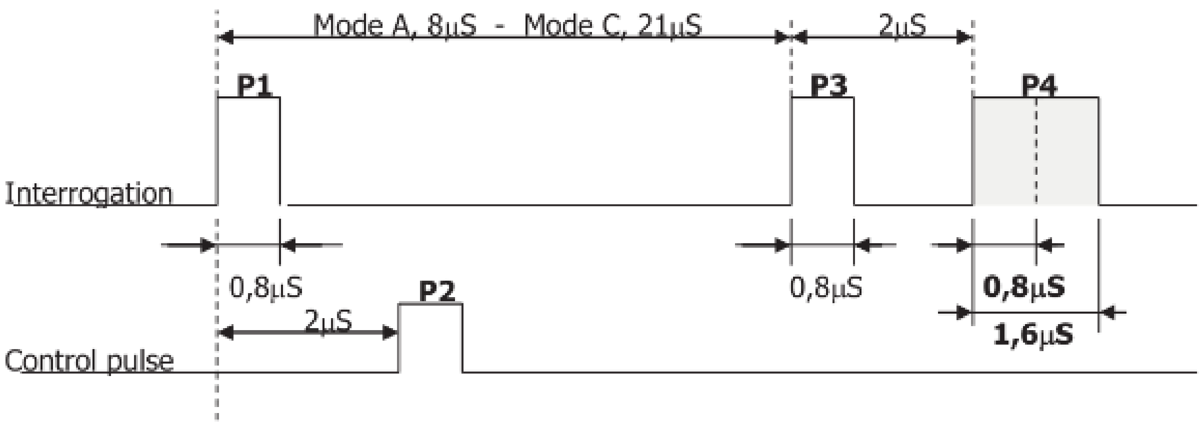
Once ATC knows an aircraft mode S address it can interrogate it specifically and thus avoid the multiple replies of the surrounding aircraft.

But mode S for now must be compatible with the classic mode A and C because a ground station with classic mode A or C must be able to interrogate an aircraft with mode S, and reciprocally an aircraft with classic mode A or C must be able to reply to a ground station mode S, hence the inter mode interrogations.

a) Inter mode interrogation

This is a juxtaposition of modes that creates compatibility (a ground station mode S interrogates in classic mode between the mode S periods).
In inter mode, there is a transmission of pulses P1 and P3 which are heard by the classic modes and **an additional pulse P4, of 0,8 or 1,6 µs, which is not perceived by the classic mode A and C transponders.**

Inter mode All-Call



This allows these interrogations:

Modes A/C only All-Call:
If P4 is short, only the classic transponders reply in mode A or C because they receive P1 AND P3 higher than P2.
Mode S is blocked by P4.

Modes A/C/S All-Call:
If P4 is long, everybody replies (A, C and S) because the classic transponders not detecting P4 still detect P1 AND P3 higher than P2, and the mode S transponders detect a P4 pulse (long) letting them know that the message is also for them.

It would be advantageous here to combine classic call and mode S call in the same general call, and thus gain time that could be used for selective calls. But this method has drawbacks (no need to expand).

b) Pure mode S interrogation

Mode S interrogations can be:

- All-Call only mode S interrogations;

The ground station interrogates all surrounding mode S aircraft and they send back their mode S address.
To achieve this the ground station interrogates the aircraft with the mode S address 1111 1111 1111 1111 1111 1111.
This address does not belong to any aircraft but is understood by all and so all the mode S aircraft reply by sending back their personal address.
These addresses are memorized by ATC.

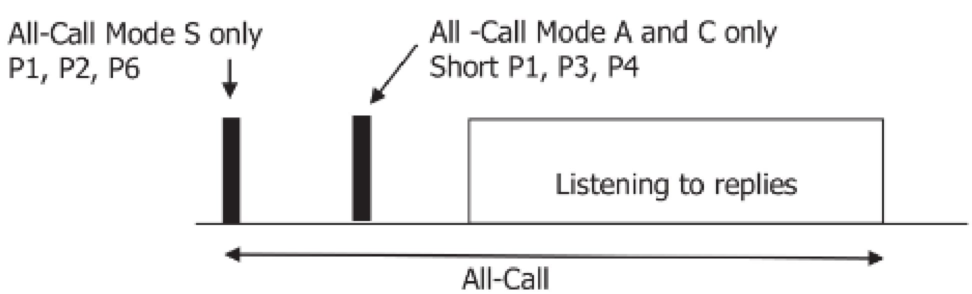
- Selective (Roll-Call) interrogations;

The ground station having memorized all the mode S addresses of the surrounding aircraft interrogates only one aircraft at a time with the use of its mode S address.

In pure mode S interrogation (selective), classic transponders must be prevented from replying.
This is done by transmitting a P2 pulse at the P1 level (classic transponders believing that they are interrogated by a side lobe do not reply).
A P6 pulse conveying **a message string of 56 or 112 bits** is then transmitted, which allows setting up a dialogue with the aircraft:

The mode S interrogation also contains-the ground station identifier code (IC) in the P6 string.

ICAO standards do not recommend the use of All-Call with a long P4 pulse, but rather a sequence All-Call mode A and C followed by a sequence All-Call mode S.



Between the All-Call (AC) periods are inserted the Selective Calls (SC) periods.

4.2.5 - Mode S replies

The mode S reply is a string made of a preamble of 4 pulses and a data block of 56 or 80 bits.

These strings can be linked to form ~~the~~ messages that support the ground-air datalink (UF, Uplink Format) or the air-ground datalink (DF, Downlink Format).

Strings are classified in three different format categories:
Surveillance format (56 bits);
Short communication format (112 bits of which 50 are for data);
Long communication format (112 bits of which 80 are for data).

There are 25 different formats (from UF00 to UF24 and DF00 to DF24).

Some UF or DF formats are exchanged from aircraft to aircraft as part of the collision avoidance system.
One aircraft plays the role of the ground station, the other (the target) keeps the role of the aircraft.
The mode S transponder is also used as the support of a new surveillance mode: ADS-B.

4.2.6 - ADS-B

The ADS-B (Automatic Dependent Surveillance – Broadcast) is, among other things, an air traffic surveillance system that can considerably improve ATC surveillance and control capacities at a lower cost because it does not require a ground-based interrogation transmitter.

ADS-B:

Automatic: No pilot action;
Dependent: Depends on the GPS or other aircraft sensors;
Surveillance: Position, altitude, speed etc…;
Broadcast: Broadcast to all (aircraft and ground station).

The principle is that each equipped aircraft broadcast automatically at regular intervals of 1 second the aircraft parameters (DAPs – Downlink Aircraft Parameters) by means of digital data formats that contain among others:

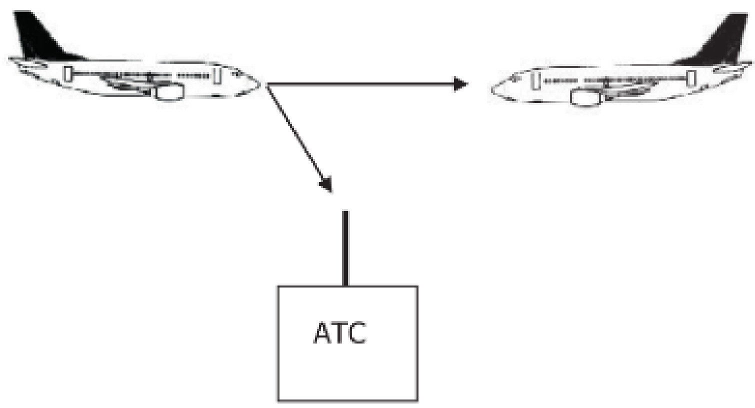
The 24 bits address of the mode S transponder (identifier that is unique in the world);
The flight number;
The latitude and longitude position of the aircraft based on a GPS position;
The altitude, the speed, the vertical speed, and the flight path.

This data is not fixed in time and does not have an identified recipient (broadcast).
Data is broadcasted to the ground and to the other aircraft, which will process, or not, the data according to their needs.
Data is transmitted on 1 090 MHz by an improved mode S transponder sometimes called 1090ES (for mode S on 1 090 MHz, Extended Squitter).

The term Extended comes from the fact that the message is broadcast with the long communication format (112 bits) of the basic mode S transponder.

Note that this system must also address general aviation and that the FAA has chosen an equipment called Universal Access Transceiver (UAT) transmitting in the DME band (978 MHz) however, Europe is moving towards a new path with no interoperability between the systems. To be followed…

The data is received by a ground station that feeds the control centers which can follow traffic more precisely than with a radar.
It must be pointed out that with this system, there is no more interrogation from the ground and therefore there is no need for a radar (in theory).
Aircraft can be equipped with a receiver that can decode the formats from the other aircraft and with a display system, they can know just like ATC, the position of the surrounding aircraft (much beyond what the TCAS allows).



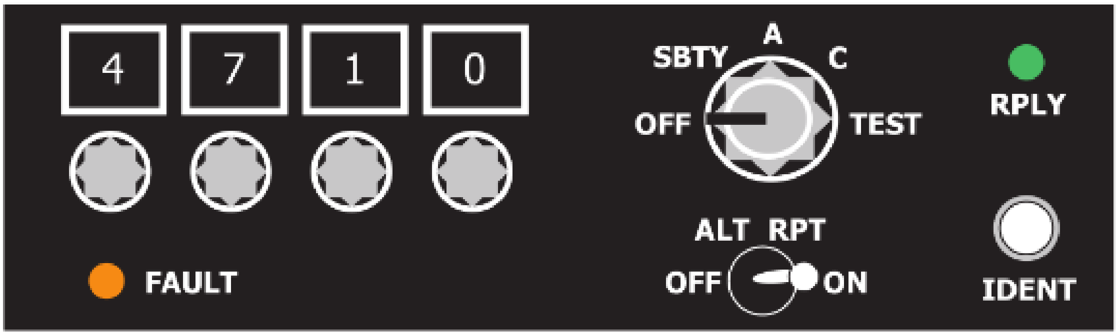
We are making our way towards the concept of free flight.

Nevertheless, first of all the concept must be validated and must answer quite a few questions: Which tools to use? Which applications to set up? What regulations to apply…?

One of the main constraints of ADS-B is its dependence on onboard navigation information and the danger that could result from the transmission of erroneous information that has not been detected.
One of the intended solutions is the use of multilateration in order to validate, via an external mean, the position information.
(See notes outside of this course on multilateration).

4.3 - Presentation and interpretation

4.3.1 – Transponder control panel



This is made up of:

- 4 digit selectors to display the assigned code;
- A mode selector knob;
- A control light (RPLY), illuminated for each response or during the TEST;
- A fault monitoring light (FAULT);
- An IDENT push button (sometimes labelled SPI) for the pilot, to be pushed upon ATC request;
- A selector knob that allows turning off altitude reporting in mode C.

If there is more than a 200 FT discrepancy between the altitude of the aircraft and the one reported by the transponder, the ALT RPT (Altitude Report) knob will be selected to OFF
On transponders that do not have this knob, the pilot will select 0000 to inform of an altitude reporting problem (mode C failure code in the UK).

Control panel of a mode S transponder.

It is possible to have, like here, a unique control panel for two transponders; the SYS knob (on the left) allows selection of the operational transponder.
Normally the transponder on the pilot flying side will be selected for reasons outside of the 062 course regarding navigation in RVSM.

We find the IDENT control, the monitoring light (ATC FAIL) and the mode selector knob including the TCAS function.

- ALT RPT OFF: mode C with altitude reporting turned off.
- XPNDR: mode S is active.
- TA ONLY: mode S and TCAS are active *(RA are inhibited, see TCAS course).*
- TA/RA: mode S and TCAS are active.

In normal operation, the selector knob is on TA/RA.

The toggle switch ABV-N-BLW is for the TCAS only and will not be explained here.

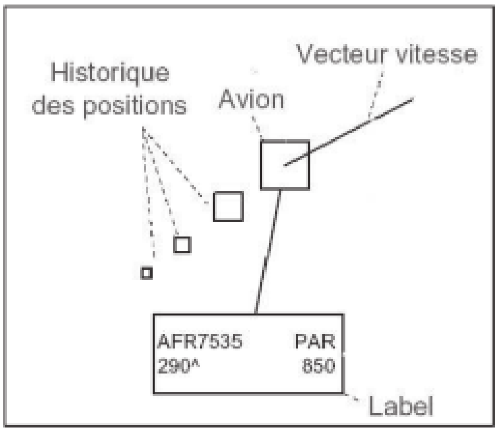
Pushing on IDENT (SPI) adds in the reply format a special identification pulse which, after takeoff, highlights the label of the corresponding aircraft on the ATC display.

A keyboard or rotating knobs allow the display of a 4-digit code assigned by ATC.
This code, with the help of the ATC computer system, allows a correlation between the flight ID and the flight plan CAUTRA number (in France).
Some outstanding codes are used at the initiative of the pilot in certain circumstances to report the situation to ATC.

Let's mention, 7500 = aircraft hijacking, 7600 = radio failure, 7700 = emergency.
As soon as one of these codes is displayed onboard the aircraft, the radar label has an additional indication (7700, EMRG or 7600, RDOF).

4.3.2 - Presentation of the information

The information is now displayed on flat screens similar to desktop personal computers and is much richer than before.



We see the aircraft, its previous positions (thus its path), its speed and speed vector (extrapolation of its path), its identification (AFR7535), its flight level (290 climbing) and more if ATC chooses to and the aircraft equipment allows it.
ATC can do primary surveillance or augmented surveillance.

4.3.3 – Elementary surveillance (ELS)

Elementary surveillance provides ATC with the position, the altitude and the identification of the aircraft.

The minimum required transponder for elementary surveillance is a Level 2 mode S transponder, which allows automatic transmission of the Flight Identity (FI) according to the ICAO format, the only one compatible with ground Air Traffic System (ATS).
This format comprises the ICAO operator code (3 letters) followed by the flight number, or by default the aircraft registration.
This transponder must be able to manage the identifier codes of the radar stations (SI for Surveillance Identifier).

This transponder must also operate antenna diversity if it equips aircraft with a maximum mass in excess of 5 700 kg or a maximum cruising airspeed in excess of 250 kts.

4.3.4 – Enhanced surveillance (EHS)

Enhanced surveillance consists of extracting the additional parameters of the aircraft from the exchanged data. This is called Downlink Aircraft Parameters (DAP).

Parameters that can be sent are:

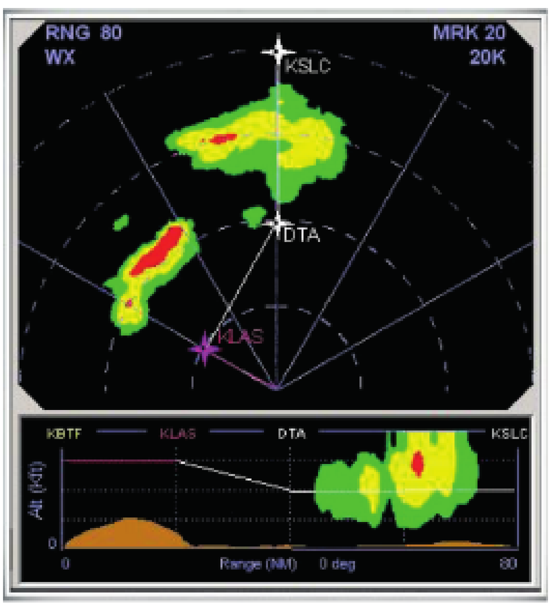
- Magnetic heading;
- Indicated airspeed or Mach number;
- Vertical speed;
- Selected altitude;
- Ground speed;
- Roll angle;
- Track and track variations.

The information sent to ATC is improved while radio communications are considerably reduced. The lower workload allows ATC to increase safety and regulate traffic better, and pilots to reduce potential errors.

Notes not included in the primary radar course

3D radar

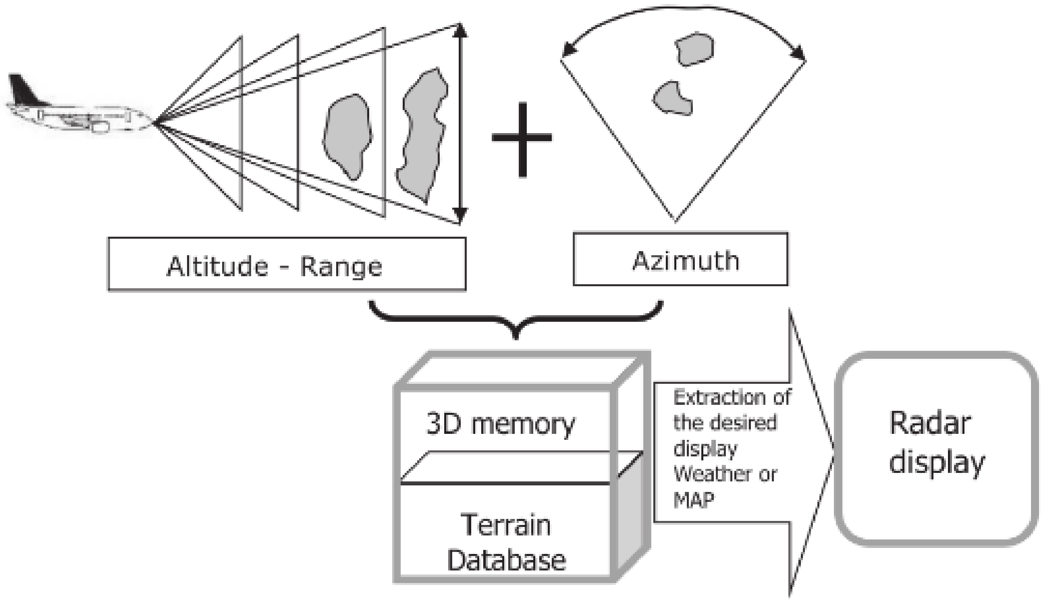
3D radars continuously scan all altitude bands, memorize the received echoes (altitude, distance, intensity) in a 3D memory and provide upon request the data from the 3D database to the display (in the selected range), with the addition of a vertical profile.



RDR 4000 display

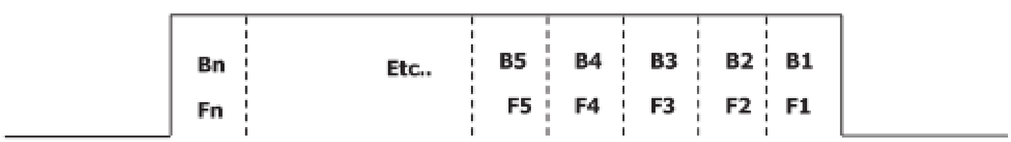
A380

Equipped with a terrain database, it can determine the ground echoes and delete them from display when necessary.



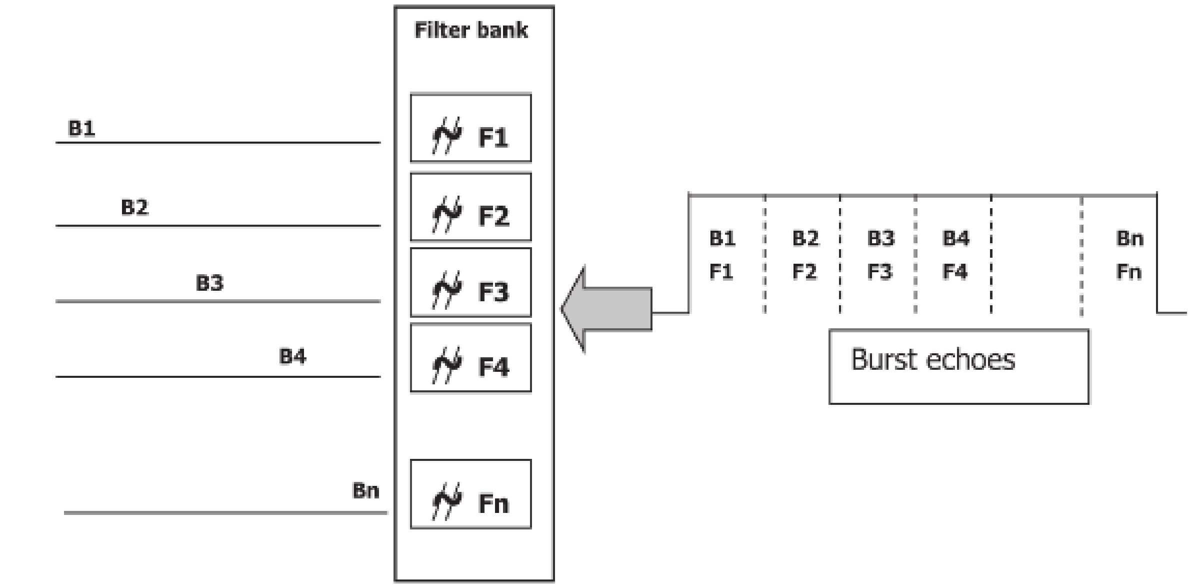
pulse compression.

We have seen that in theory, the radar pulse must be of a short duration in order to limit the blind area and increase the range discrimination.
One of the disadvantages of these pulses is that they release less energy than a long pulse.
We now use, in addition to the short pulses, long pulses and a technique called pulse compression
The (gross) principle is as follows:
The long pulse is, in fact, a succession of bursts (B) transmitted with different frequencies (F).

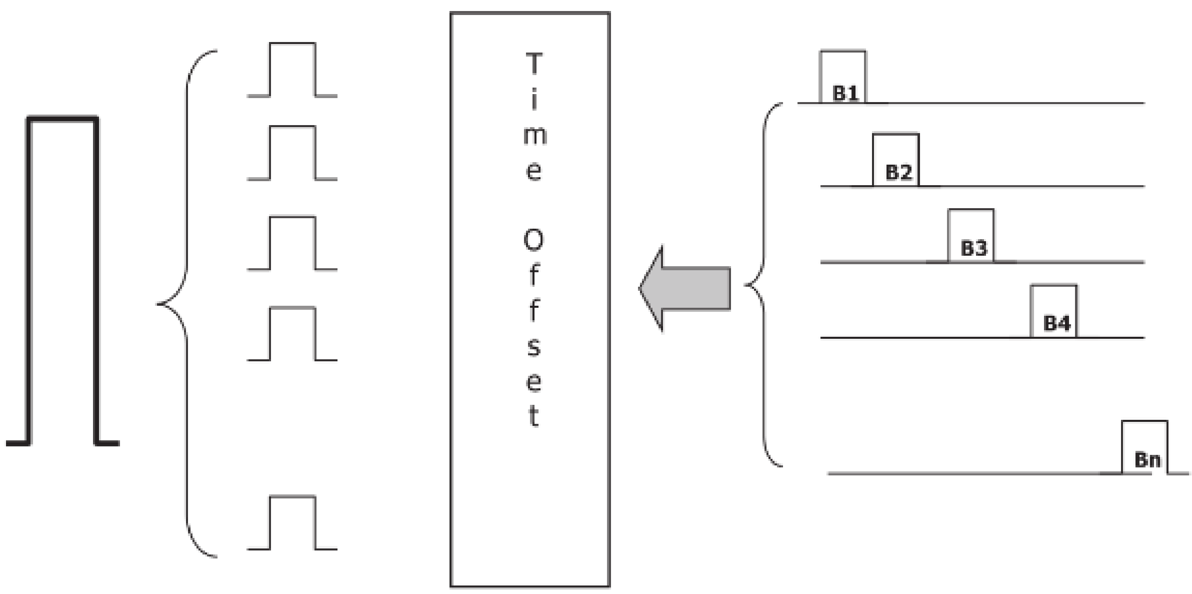


The pulse, after having hit an obstacle, is reflected, but B1 returns before B2 which returns before B3, etc.
It is therefore necessary to have this succession of echoes go through a system synchronizing all these pulses back on the same origin time.

The echo will cross a filter bank tasked with classifying each burst, and then each signal exiting its own filter will be subjected to a time offset that will bring it back to the same ~~origin~~ time before treatment, giving as a result the equivalent of a pulse, of time equal to B1, but of strong amplitude, sum of B1+ B2+ B3 + etc…



Separation of the bursts by filter banks



Realignement of the bursts on the same time of origin and creation of the equivalent pulse

06

GLOBAL NAVIGATION SATELLITE SYSTEMS (GNSS)

01 GENERAL

1.1 – Principles

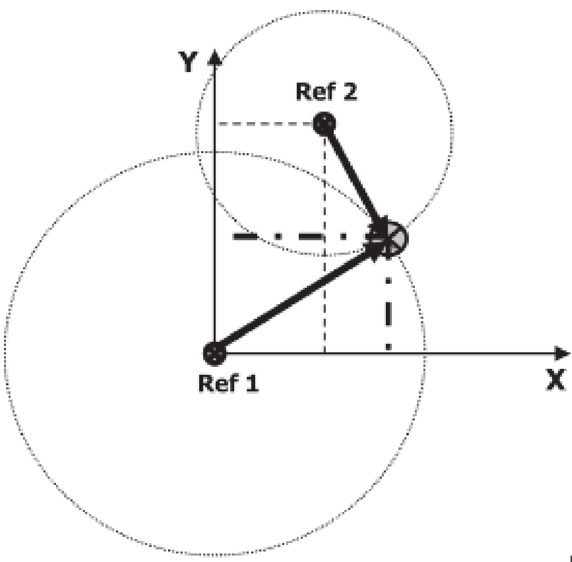
At the moment, there are four satellite assisted navigation systems:

- NAVSTAR GPS, American system;
- GLONASS, Russian system;
- GALILEO, European system under construction;
- BEIDOU, Chinese system under construction.

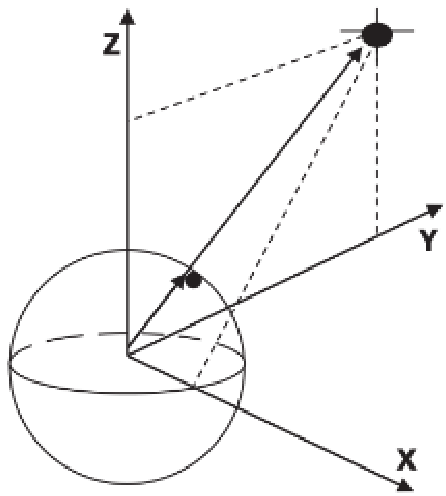
Those systems, even though they are based on different satellite constellations, should, with time, be inter-operable with the use of a proper receiver.

The systems are based on constellations of satellites broadcasting signals that can be used by suitable receivers in order to provide a position.

There are many possibilities to determine the position of a point in a coordinate system.
One of them consists of starting from reference points (at least 2) and having the knowledge of at which distance the unknown point is located from each of these reference points.



DME/DME positioning



Satellite positioning

This method requires agreeing on the referential, and on the other hand being able to measure D with sufficient accuracy.
We had a quick look at this notion in the DME chapter (position determination).
The measure of D was based on a measurement of time of radio wave propagation.
The GNSS applies a similar principle, namely a measurement of time (from which it deducts a distance) and the positioning of a user in a three-dimensional reference system, starting from this measurement of time which must be very accurate.
The reference points are satellites in orbit around the earth.
The reference datum is different for each system, but we can go easily from one to another one.

The one used by the GPS NAVSTAR is the WGS84 (World Geodetic System 1984).

1.1.1 - WGS 84

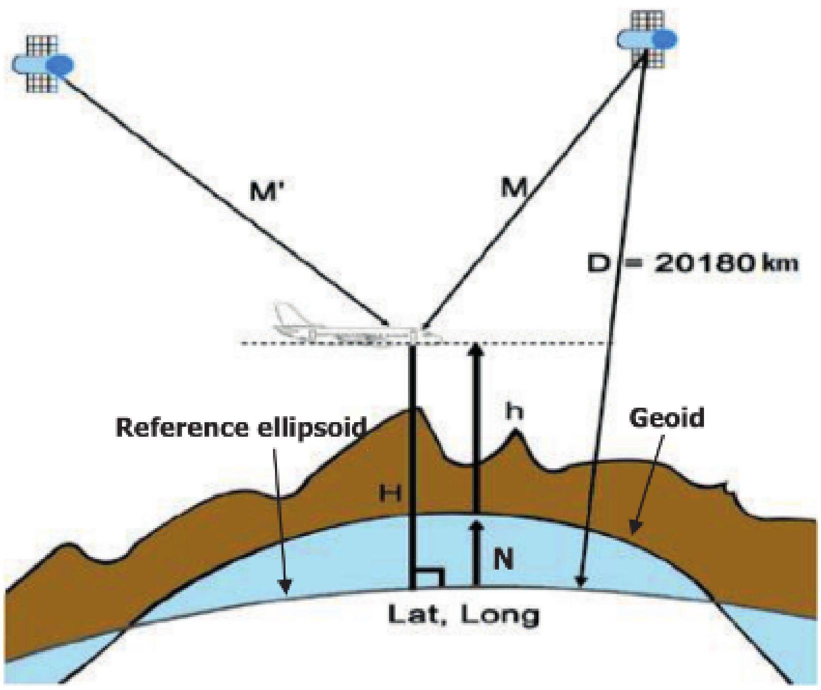
The WGS84 defines in three dimensions the coordinates of a point on the earth's surface by applying corrections to two conventional surfaces:

The geometric ellipsoid, the so-called reference;
The geoid, an equipotential surface meant to coincide with the mean sea level.

The ellipsoid is an approximate shape of the earth defined by convention.
The geoid is a particular equipotential surface, since it coincides with the mean sea level, artificially extended below the land.
It refers to the hypothetic extension of the mean sea level below the land and constitutes the reference surface used for aeronautical charting.

The geoid is deducted from the reference ellipsoid by the correction, calculated at all points, of N meters above it or below it.

N is called the undulation of the geoid. It is the GUND found on visual approach charts.



The value of N is obtained from a database.
The entry into the table is done with the Latitude/Longitude calculated by the GPS receiver.
Hence, a GPS position error will provide a false correction N, and therefore a false altitude.

1.1.2 – Physical structure

Starting from the stated positioning principle, all GNSS systems need: reference stations (the satellites), a communication protocol, a surveillance and correction system (remember that the measure of distance must be precise) and a measurement system (the receiver), all used in a common referential (Ex: WGS 84).
The system is made up of 3 parts:

- The Space Segment;
- The Control Segment;
- The User Segment.

1.2 - Operation

1.2.1 - GPS NAVSTAR

The Global Positioning System (GPS) is a worldwide system (*Global meaning around the globe*).
The GPS is the result of a project launched at the beginning of the 1960's in the USA and called NAVSTAR.
NAVSTAR stands for NAVigation Satellite Timing And Ranging.
There are two operating modes:
the Precise Positioning System (PPS): restricted to the US military or other approved users, and providing a higher accuracy than the SPS;
the Standard Positioning System (SPS): available to civilian users and providing a positioning accuracy lower that of the PPS by introducing deliberate errors in some of the parameters transmitted by the satellites.
Note:
This is not true anymore, the voluntary degradation of the signal is now suppressed, but nevertheless a return to it is always possible, and this is why we will pretend that it still exists.
The system remains entirely under the control of the USA because the satellites and the control stations belong to them.

1.2.2 – Space segment

This is the group of satellites in use.
The satellites are spread over a certain number of orbits allowing a global coverage.
The altitude of the satellites varies from one system to the other (around 20 00 km) and therefore the revolution period depends on this altitude.
Because they are low, the orbits of these satellites are practically circular.
Like all orbital satellites, they obey ~~to~~ Kepler's laws.

1.2.3- Frequencies and codes

The satellites are equipped with a transmitter-receiver and with atomic clocks (Caesium, Rubidium or Hydrogen Maser) to maintain a very accurate time. (The drift for a Hydrogen Maser clock is 1 second over 3 million years).
Satellites transmit, on two carrier UHF waves, different coded PRN signals (Pseudo Random Noise) relative to the time, and the parameters necessary for positioning and navigation.

The first signal is a code named C/A (Coarse Acquisition, sometimes called Clear Access) which forms the basis of the Standard Positioning Service (SPS) used by civilians.
The second code is called P code (Precise or Protected).
The Precise Positioning Service (PPS) is more accurate than the SPS.
These two codes are PRN encoded (see the notes outside of this course at the end of this chapter).
Each satellite uses a fixed PRN sequence, different from the ones of other satellites.
Each satellite can be identified by its PRN.
The carrier frequencies in use are in the L band:

- L1: *1 575,42 MHz (sometimes called Link 1);*
- L2: *1 227,60 MHz (called Link 2).*

L1 is modulated by P and C/A codes.
L2 is modulated by P code only.

Only the C/A code is accessible to the civilian community.
The SPS uses the L1 frequency and the PPS uses both the L1 and L2 frequencies.

Both of these frequencies also transmit a navigation and system message, detailed later, superimposed to the C/A and P codes.

NOTE:
Since 25 September 2005, the first ~~satellite~~ of a new generation of satellites has been launched.
It is a group of satellites called Block IIR-M.
These satellites (and the next ones of the block) transmit a new M code reserved for the military on bands L1 and L2, a code that is more resistant to interference.
A new signal intended for civilians is transmitted on the L2 band which becomes the L2C signal.
This signal is transmitted at a higher power and can be more easily extracted from the radioelectric noise.
It must provide ~~a~~ greater accuracy, ~~a~~ better integrity and ~~a~~ better coverage, in particular in shaded areas (undergrowth, parking, mountainous area).
Having two signals on L1 and L2C gives better correction of the ionospheric delay, directly by the receivers (precision factor).

The system managers can at all time reactivate the voluntary precision degradation (abandoned by President Clinton) provided by the C/A code *(but also the P code)* by means of a device called Selective Availability (SA).
This consists of manipulating the digits that provide the satellite time, or the precise positioning of the satellite on the orbit (ephemeris), with a secret key.
This function introduces pseudo random errors which blend perfectly with all types of errors from the GPS and which cannot be corrected without the key.

It is possible to counter this difficulty with the implementation of the differential GPS seen at the end of this chapter.

1.2.4 – Navigation message

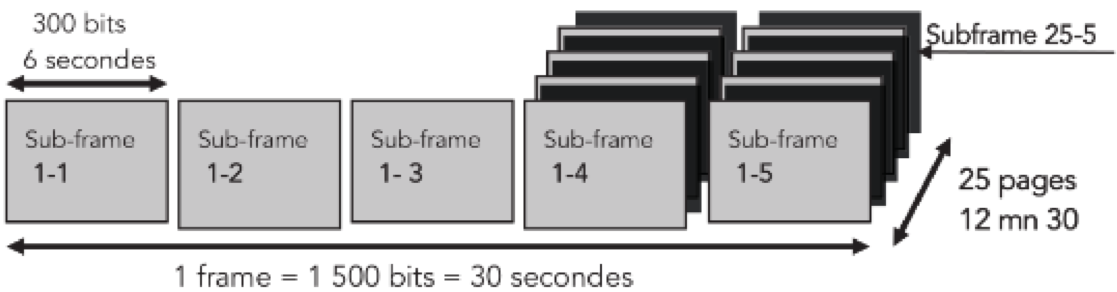
A navigation message superimposed to the C/A (*and P*) code, sent by each satellite contains information such as:

- The time of transmission of the frame;
- The almanac of the constellation (approximate position of the group of satellites);
- Its own ephemeris (precise position of the satellite on its orbit);
- Ionospheric model coefficients;
- The satellite performance level (health condition);
- The difference between the GPS time and the UTC UNSO time with an accuracy of 0,1 µS;
- The good working condition of the satellite.

Note: *UTC UNSO: UTC time calculated by the US NAVAL OBSERVATORY.*
GPS time: UTC time on the date of 06 January 1980 at 00:00.
GPS time is related to UTC time but is not synchronized by it, hence a difference transmitted in the navigation message (13 seconds difference on 07/01/2000).

NAVIGATION MESSAGE FORMAT

Consists of 25 frames (pages) distributed in identical subframes, with the exception of subframes 4 and 5.



1 frame = 5 subframes

CONTENT OF THE MESSAGE

Subframe 1	**TLM**	**HOW**	Satellite clock correction
Subframe 2	**TLM**	**HOW**	Ephemeris 1
Subframe 3	**TLM**	**HOW**	Ephemeris 2
Subframe 4	**TLM**	**HOW**	UTC, ionospheric model, …
Subframe 5	**TLM**	**HOW**	Almanac satellite 1 to 24

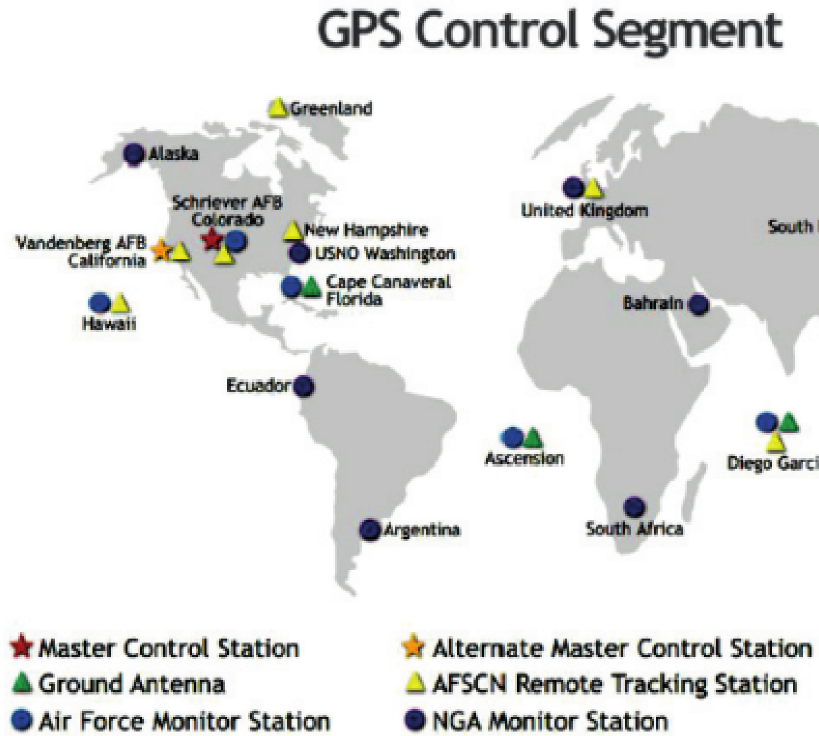
The recovery of the complete almanac requires the reception of subframes 5-1 to 5-25, hence a recovery in 12 minutes 30.

As we said before, the movement of the satellites on their orbit obeys Kepler's laws.
We can, from those laws, in a more or less near future, determine the theoretical position of the satellites.

The collection of these theoretical orbital data of all satellites form the almanac.
This theoretical path is influenced by numerous other parameters such as the gravitational influence of the sun, of the moon or of other planets, and by solar radiation.
The real position of the satellites detected by the ground station is thus not precisely the one determined by Kepler's laws.
This very accurate data is called ephemeris.
The almanac data is therefore less accurate than the ephemeris data.
They are used primarily to calculate if, for a given user, such or such a satellite is visible.
The master station transmits, the clock corrections and the computation of the ephemeris at a minimum every 12 hours, to the satellites. The satellites store and then pass on the information to the users.
(The satellites that will soon be launched compute their ephemeris internally).
The orbital data are broadcasted by each satellite in the following way:
ALL the satellites transmit the ALMANAC (theoretical orbital data of all satellites).
This information is not very accurate but is valid for many months (180 days).
EACH satellite transmits its EPHEMERIS (actual orbital parameters) every 30 seconds (one time per frame).
This very accurate information is valid from for 4 to 6 hours.
We will come back on those points in the chapter "User segment"
Time differences are also detected.
The detected time difference is stored and then transmitted to all users in the navigation message.
The health status, provided by the validity of the navigation data, is used to exclude unhealthy satellites from the position calculation.

1.2.5 – Control segment

This refers to all the control bases that follow at each second the path of each satellite.
It is composed of a master control station and of monitoring stations located around the world in order to obtain a permanent tracking.
The COLORADO SPRING station located on the Falcon Air Force Base in Colorado is the master control station.
The stations dialogue with each other, either via a ground network or via satellites.
The role of these stations is to ensure satellite monitoring, detect positioning and time errors, but also to send satellites the error corrections, via the master station, which also computes the ionospheric propagation time.



1.2.6 – User segment

This is the best known segment because it is made up of all the GPS receivers used around the world.

There is a very significant variety of receivers because there is a GPS for every type of use (mountain hiking, sea navigation, etc…).

A GPS receiver provides:

- a position in a three-dimensional space (by computation);
- a speed (from measures of the Doppler effect on the carrier (and) or from movement as a function of time.
- A time (it is the very principle of the GPS).

- The GPS receiver thus makes it possible:
- to know its position;
- to plan a route;
- to follow a route;
- to memorize points on a chart (database).

The GPS receiver can be used on its own or as part of a navigation system.

a) Receivers

The receiver is able to determine the distance satellite-aircraft by measuring the elapsed time from the transmission of a satellite signal to its reception onboard the aircraft.

> *Analogy to understand better:*
>
> *A person A and a person B are separated by a distance D.*
> *A shouts to B: « it is 10 hours 25 minutes and 30 seconds on my TOP ».*
> *When B hears the TOP he looks at his watch and reads 10 hours 25 minutes 32 seconds.*
> *Knowing the time that A was transmitted, because it is included in the transmitted message, B deducts that the information took 2 seconds to reach him and that, with a theoretical speed of sound of 330 m/s, he is located at a distance D = 660 meters, from A.*

As we will see in more details, the position of a user is determined by the intersection of 4 position lines, each centered on a satellite.
We understand easily that the simultaneous acquisition of these 4 position lines is adapted to users moving at high speed.

GPS receivers used in aviation are therefore multichannel.

Some of these receivers are conceived to receive all the satellites within sight of the user and are named "ALL IN VIEW" receivers.
They are able to pursue up to 12 satellites simultaneously.
They elaborate a position using all satellites in sight.
Therefore, if one of the satellites is momentarily masked, the calculation of the position continues with practically no accuracy degradation.

A satellite in considered in sight if it is 7,5° above the horizon.

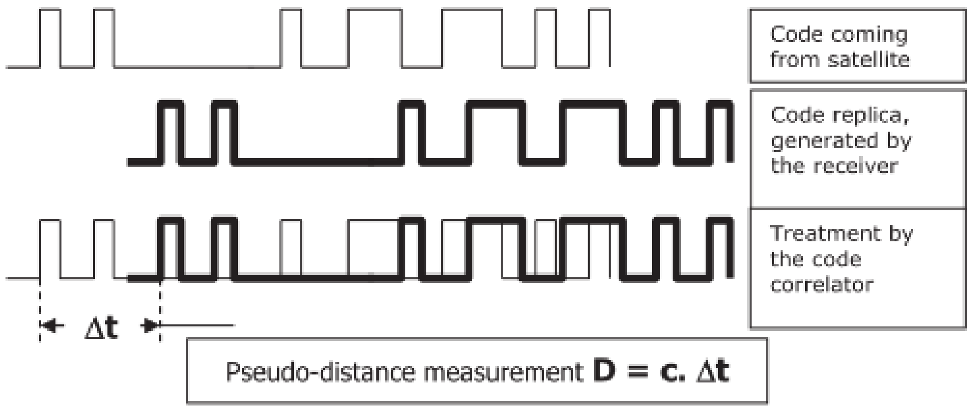
To achieve this, GPS receivers are equipped with a mask angle setup set at 7,5°.
The objective is to calculate ignoring the satellites that are too low on the horizon where the signals would inevitably be of poor quality (noisy and unstable signal).

The antennas for the GPS receivers are located on top of the aircraft.
The receiver must be located close to the antennas to minimize the line losses on a signal that is already very weak.
About the antennas, studies are underway to obtain an aircraft attitude reference, like with an IRS but without the drift over time and at a lower cost, with a GPS system and many antennas for which spacing is known.

b) Calculation of the distance satellite - aircraft

The receiver extracts from the UHF carrier the binary codes transmitted by the satellite.
The receiver generates a replica of the C/A (or P) code from the satellite and measures, by correlation, the shift between this replica and the code transmitted by the satellite.

This shift is a measure of the propagation delay of the signal from the satellite to the receiver. Multiplying this propagation delay by the propagation velocity will provide the distance to the satellite.

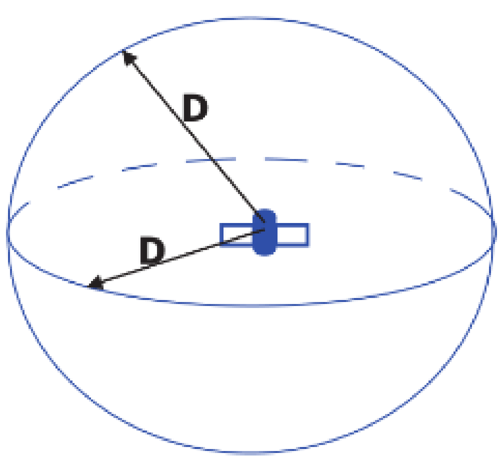


C: Wave propagation velocity

The receivers that work on that principle are called code correlator receivers.

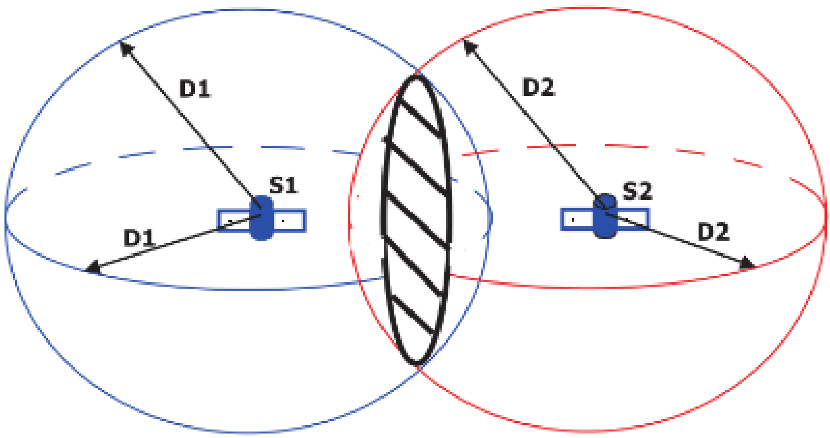
Other receivers called phase shift receivers (Carrier Phase GPS) work by using the carrier phase and provide greater accuracy (used in geodetic).

The calculation of Δt does not take into account the fact that the satellite clock and receiver clock are not synchronized.
For technical and cost reasons, the receiver clock is a simple quartz clock and its drift is higher than that of the satellite, the calculated Δt is therefore not rigorously correct.
Also, the wave propagation is not done in a homogeneous medium (ionosphere, troposphere).
Propagation velocity varies while we consider it to be constant in our calculations.
This is referred to as the ionospheric delay.
The calculated distance is therefore not the physical distance satellite-receiver.
This is the reason why this measure is called « pseudo distance » in the GPS dialect.
If this calculation allows us to say that the receiver is at a distance D of the satellite, it does not allow the determination of a position, but only of an area of position.
The area of position is a sphere of radius D centered on the satellite.
The receiver can be anywhere on that sphere.

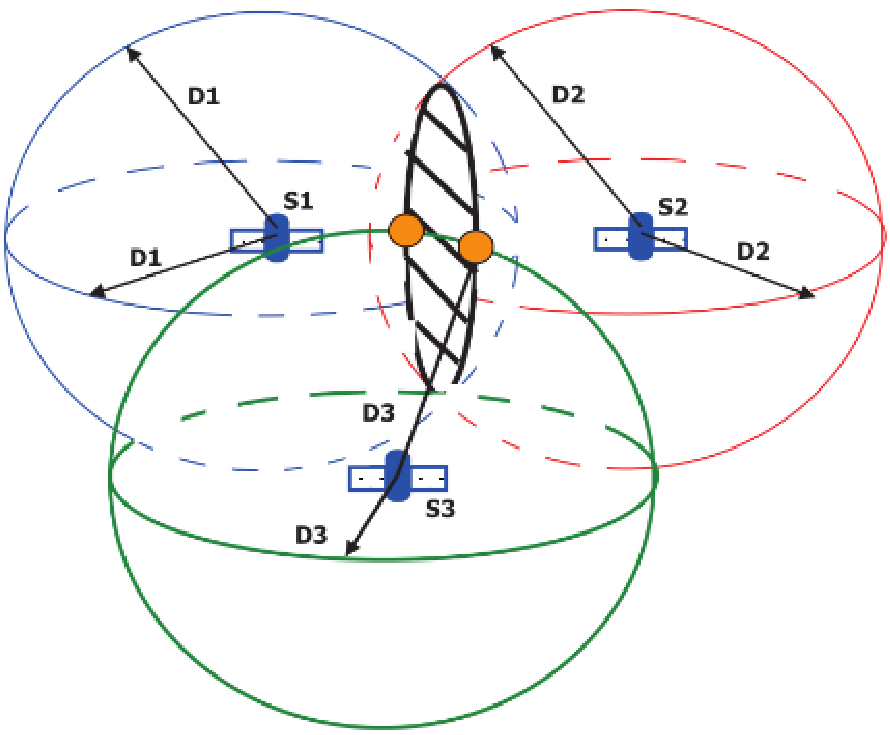


c) Position determination

If the receiver acquires two pseudo distances, from satellites S1 and S2, the area of position is then the intersection of the two spheres.
This area of position is a circle.



By repeating the operation with a third satellite, only two points of the previous circle correspond to the position of the receiver.



Since in our application, the receiver is either on earth or in an aircraft evolving in close proximity to the earth, and that the two possible points are located one near the earth and the other one very far in space, the irrelevant point can easily be excluded from the calculation.
However, we have not resolved the problem of time delay, and the found position, based on three pseudo distances, is not rigorously correct.

A time delay of 1 microsecond brings an error of 300 meters.

It is necessary to find a way to eliminate this clock error to be synchronized on the reference GPS time.

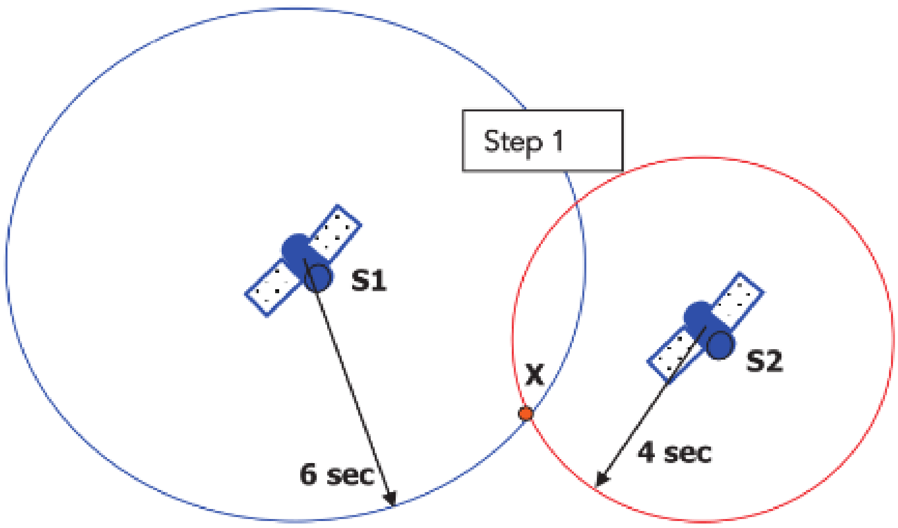
The solution rests on the use of a fourth satellite.

Let's try to understand how 4 inaccurate satellite measures can provide an exact positioning, or at least one emancipated from the clock error.

The previous figures show the difficulty to reason in 3D in the 2D space, this is why we will analyze the problem in this 2D space with 3 satellites, which will be equivalent to a 3D demonstration, with 4 satellites.

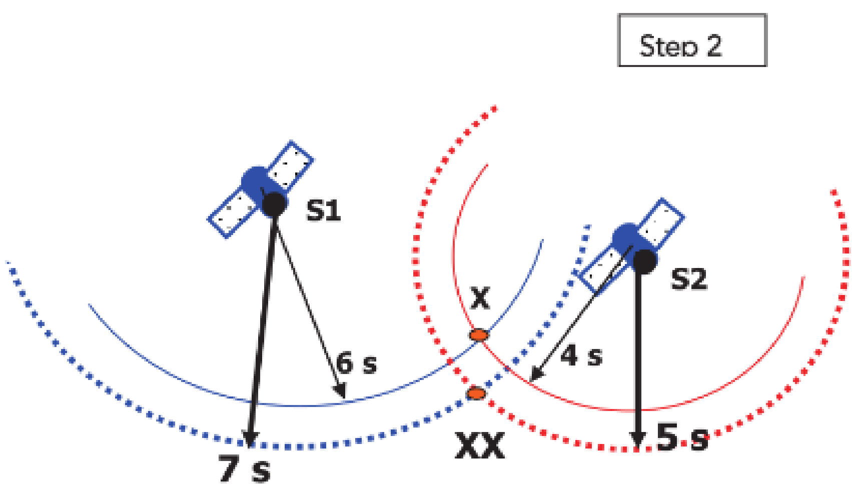
Two exact measures of distance would be sufficient to determine a position.

Point X is at the intersection of the two measures deemed to be exact (6 seconds and 4 seconds).



Let's suppose that each measure is affected by a 1 second error due to the receiver.

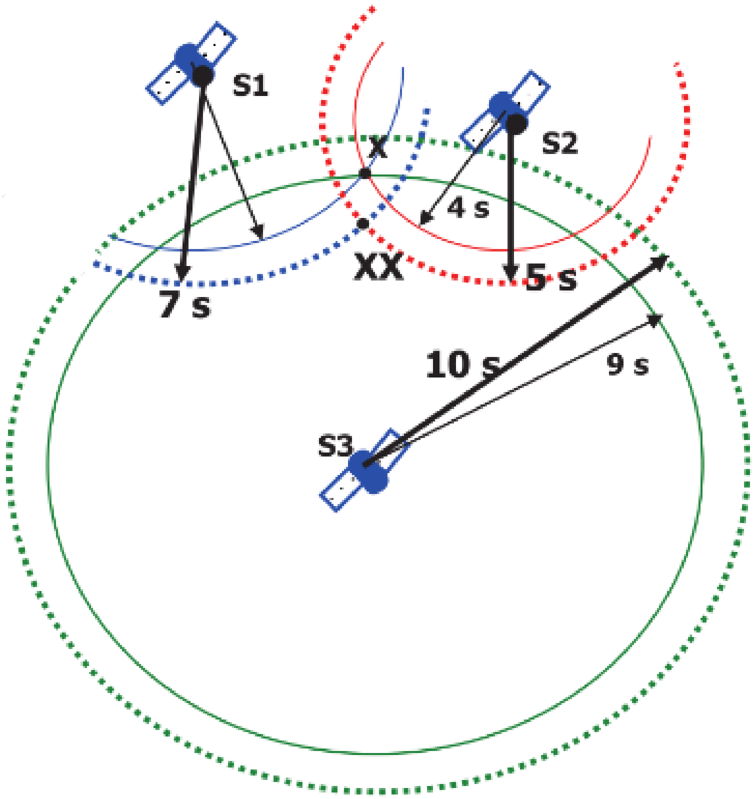
We will then measure 7 seconds to S1 and 5 to S2.



After the measurement of the two pseudo-distances, the position indicated by the receiver is XX.

A third satellite is thus necessary.

Step 3



By considering that the error made by the receiver in measuring the propagation delay from S1, S2 and S3 is the same, since it is imputable to the receiver clock, we just need to find which value to subtract from each measurement (which corresponds to the error) so that the dotted circular arcs intersect in a common point which will result in position X.

By reasoning in 3D space, we understand **that we must resort to a 4th satellite.**

We can then calculate the position of the receiver by resolving a system of 4 equations and 4 unknown parameters with this model:

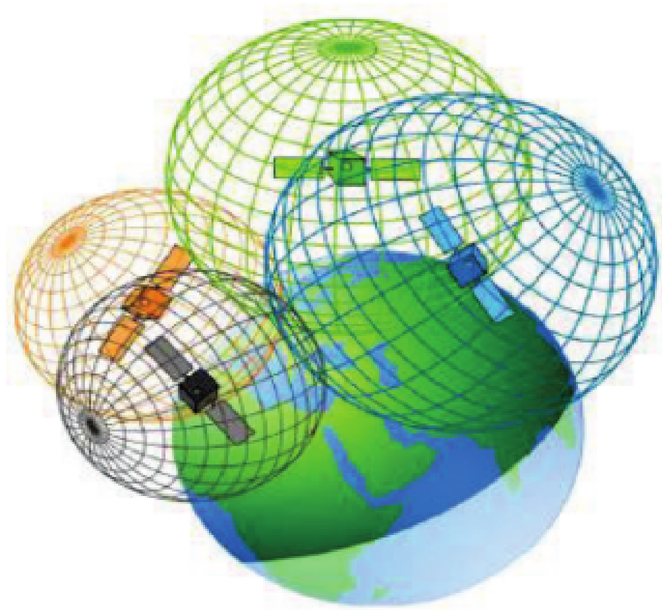
$$(X-\alpha)^2 + (Y-\beta)^2 + (Z-\delta)^2 = (R-CB)^2$$

X, Y, Z are the coordinates of the position of a satellite;

α, β, δ are the coordinates (unknown) of the receiver;

R represents the pseudo-distance receiver-satellite;

CB represents the delay imputable to the receiver clock in relation to the GPS time (CB = Clock Bias).



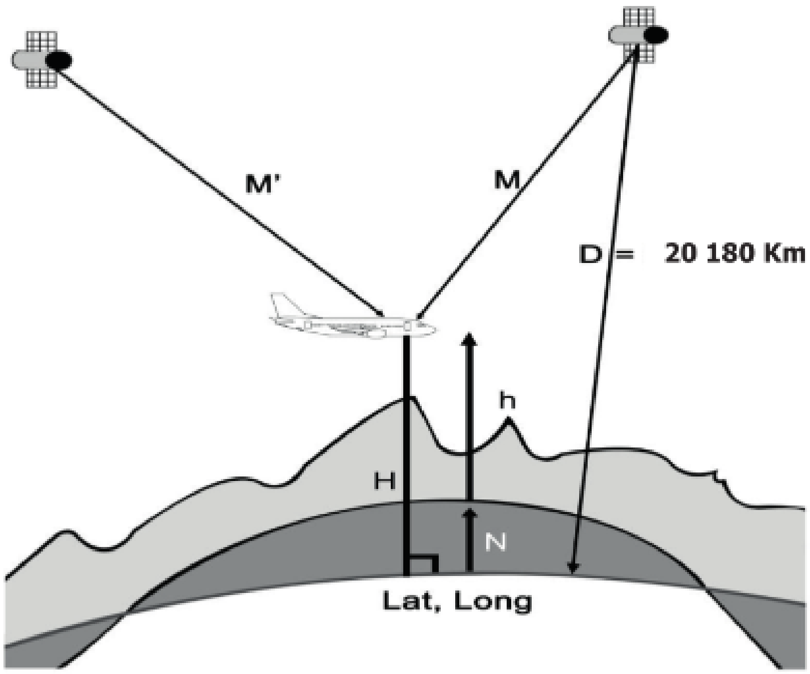
Therefore, 4 satellites are needed to determine a position in 3D, but if the altitude is known (in the case of maritime navigation for example) there is one less unknown parameter and the system of 4 equations and 4 unknown parameters becomes a system of 3 equations and 3 unknown parameters.

Hence this recurring question:
How many satellites are needed to elaborate a position in 2D?
Answer: 3 YES BUT… Provided that we know the altitude, which the question never specifies.
This all seems perfect and should allow us to obtain an accurate positioning. However, another detail needs to be settled: the very accurate position of the satellites must be known.
To accurately calculate the intersection of the position areas, we must know exactly the position of the satellites in space (see parameters X, Y, Z of the previous equation).
For that, the control station performs the ephemeris calculations every 12 hours, transmits the results to the satellites and via the navigation message allow the users to retrieve the data.

d) Ionospheric delay

Furthermore, with have seen that the wave does not propagate in a homogeneous medium and an error due to crossing of the ionosphere must be corrected.
The method will be treated in Chapter E – Errors and factors affecting accuracy.

e) Vertical GPS indication



The GPS receiver determines a 3D position based on pseudo-distances (M, M', etc.).
This position will allow determination of the orthometric height H in relation to the reference ellipsoid.
Knowing the position in Latitude, Longitude, the receiver will read in the database the factor N to be added algebraically to H to determine the height h above the geoid.
The GPS receiver calculates a height above the reference ellipsoid but provides the crew with an indication of height in relation to the mean sea level (geoid).
It is always possible, when necessary, to go to another model than the WGS84 by applying mathematical transformations.

1.2.7 - GPS NAVSTAR integrity

The integrity measure consists of giving the receiver the means of verifying the integrity of the received navigation data by analyzing the coherence of the pseudo distance measurements.
There is a software program (RAIM for Receiver Autonomous Integrity Monitoring), integrated to the receiver, and it is tasked with monitoring the integrity of the GPS and eventually to predict a loss of integrity (PRAIM).
It requires the use of more than 4 satellites.
RAIM:
There are two levels, each requiring a different number of received satellites.

FDI (Fault Detection Identification):
The user is informed about the true accuracy of the calculation.
REQUIRES AT LEAST 5 SATELLITES IN VIEW.

FDE (Fault Detection Exclusion):
The system eliminates the faulty satellite from the calculation.
REQUIRES AT LEAST 6 SATELLITES IN VIEW.

The decree of 19 June 1998 art.5-3 paragraph a) and b) specifies:

a) In case of loss of the RAIM detection function, the flight crew may continue navigating with the GPS equipment. The flight crew should attempt to cross-check the aircraft position with the information provided by the VOR, DME and ADF, in order to confirm the existence of a required level of precision. In other case, the crew must revert to an alternative navigation means.
b) In case of exceedance of the alarm limit, the flight crew must revert to an alternative means of navigation.

The decree of 19 June 1998 art.5-2 paragraph a) specifies:

a) During the flight planification phase (pre-flight), given a GPS constellation of 23 satellites or less (22 satellites or less for an autonomous GPS equipment using pressure-altitude information), the availability of GPS integrity (RAIM) must be confirmed for the planned flight (route and time). This must be obtained from a prediction program **either ground based, or integrated to the equipment**, or by another acceptable means for the minister in charge of civil aviation.

The flight dispatch must not be authorized in case of predicted continuous loss of RAIM of more than five minutes on any segment of the planned flight.

If the GPS is assisted by receiving the barometric altitude and ~~that~~ this altitude is taken into account by the RAIM, the same functions will be available with one less satellite.
FDI: 4
FDE: 5

The RAIM guarantees a certain integrity but does not provide any improvement of the navigation accuracy unlike GBAS and SBAS treated later in 062 06 02 00.

a) Acquisition mode

When the receiver is started, if it has not worked for a certain time (or if it has been moved 100 NM in a short time), it does not know the position of the satellites.
A function called by some manufacturers SEARCH THE SKY or AUTOLOCATE (Garmin source) allows the receiver to listen to the visible satellites.
A satellite is considered visible if it is at least 7,5° above the horizon.
When the receiver is locked on a satellite, it will acquire the almanac.
Reminder: Il will take 12 min 30 sec to retrieve the almanac.
After retrieving the almanac, the receiver memorizes the satellite constellation, it then scans the position where the satellites should be VISIBLE.
It can then choose the satellites presenting the best configuration (azimuth and elevation) for the position calculations.
It then acquires the ephemeris and other parameters of the navigation message of each satellite and goes into calculation of position (navigation mode).
In such a situation, with absence of a position and absence of almanac or time inaccuracy, it will take around 15 minutes for the receiver to go into navigation mode (TTFF: Time To First Fix).

At the end of the mission, the GPS keeps the constellation in its memory.
For its next use, it will directly scan the useful satellites.
It will go to navigation mode in 30 seconds.

b) Interoperability

Even though they use different reference systems than the WGS84, different satellite constellations (satellite altitude, inclination, revolution period) and they can provide different services, the positioning principle of the other GNSS constellations is similar to the one of NAVSTAR.
Agreements were made between appropriate agencies to ensure interoperability of NAVSTAR and GLONASS.
Interoperability is also planned between GLONASS and GALILEO, the future European constellation.
See notes outside of this course.

1.3 - Errors and factors affecting accuracy

Errors affecting the measure of the pseudo distances and therefore the accuracy are mostly due to:

- Ionospheric delay;
- Dilution of position;
- Satellite clock errors;
- Satellite orbital variations;
- Multipath effects.

The impact of these errors is dimensioned by the UERE (User Equivalent Range Errors).
UERE is defined as being the square root of the sums of the square of each of these errors.

1.3.1 – Orbital variations

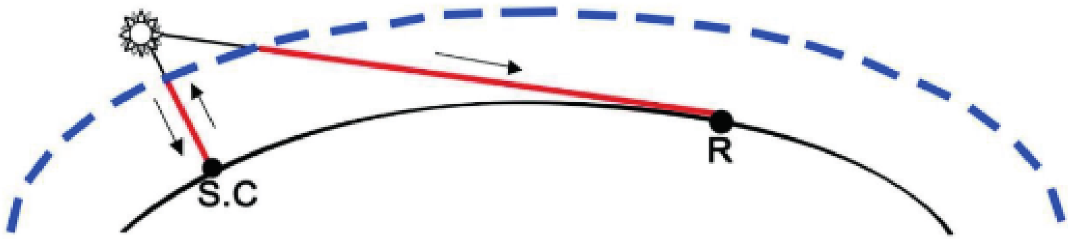
Orbital variations of the satellites are caused by solar winds and by the sun and moon gravitational forces.

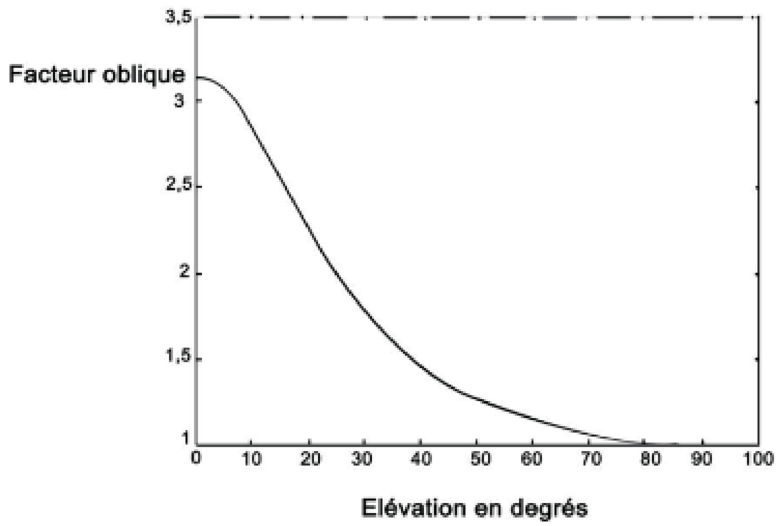
1.3.2 - Ionospheric propagation delay (IPD)

The ionospheric delay depends on the ionosphere TEC (Total Electrons Count) and it varies according to the amount of solar radiation, the geomagnetic latitude and altitude of the satellite.
The ionospheric delay can vary from a few nanoseconds at night vertically, to many hundreds of nanoseconds during the day for a satellite low on the horizon.

In order to calculate the delay, control stations that receive L1 and L2 frequencies perform measurements on those two frequencies.
It is demonstrated that the ionospheric delay is inversely proportional to the square of the frequencies.
The delay that they detect is valid only for their position.
They cannot transmit this delay to users who are subjected to a different delay.
Indeed, the path of a wave in the ionosphere is greater for a satellite on the horizon than for one vertical of the user.
The control stations extract coefficients (*8 parameters from the Klobuchar model)* from an equation representing the vertical ionospheric delay, which is valid for all (the ionospheric model).
They then charge the parameters of the ionospheric model on the satellites which pass on the data via the navigation message and, the user (the receiver) applies the ionospheric model affected

with the correct parameters and an oblique factor (depending on the satellite altitude) in its calculation of the ionospheric delay.





This method has limitations because, ionospheric variations can be rapid and the navigation message is not updated rapidly with regards to parameters, and also the ionospheric model in this method does not reflect exactly the very complex reality of the state of the ionosphere in relation to its location.

This method corrects only around 50% of the ionospheric error.

The ionospheric delay is the most significant residual error.

The transmission of a new code on L2C will enable to locally make the calculation done by the control stations (since the user will receive L1 and L2C) and therefore improve accuracy.

1.3.3 – Multipath error

The wave coming from the satellite to the receiver can be direct or reflected on conductive obstacles.

In general, the pseudo distance calculated with the wave path shows, during a relatively short period of time, a difference with pseudo distances acquired on other satellites.

The error will vary rapidly over time and a Kalman filter will take into account only a small portion of the error without real incidence on the average value if the multipath lasts only for a brief moment.

If the problem persists, the RAIM can, in case of an anomaly, deselect the satellite that is the source of the error and replace it by another satellite instantly without degradation.

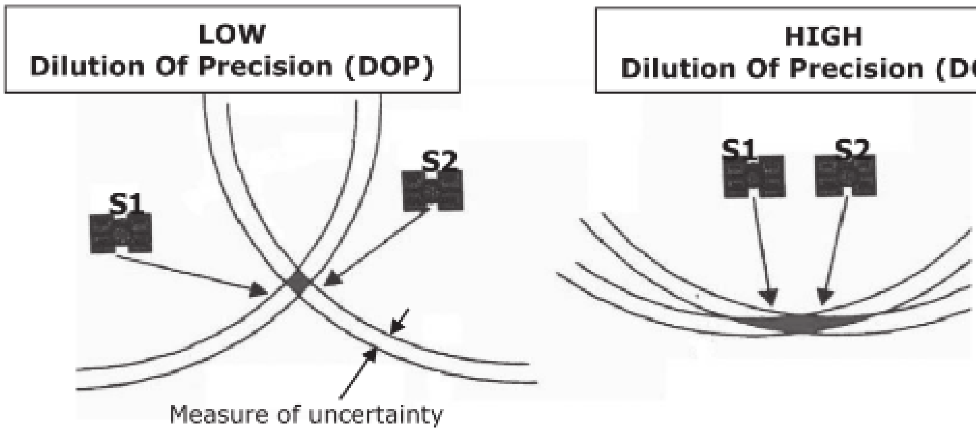
To minimize multipath problems, the location of the antennas will have to be chosen carefully.

We can also use special antennas called « choke ring » antennas, but they are mostly suited for ground stations (for a DGPS reference station for example).

Choke ring antenna

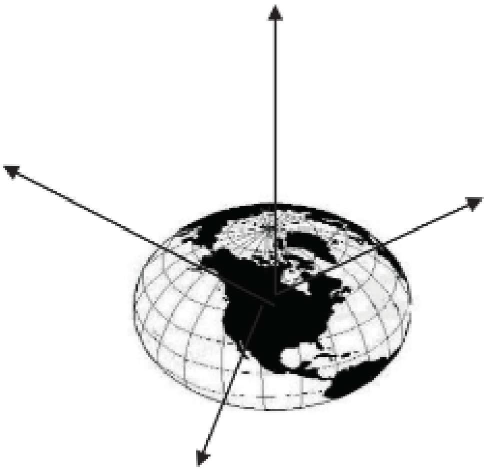
1.3.4 – Dilution of position (DOP)

The precision of the measurements also depends on the overlap of the areas of position, and thus on the geometric configuration of the satellites in relation to the user.
The factor to take into account is the GDOP (Geometric Dilution Of Precision), a variable with no dimension calculated from the angle formed by the point and the satellites.
If the DOP increases, the accuracy decreases.



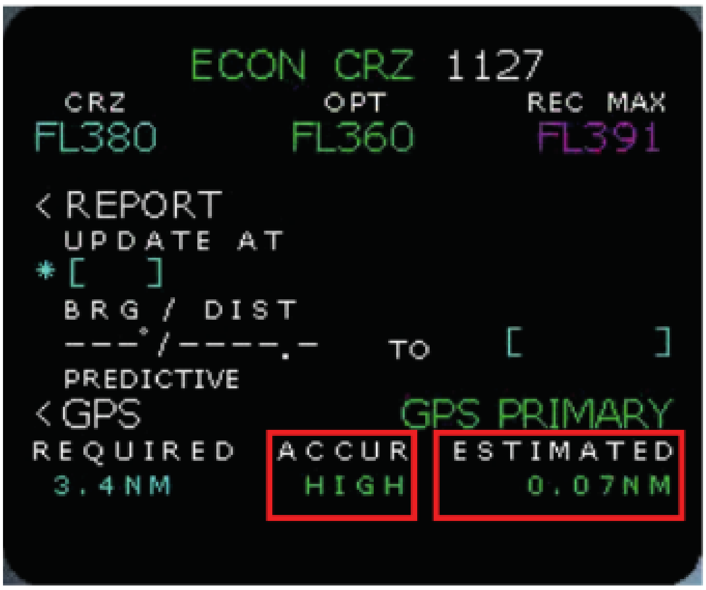
The DOP will be greater when the satellites are (relatively) close to each other.

The ideal situation is to have one satellite vertical of the observer and the other three low on the horizon (but with a minimum of 7,5° of elevation) spread out every 120 degrees.



Other than this ideal case, the reception of a high number of satellites (7 or 8) favors a low DOP because there is more choice using the 4 satellites presenting the best configuration.

The product of the UERE by the DOP provides an estimation of the position accuracy (EPE - Estimated Positional Error).



02 AUGMENTATION SYSTEMS

These systems, built around satellite navigation systems that were previously studied, are intended to increase integrity or integrity and accuracy.

Integrity signifies:
The trust in the validity of the information;
The capacity of the system to warn of a failure.
ICAO specifications require a maximum delay of 2 seconds to warn the user of a failure of a precision system (for example the ILS) and of 8 seconds for a non-precision system.
A receiver that produces a 3D position, from the measure of 4 pseudo-distances, has no means of detecting that the data received from a satellite is corrupted and can therefore make a significant positioning error (many hundreds of NM).
It is the same problem as with the IRS, if there are only 2 IRS and they do not agree, it is not possible to figure out which one is right.

Additional techniques are needed to guarantee integrity.
Some already exist and others are under development.
For civil aviation there is:

1) Air Based Augmentation System (ABAS):
Based on multichannel receivers and an appropriate computation program, RAIM or more evolved (AAIM).

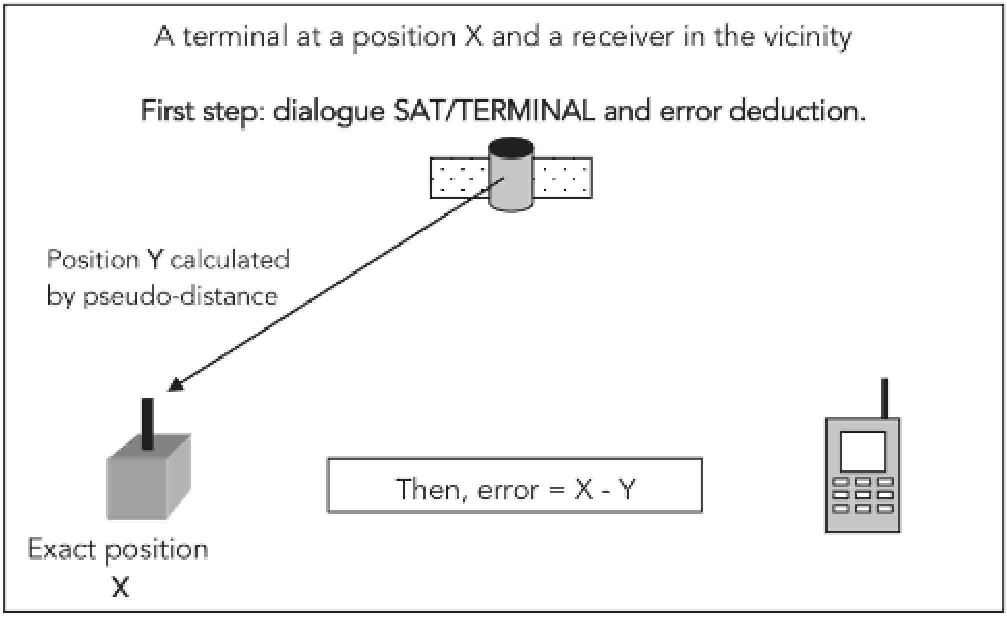
2) Ground Based Augmentation System (GBAS):
Local area DGPS (Differential GPS), based on ground implementation of pseudo-satellites.

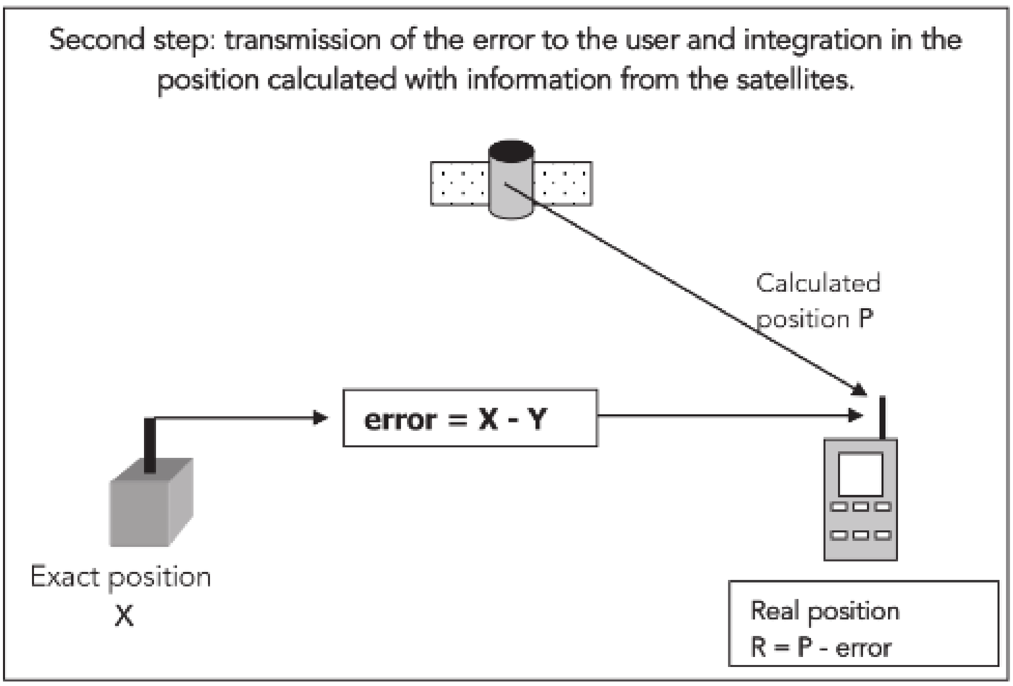
3) Space Based Augmentation System (SBAS):

Using geostationary satellites.

2.1 - GBAS and DGPS (Differential GPS) principle

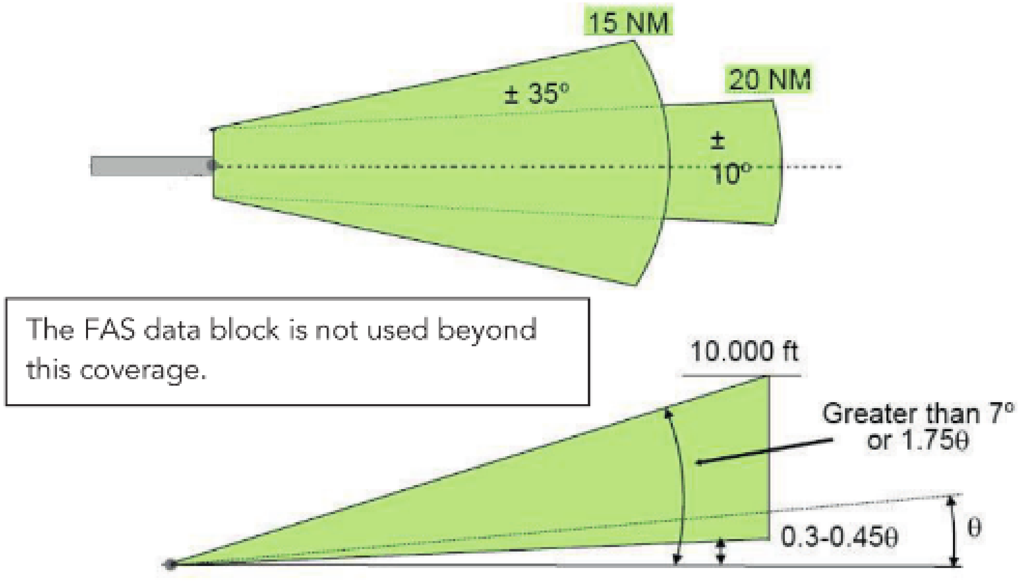
GBAS uses terminals (called pseudo satellites or pseudolites), which know exactly–their position and are installed in the vicinity of airports.
A GPS receiver is fitted with a device that allows it to receive the information coming from those terminals, in addition to the GPS signals.





The terminal knows exactly its geodesic position and sends the error correction via data link to users through a special channel (VHF VOR or ILS channel).

The minimum coverage area IN AZIMUTH is 35° on either side of the runway centerline up to 15 NM, and 10° from 15 to 20 NM.

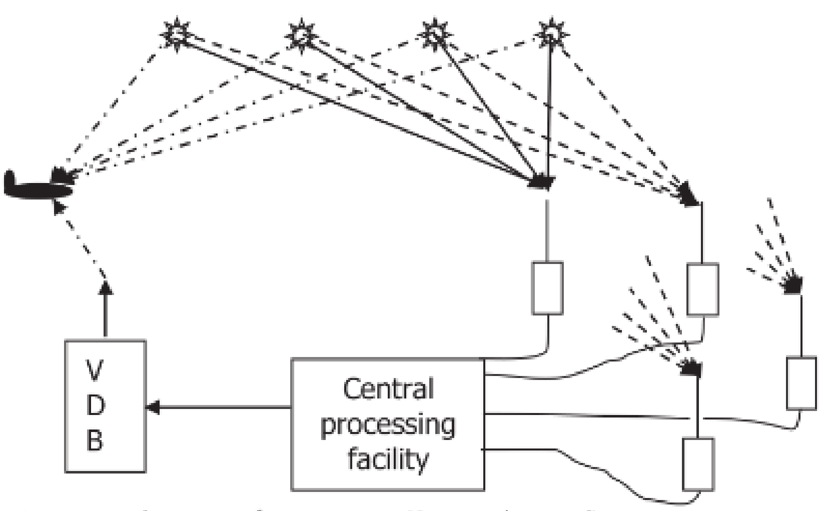


Minimum Service Volume

From the GBAS principle we can deduce that the correction of error that it brings is better when the user is close to the ground terminal.

This type of augmentation is also designated by the term LAAS (*Local Area Augmentation System*).
The station elaborates the distance differential (error correction), verifies the integrity, elaborates approach data for one or several runways, and then transmits to users via a VHF station (VDB for VHF Data Broadcast) using VOR/ILS channels.
The onboard system can then correct its position, giving more integrity and calculate lateral and vertical deviations in relation to an approach axis.

In order to increase coverage and system capacities and have information redundancies, it is possible to group several monitoring stations in one area.
These stations collect information from satellites and transmit them to a central processing facility.
This system is called GRAS (*Ground based Regional Augmentation System*).



The VDB transmits several types of messages (8 are planned).

Messages are identified by the onboard receiver.

Type 1 message broadcasts corrections to make to pseudo ranges.
Type 2 message broadcasts data pertaining to integrity.
Type 4 message broadcasts data from final approach segment (FAS) and maximum tolerable deviations.

FAS (Final Approach Segment) messages transmitted by the VDB include the approach window of values and the alert thresholds.
Typically, the FAS is a collection of data providing the necessary information to define the geometry of the final approach.

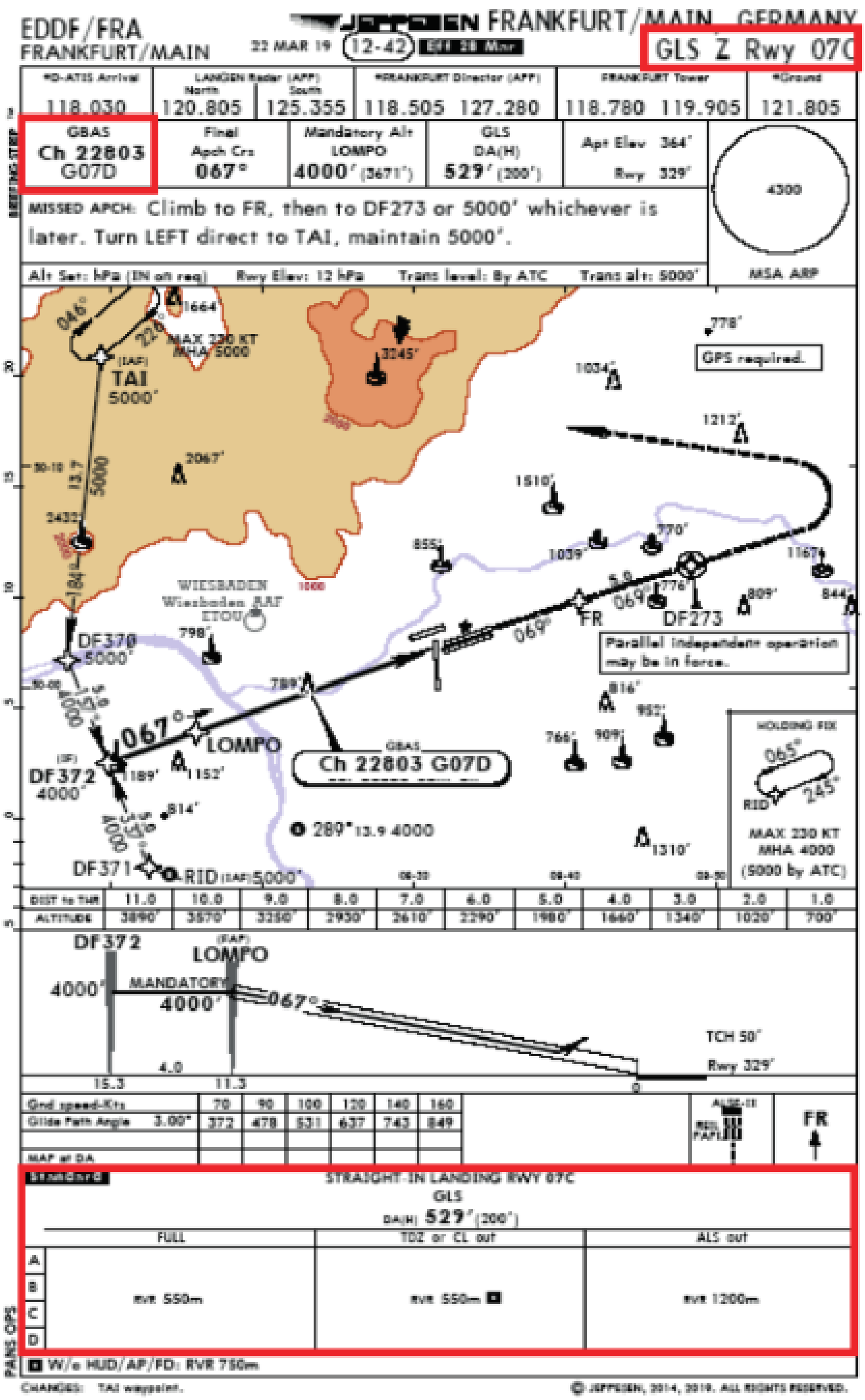
Only for information, here are some of the data among the 20 data fields of an FAS.

- Operation Type (i.e. APV)
- FPAP (Flight Path Alignment Point)
- LTP (Lat/Lon WGS84)
- LTP ellipsoidal height
- GP Angle (VPA)
- GP course width
- HAL (Horizontal Alert Limit)
- VAL (Vertical Alert Limit)
- Others

The GBAS (LAAS and GRAS) provide via these messages two types of information services used for guidance:

- A service of horizontal positioning in terminal area.
- A service of 3D positioning for the final segment *(applicable to several runways).*

An approach based on a GBAS system is called **GLS** approach (GLS – GNSS Landing System).



GLS 07 Approach FRANKFURT

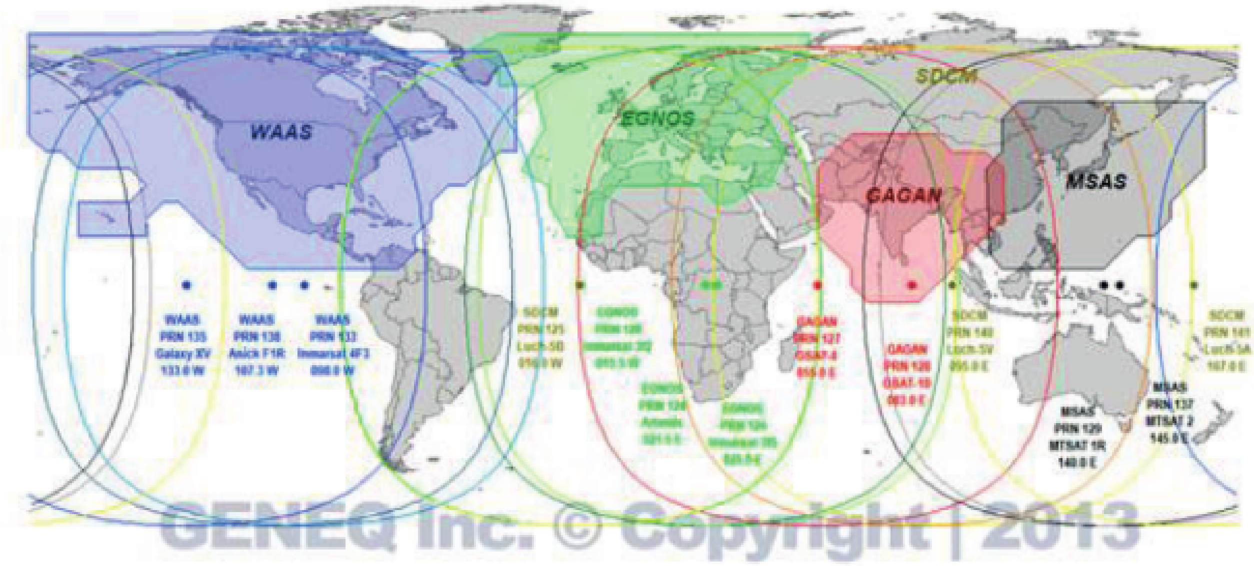
The pilot must select in the FMS (or manually) the GBAS channel (here Ch 22803) specific to each airport to tune the MMR (Multi mode receiver) on the VDB channel and check that the G07D procedure is loaded.

ILS and GLS use the same navigation instruments.

Minima here are 200 feet above ground level and RVR according to operating runway lights.

2.2 - SBAS

In order to increase the coverage of the GBAS systems, there are several SBAS programs which allow coverage on the scale of a continent.
They are the following systems:
WAAS (Wide Area Augmentation System) in the USA;
EGNOS (European Geostationary Navigation Overlay System) in Europe;
MSAS (Multi function transportation Satellite based Augmentation System) in Japan;
GAGAN (GPS And Geo Augmented Navigation) in India.



These systems, which will be entirely compatible in the future, are each made of:

- A network of reference ground stations monitoring satellite constellations.
- Geostationary communication satellites covering entire continents.
- And of course, SBAS receivers onboard the aircraft.

The role of the reference stations is to establish in real time a diagnostic on the good operation of visible GNSS satellites and geostationary satellites, to compute differential corrections to apply to the information provided directly by those satellites, to monitor integrity, and to upload this information to the geostationary satellites.

Geostationary satellites, of which, by definition, the position is perfectly known, pass on to the users the correction and integrity messages via broadcasted signals similar to those of the GNSS (for the users it is as if there were additional GNSS satellites that were always visible).

The whole system allows users equipped with a receiver intended for these services (GNSS receiver) to have, not only at least one additional satellite, always useful in areas where visibility towards the sky might be reduced, but also error corrections and warnings in case of malfunctions.

The SBAS system allows development of 3D Type A and Type B approaches.

- 3D Type A: DH ≥ 250ft
- 3D Type B: DH ≥ 200ft
3D procedure: procedure with vertical guidance (APV which will be seen in the PBN chapter)

2.3 - EGNOS (European Geostationary Navigation Overlay System)

The system is based, as all SBAS, on a ground segment made of 34 RIMS (Ranging and Integrity Monitoring Station) spread around the world.

Those stations send their data to 4 European control centers.

The master stations compute the corrections for positioning, data integrity, signal continuity and availability.
The stations also receive weather and atmospheric parameters.
They update an ionospheric delay grid (world grid 5° by 5°).
They upload the parameters on three geostationary satellites (2 Inmarsat III which are AOR-E and IOR-W and an Artemis satellite which belongs to the European Spatial Agency).

The spatial segment (the three satellites providing coverage of western Europe and the Mediterranean) pass the information on to the users (SBAS receivers of the user segment).
Concerning the ionospheric delay, the user chooses (automatically of course) 3 or 4 points on the grid that are best adapted to its position and interpolates, if necessary, and then applies the correction.

The accuracy of this system is 1 to 2 meters laterally and 3 to 5 meters vertically, with a time accuracy of 10 nanoseconds.

In case of a problem (loss of integrity for a satellite), the user can be alerted in less than 6 seconds.

2.4 - ABAS

The GPS receiver can be used alone or integrated to a navigation system.
When used alone, the augmentation rests on the RAIM studied previously.
When integrated to a system it can be aided if information from external sources (IRS, CADC) can be introduced in its calculations, for example information of time, speed or barometric altitude.
Three satellites are sufficient to determine a 3D position if the receiver is provided with the altitude for example.
Integrated in a navigation system, it can either be aided or aid other elements of the system.
When powered up for example, if the GPS cannot yet acquire satellites, it can use the position parameters entered when the INS is initialized. Afterwards it is the GPS that will limit the influence of the INS drift on the position calculation.

Based on the GNSS and other aircraft sensors (IRS for example) the ABAS system is designated by the term AAIM (Aircraft Autonomous Integrity Monitoring).

In an FMS system, if the GPS at some point only receives 3 satellites, it only provides the FMC with the time and the altitude.

Unlike the GBAS and SBAS systems which bring integrity and accuracy augmentation, an ABAS system does not improve accuracy, it only completes the integrity monitoring function.

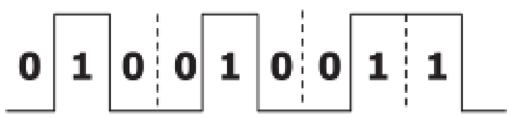
Notes not included in the course

A - Constellation interoperability

L1C is the fourth civilian GPS signal, designed to enable interoperability between GPS and international satellite navigation systems.
Its name refers to the radio frequency used by the signal (1575 MHz, or L1) and the fact that it is for civilian use. There are also two military signals at L1, as well as the legacy C/A signal. L1C should not be confused with L1 C/A.
L1C features a Multiplexed Binary Offset Carrier (MBOC) modulation scheme that enables international cooperation while protecting U.S. national security interests. The design will improve mobile GPS reception in cities and other challenging environments.
The United States and Europe originally developed L1C as a common civil signal for GPS and Galileo. Japan's Quasi-Zenith Satellite System (QZSS) and China's BeiDou system are also adopting L1C-like signals.
The United States will launch its first L1C signal with GPS III. L1C will broadcast on the same frequency as the original L1 C/A signal, which will be retained for backwards compatibility.

B - PRN

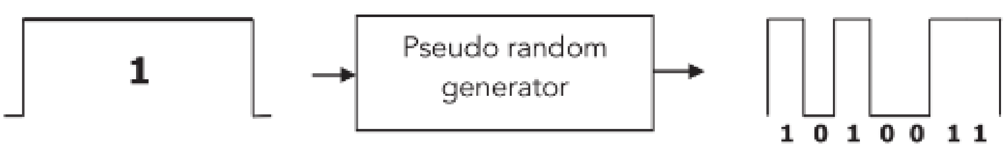
The C/A and P codes transmitted by the satellite are a sequence of 0 and 1 that constitute the elements of information given.



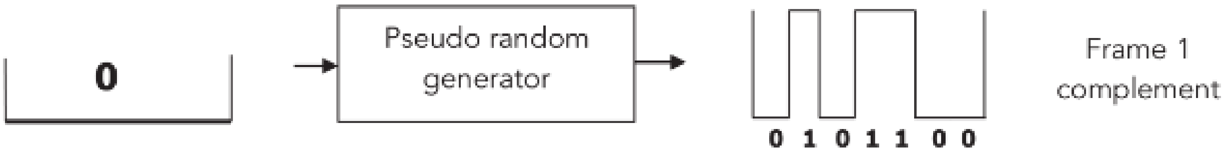
Instead of having a carrier frequency modulated directly by these bits, ~~the~~ information is transmitted through a special circuit called pseudo random generator.
This generator will transform a high level (here a 1) in a known sequence of bits and will transform a 0 by the complement to this seqence.
It is the sequence of bits (a chip) that will modulate the carrier.

Here is an example limited to 7 bits, to make it short, because the PRN of the GPS is made up of 1 023 bits.

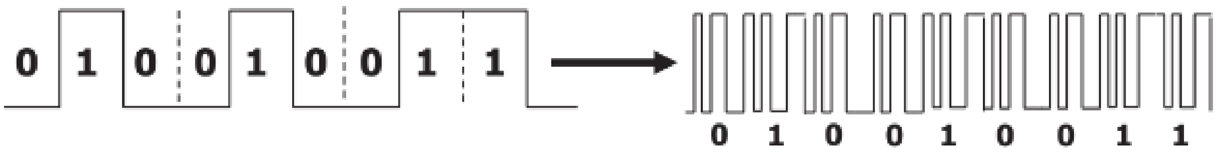
Transformation of a logical 1 into a pseudo random sequence.



Transformation of a logical 0 into a pseudo random sequence.



Information presented earlier in the chapter and its PRN code.



The initial information is coded pseudo randomly and the signal, after coding, presents a very wide and spread out spectrum that ressembles a white noise, hence the name Pseudo Random Noise.

This coding system creates information redundancy and allows, during transmission, a better error control, and even error corrections.

Small throwback on notions of computer logic from Subject 020.

Extracts from an article reproduced with the kind permission of Mr J.F. FOURCADIER, radio amateur F4DAY.

Website: http://jf.fourcadier.pagesperso-orange.fr

What is a pseudo random generator? What can it be used for?

A pseudo random generator produces a sequence of known length of logical 0 and 1. It is called random because the sequence is arbitrary. However, when the sequence arrives at its end, the generator does not stop operating. The sequence already transmitted is repeated again. Hence the qualitative pseudo random.

Example: 1 1 1 0 0 1 0 1 1 1 0 0 1 0 1 1 1 0 0 1 0 1 1 1 0 0 1 0 1 1 1 0 0 1 0...

The elementary sequence presented above is short: it has a length of 7 bits. We will see that we can easily create much longer sequences…

The pseudo random generator for data transmission plays the same role as a television test card. Even if the sequence produced by a pseudo random generator seems arbitrary, it is, by construction, perfectly defined. Specific measuring instruments, expensive but commercially available, can recognize normalized sequences and display the transmission error rate. The transmitted signal presents a very wide and spread out spectrum that resembles a noise.

C - Spread spectrum

This technique consists of using a wide frequency band to transmit data with accrued low temporal resolution power.
This way the receiver can separate the information transmitted via different propagation paths and reconstitute the starting signal.

There are two methods of spectrum spreading:

- By frequency hopping:

A band is cut in 1 MHz channels for example, then the information modulates a certain channel during a very short period of time by following a combination known by the receiver which does the same channel hopping.
A disadvantage of this method is the interruption of information during the frequency hopping.

This technique has been used for a long time by the military to prevent communication interceptions.

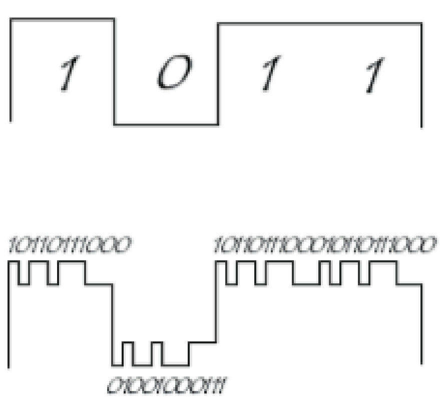
The band used is wide, but the modulation uses a narrow band technique

- By direct sequence:

The data to be transmitted, 1 or 0, are multiplied by a predefined sequence in order to obtain a wide band signal.

Example: if I must transmit some 0 and 1 at a rhythm of 10 Hz (it is my information) I create a system that correlates my 0 and 1 with a predefined sequence of bits at a rhythm of 1000 Hz (does this remind you of the PRN?) and it is this sequence of bits that modulates my carrier.
(We call chip each bit encoded by the pseudo random sequence).

Thanks to the chipping, redundant information is transmitted, which allows error controls on the transmissions, and even error corrections.



07

PERFORMANCE-BASED NAVIGATION (PBN)

01 PBN CONCEPT (ICAO DOC 9613)

According to the ICAO PBN Manual (Doc 9613), PBN aims to ensure global standardization of RNAV and RNP specifications and to limit the proliferation of navigation specifications in use world-wide.

The PBN concept, by definition, rests on the implementation of navigation specifications associated to each phase of flight and based on the concept of area navigation.

Area navigation is a method of navigation which uses a position of the aircraft that is independent of the location of the ground infrastructures.

The determination of the position of an aircraft is usually based on the following means:

- autonomous aircraft systems (IRU or INS inertial positioning);
- ground systems (DME/DME or VOR/DME type positioning);
- satellite systems.
- This navigation is based (in general) on an inboard database which includes among other things:
- waypoints defined in the worldwide reference WGS 84 (latitude and longitude);
- transitions between those waypoints;
- specific constraints (altitude, speed).

A Waypoint or WP identifies a point along a flight path. A segment on a flight path is confined and defined by the two WP at its extremities.

Area navigation differs from a method based on ground systems only (conventional navigation) and allows us to consider more direct and more efficient routes than those that can be obtained by conventional navigation.
The implementation of RNAV-RNP systems (developed below) based essentially on means of satellite radio navigation (GNSS), should enable optimized utilization of the available airspace while taking into account cost control.
To reach these objectives, we must extensively use the navigation (GNSS), communication (CPDLC) and surveillance (ADS-C) systems that are already available with various onboard equipment, to avoid expensive aircraft modifications.

Note:
Specifications exist for all areas in aviation (ATC, aircraft, crew).
For ATC, specifications exist for navigation, communication (CPDLC available or not for example) and surveillance (Radar equipment, ADS-C, ADS-B).
For aircraft (monitoring, long-haul flight means or not, and MEL).
For flight crew (composition of the crew, procedures, training, experience).

Flight crew and ATC must be aware of the capability of the on-board RNAV/RNP system to determine if the performances of the system are appropriate for the specific airspace requirements.

1.1 – PBN Principles

RNAV is the abbreviation for Area Navigation.
ICAO Annex 11 defines Area Navigation (RNAV) as a method of navigation which permits aircraft operation on any desired flight path within the coverage of station-referenced navigation aids or within the limits of the capability of self-contained aids, or a combination of these.

This concept applies to en-route navigation (permanent RNAV routes, occasional routes, published or not), but also to navigation in terminal areas (SID, STAR, RNAV approaches).
A~~n~~ RNAV navigation system allows navigation on any flight path with a required navigation performance (RNP), without having to fly over ground stations.

RNP (which responds to criteria of quality and precision gives a value of navigation precision that must be obtained during at least 95% of the time by all aircraft flying within an airspace.
The precision reflects the conformity of the true position with regard to the required position.
For instance, RNP 1 gives a navigation precision of 1 NM, meaning that within a designated airspace, the navigation precision of aircraft flying is 1 NM, with a containment probability of 95%.

But other factors also define the RNP concept:

- accuracy of the precision of the position
- integrity of navigation information
- Continuity of the position precision
- availability of navigation capability.

RNP can be specified for a route, a number of routes, an area, an airspace, or any space with dimensions defined by the competent authority in charge of planning that airspace.

A RNAV system having an onboard alerting system of the RNP that can gives an alerts to the crew in the case of excessive deviations is called ~~an~~ RNP system.

Therefore, RNAV and RNP systems are fundamentally the same with the exception of the monitoring and alerting systems.
RNP systems allow minimum variability of navigation and ensure operation of reliable, replicable and predictable flights.
Continuity is defined as the capacity of the system to carry out its functions (precision and integrity) without interruption.
Integrity is the measure of trust that can be placed in the accuracy of the information provided by the whole system.

Integrity includes the capability of a system to provide valid and timely alerts to the user.

- Three main parameters define the notion of integrity:
  - the alert limit;
  - the integrity risk (probability of exceeding the alert limit);
  - The alerts delay

- The measure of integrity of the GNSS can be based on:

  - regional or local augmentation systems (ex: EGNOS, DGNSS);
  - autonomous mechanisms using redundancy of signals and sensors (ex: RAIM, AAIM);
  - systems that supply their own integrity service such as Galileo.

Availability is defined as the percentage of time (annually) during which the system is available for use.

PBN is not specific to the aircraft sensors.

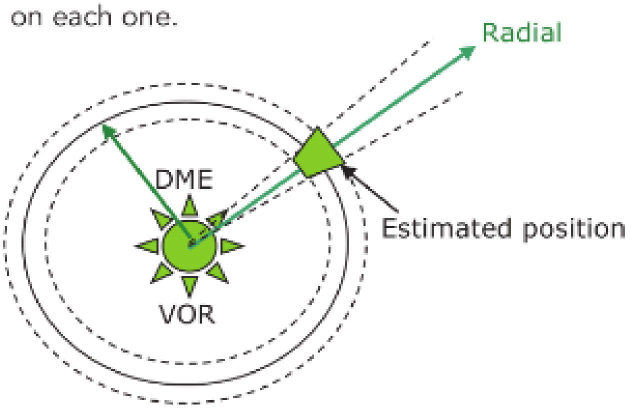
Initially, conventional navigation was based on specific aids (NDB, VOR, VOR/DME, etc.).
Except when flying overhead a radio aid, navigation precision decreased progressively as distance was increased from the radio aid
PBN, being based on the concept of area navigation, solves this problem and provides the same precision on all points of the flight path.
The big feature of the PBN concept is the specification of a minimum performance requirement without taking into account specific equipment but by imposing navigation capabilities.

PBN offers quite a few advantages in comparison to the sensor specific method:

- it reduces the need to maintain sensor-specific routes and procedures, and their associated costs;
- it avoids the need for developing sensor-specific operations with evolving navigation systems, which would be cost-prohibitive;
- it allows for more efficient use of airspace (route placement, fuel efficiency and noise abatement);
- it clarifies how RNAV systems are used; and
- it facilitates the operational approval process for operators by providing a limited set of navigation specifications intended for global use

Area navigation uses computed data.

Navigation raw data is information delivered by specific radio navigation aids such as VOR, DME, ILS or GPS (Navigation sensors).
The provided data is relative to the installation of the ground navigation aids (VOR, DME, ILS).
These aids do not share their information and provide a position precision that depends entirely on each one.



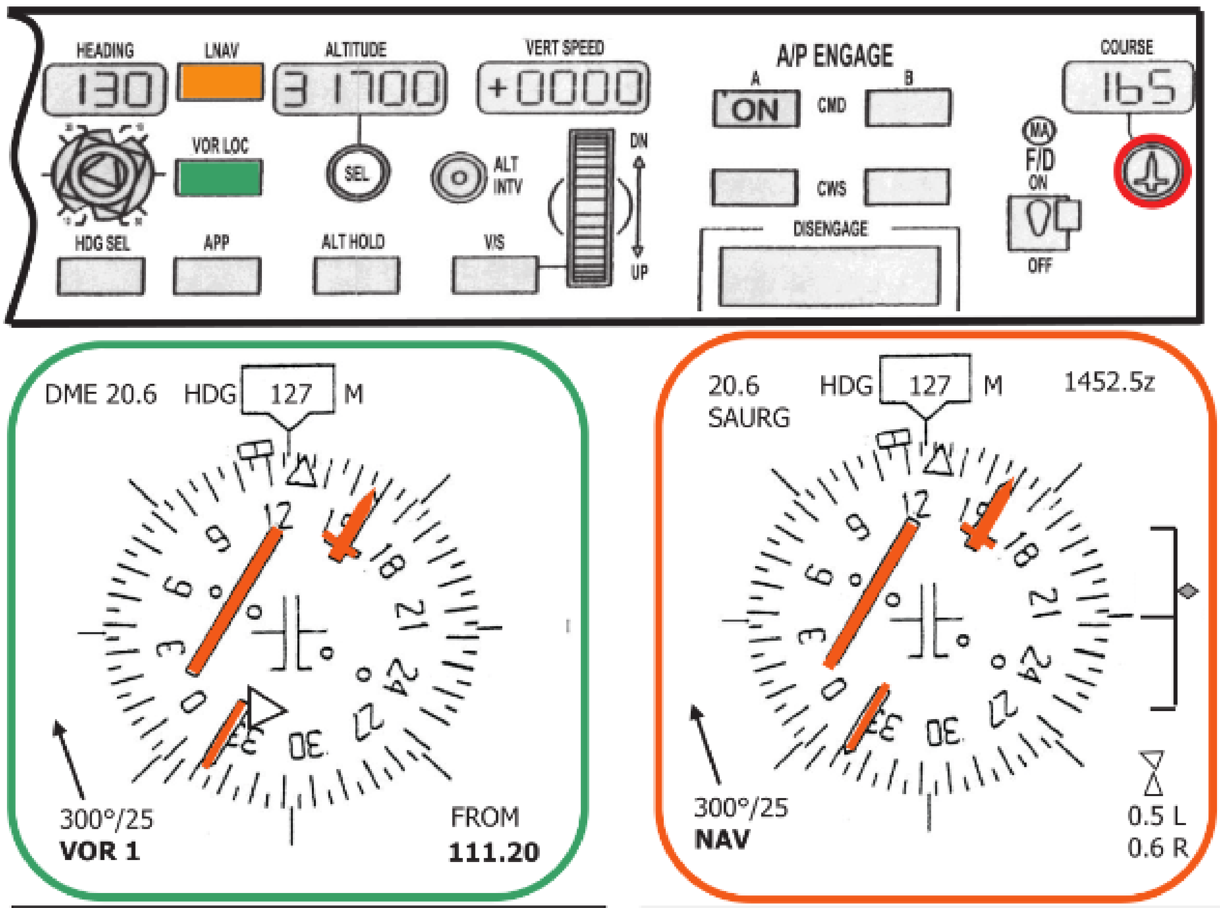
An example, on the left.
A position based on a VOR/DME is dependent on the imprecision of the slant distance provided by the DME and the radial provided by the VOR.

Computed data is data calculated by a computer executing numerous navigation algorithms. This computer, after validation of the available raw data, gives flight crew the most probable exact position in relation to a desired route irrespective of the location of the ground aids.
Most systems also compute data consistency.

Example of computed data:

- Information on distance will be a ground distance calculated from the slant DME range because the computer takes into account the altitude of the aircraft.
- The position is not displayed by using a lateral deviation in degrees with respect to a desired route but by a lateral deviation expressed in NM.

Example B737, two almost identical displays and yet…:
VOR/LOC selected on AP/FD control panel and raw data display on ND in Full mode.
LNAV selected on AP/FD control panel and computed data display on ND in Full mode.



Upper left, the DME slant range is 20.6 NM.
The course deviation indicator is activated by the pilot with the COURSE knob (165).
The deviation indicator indicates a deviation expressed in degrees in relation to the selected course (165).
All of this data comes from the VOR/DME selected by the flight crew (VOR1 – 111.2).

Upper left, the ground distance from the present position to the TO Waypoint (SAURG) is 20.6 NM.
The course deviation indicator is activated by the FMC according to the active leg of the flight plan (COURSE knob is deactivated).
The deviation indicator indicates a lateral deviation in NM in relation to the Flight Plan route.
All of this data comes from the Flight Management Computer.

1.2 - PBN components

PBN is based on area navigation and comprises three components:

- Navigational aid (NAVAID) infrastructure,
- Navigation specification,
- Navigation application.

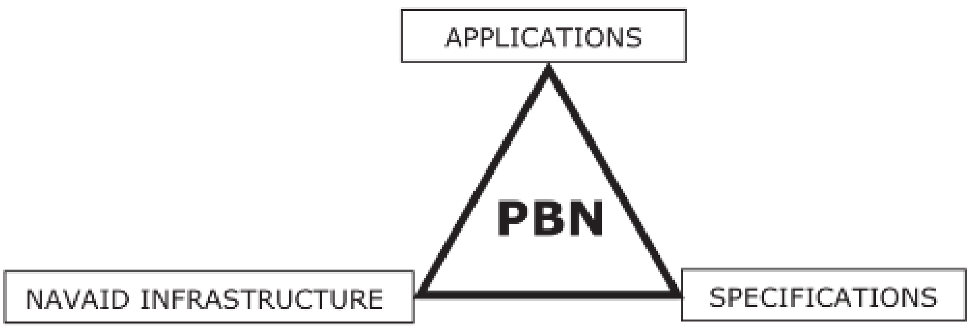
The infrastructure of navigational aids refers to navaids on the ground and in space.
The NDB is excluded from PBN.
Precision approaches (ILS, MLS or GBAS) are not part of PBN.

Navigation specifications are technical and operational specifications that define ~~the~~ navigation performance and the required functionalities of the RNAV system.

They also determine how navigation equipment is supposed to work to operate within the infrastructure of navigational aids in order to meet the operational needs that apply to the designated airspace concept.

Or more simply:
A navigation specification is defined as a set of conditions that an aircraft and its flight crew must fulfill to carry out a PBN flight in a defined airspace.

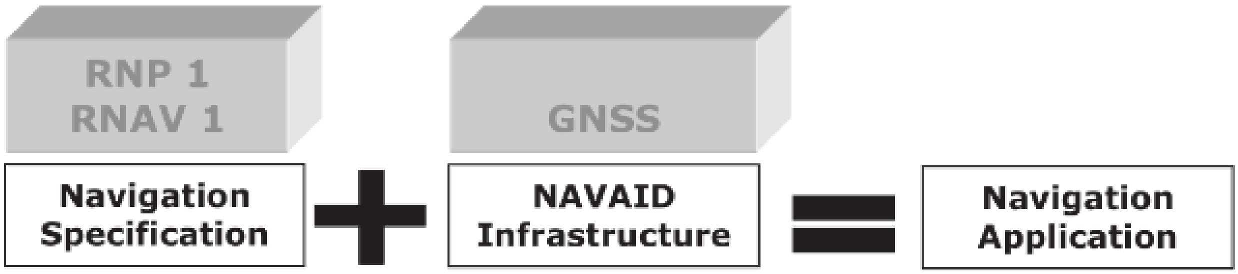
Navigation applications are based on the combined use of NAVAID infrastructure and navigation specifications.



There are two types of navigation specifications:

RNAV specifications:
Navigation specifications that do not require on board monitoring and alerting systems.
RNP specifications:
Navigation specifications that do require on board monitoring and alerting systems.

Example:
As expressed earlier, the application (creation of an ATS route for example) is based on RNAV or RNP specifications (choice of the State), which depend among other things on the surrounding infrastructure (DME sensors, GNSS).



1.3 - PBN Scope

Oceanic or remote areas and terminal phase of flight.

In oceanic or remote airspace, en-route and terminal phases of flight, PBN is limited to operations with linear lateral performance requirements and time constraints.

Approach phases

During approach phases, PBN allows operations with linear lateral guidance and also angular guidance.

Annex I of Commission Regulation UE n° 1178/2011 specifies that the following definitions that are inserted in Part-FCL.010:

A "linear operation" means an instrument approach operation in which the maximum tolerable error/deviation from the planned path is expressed in units of length, for instance nautical miles, for cross-track lateral deviation.

An "angular operation" means an instrument approach operation in which the maximum tolerable error/deviation from the planned path is expressed in terms of deflection of the needles on the Course Deviation Indicator (CDI) or equivalent display in the cockpit.

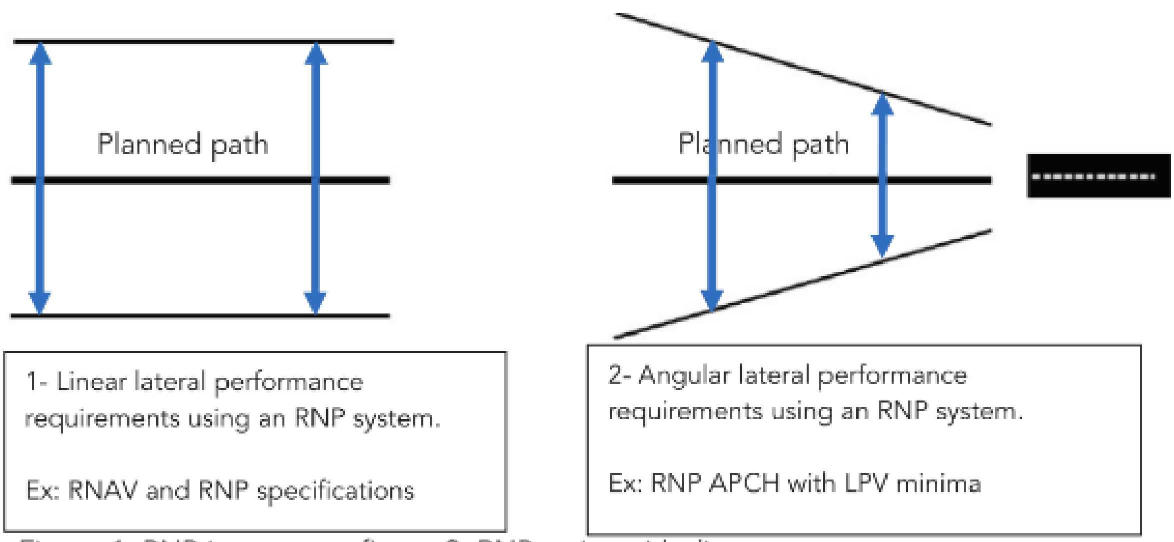
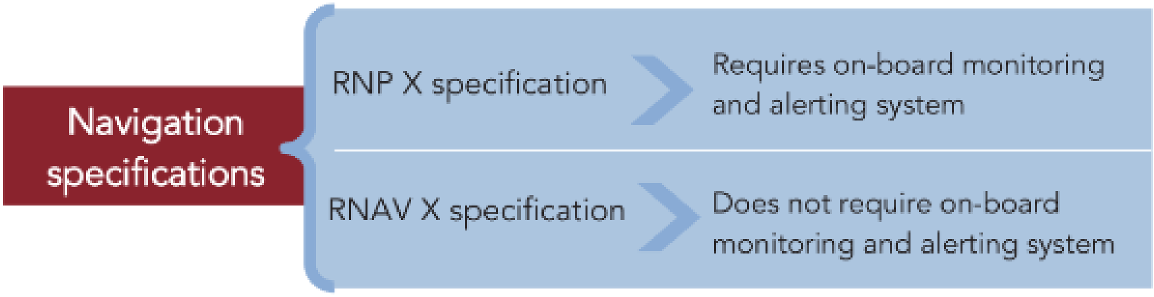


Figure 1, RNP is constant, figure 2, RNP varies with distance.

Procedures with angular lateral guidance have the appearance of ILS guidance, but only the appearance, because they do not require navigation aids on the ground.
LPV: Localizer Performance with Vertical Guidance.

02 NAVIGATION SPECIFICATIONS

2.1 - RNAV and RNP



Navigation specifications based on satellite positioning are called RNP.

They normally differ from each other by a number (RNP "x") expressing the associated navigation precision in NM.
Just as it manages the degree of confidence, the positioning system estimates the maximum error it makes.

As the error can exceed the threshold equivalent to the number associated to the RNP, a monitoring and alerting system will inform the flight crew.

With some exceptions, certified onboard GNSS equipment is required to comply with an RNP x navigation specification.

When the system that computes the position of the aircraft is not able to determine the integrity of the computed position, the navigation specification is called RNAV.

This is also associated to a numerical value that represents the precision as a maximum estimated error. It is the case of navigation systems which compute an aircraft position using radio beacons only such as DME or VOR-DME.

The flight crew however are not informed by the navigation system when it deviates from its defined path because of an erroneous position computation (absence of "likelihood tests").

Navigation specifications (such as RNAV "x") can be still be used when other means allow monitoring of potential deviations that are not controlled by the system or by the flight crew (radar monitoring by a controller for example).

2.2 - Navigation functional requirements

Again, RNAV system and RNP system function in the same way except for the crew monitoring and alerting system.
They must both be able to follow with precision a ground track, whether straight or curved.
Basic minimum indications that must be presented by both systems are:

- A continuous indication of lateral deviation from the active flight plan route (Navigation Display for example).
  The indicator must be calibrated according to the required precision.
- Distance/bearing from present position to active waypoint.
- Ground speed or time to active waypoint (also called TO waypoint).

And of course the system generally includes a navigation database to benefit from all the applications (possible exception with RNAV 5).

The monitoring and alerting component of an RNP system must provide in some form:

- The required and estimated precision levels.
- Monitor the system in terms of precision (lateral and longitudinal) and integrity.
- Alert the crew if RNP requirements are not met.

NB:
An ICAO document states that FAA and EASA have slightly different definitions of what constitutes an RNP system, but I have not found trace of the slight difference.

2.3 - Designation of RNP and RNAV specifications

RNAV or RNP performances are given a value x (RNAV x or RNP x).
This value refers to a lateral accuracy which a group of aircraft operating within a given airspace, route, or procedure is expected to achieve during at least 95% of the flight time.
The value of x depends on the total error of the system (TSE – Total System Error) developed in section Chapter 07 § 4.1.
Performance requirements are defined for each specification type.
An aircraft approved for a specification is not necessarily approved for another specification.
An aircraft approved for an RNP or an RNAV specification with a stringent accuracy requirement (ex: RNP 0.3) is not necessarily approved for a specification that has a less stringent accuracy requirement (ex: RNP 4).

For example, RNP 1 approval requires one GNSS with additional constraints.
RNP 4 approval requires one GNSS and two long-range navigation systems with additional constraints.
Therefore, an RNP 1 approved aircraft will not be RNP 4 if it does not have a long-range navigation system.

RNAV 10 (RNP 10) and RNP 4:

RNAV 10 and RNP 4 specifications are used in the oceanic and remote airspace.
The total error of the system must be less than or equal to 10 NM or 4 NM during at least 95% of the flight time.
RNAV 10 specification is based on either INS, IRS or GNSS equipment and allows aircraft spacing of 50 NM (lateral and longitudinal).
RNP 4 specification is based on GNSS and allows aircraft spacing of 23 NM (lateral and longitudinal). -amendment 7 PANS-ATM of 10 November 2016-

RNAV 5:

RNAV 5 is used in the en-route and arrival phases of flight.
It is based on the VOR/DME, DME/DME, INS, IRS or GNSS.
It is the equivalent of B-RNAV (as it was formerly called by ECAC) initially designed to use conventional radio navigation aids.

RNAV 2 – RNP 2:

RNAV 2 and RNP 2 are also used as navigation specifications.

RNAV 2 might be used in the en-route, continental, arrival and departure phases of flight.
It is based on a DME/DME, DME/IRU or GNSS infrastructure.

RNP 2 is used in the en-route and oceanic and remote phases of flight.
It uses the GNSS as its means of navigation.

RNAV 1 – RNP 1:

For phases of flight that correspond to arrival or departure procedures, airports can implement RNP (RNP 1) or RNAV (RNAV 1) type navigation specifications depending on traffic density, radar equipment, or even communication means.

RNAV 1 is used for the arrival and departure phases of flight.
It is the equivalent of P-RNAV (as it was formerly called by ECAC).
It is based on the DME/DME, DME/IRU or GNSS.

RNP 1 is used for arrival and departure phases of flight in areas of reduced traffic with or without reduced ATS monitoring.
It uses the GNSS.

RNP APCH - RNP AR APCH:

For approach phases, only RNP navigation specifications can be implemented.

They are designated by: RNP APCH and RNP AR APCH. (AR stands for Authorization Required)

The main differences between the RNP APCH and RNP AR APCH specifications are:

- The RNP value that can be used on the final approach segment is 0,3 NM for RNP APCH specification and varies between 0,3 NM and 0,1 NM for RNP AR APCH specification.
- To carry out a procedure based on the RNP AR APCH specification, the navigation system must have an "RF" (Radius to Fix) capability (explained later), to execute constant radius turns.
- Vertical guidance during final approach is systematically associated to an RNP AR APCH specification procedure.
- RNP AR APCH operations require special authorization for aircraft and flight crew.

Note:
"Authorization Required" does not mean that the pilot must request an authorization for each approach but rather that the operator has received approval for its aircraft and flight crew to perform this type of approach.

RNP 0.3:

Both RNP APCH and RNP AR APCH specifications at some point use a precision of 0,3 NM, but there is also an RNP 0.3 specification.
It is primarily for helicopters and uses the GNSS.
RNP 0.3 can be used in all phases of flight except for oceanic and remote areas, and final approach.
RNAV 1, RNP 1 and RNP 0.3 can also be used for en-route phases of low-level instrument flight rule (IFR) helicopter flights.

In summary:

Navigation Specification	Flight Phase							
	En Route Oceanic Remote	En Route Continental	ARR	Approach				DEP
				Initial	Intermed	Final	Missed	
RNAV 10 (RNP 10)	10							
RNAV 5		5	5					
RNAV 2		2	2					2
RNAV 1		1	1	1	1		1	1
RNP4	4							
RNP2	2	2						
RNP1			1	1	1		1	1
Advanced RNP	2	2 or 1	1	1	1	0.3	1	1
RNP APCH				1	1	0.3	1	
RNP AR APCH				1-0.1	1-0.1	0.3-0.1	1-0.1	
RNP 0.3		0.3	0.3	0.3	0.3	-	0.3	0.3

03 USE OF PERFORMANCE-BASED NAVIGATION (PBN)

3.1 - Intentionally left blank
3.2 - Intentionally left blank

3.3 - Specific RNAV and RNP system functions

RNAV and RNP systems must have specific functions that ensure reliability, repeatability and predictability of air navigation. This is essential to the development of more efficient airspace routes and procedures.

To achieve this, area navigation systems must define waypoints by latitude and longitude and must allow navigation on any flight path, in particular those with a fixed radius (FR – see below).

The system must allow to create and monitor limited and standardized route segments so that all aircraft fly on the same flight path on an airway or on a given procedure.

This function is provided by the Flight Management Computer (FMC) associated to a suitable database within the Flight Management System (FMS).

Flight path monitoring function.
This is normally assigned to the pilot using a lateral deviation indicator with adequate sensitivity for the required RNP.
Certain route segments, for example Radius to Fix (RF) below, require an autopilot able to follow a circular arc.

Navigation system monitoring function.
Normally automatically performed by the FMS and the GNSS augmentation systems.
The system must be able to display in real time the "estimated position uncertainty" (EPU) and the required RNP.
The system must also be able to display in real time an alert message like UNABLE REQD NAV PERF–RNP on the CDU (Control Display Unit) of the FMS or on the ND (Navigation Display) or on the PFD (Primary Flight Display) as soon as the EPU exceeds the required RNP.

Path terminator:
ARINC 424 specifications are a standard used to translate paper procedures published by all countries states into coded computer language that is understood by the navigation database.
Those procedures (SID, STAR, IAP) are made up of navigation segments called Path Terminators that are connected to each other.
There are 23 types of Path Terminators used for departure, arrival and approach phases of flight that can be stored in the database and are designated by a 2-letter code (for example RF for Radius to Fix).
They are defined by their path and a point called "terminator".
The path describes the flight path to follow to reach the terminator.
The terminator is the event or the condition that forces the navigation computer to move on to the next step (reach another point, an altitude, intercept, etc.).

Example:

The departure procedure indicates to climb following the runway axis until 3000 ft and then proceed direct to MOCCA.
The path is the runway axis (a track) and the terminator is altitude 3000 ft (any point where 3000 ft is reached).

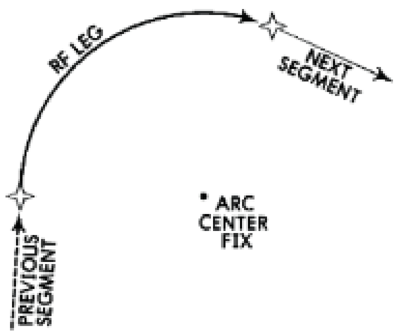
Radius to Fix

Radius to Fix (RF) is a flight path that follows a circular arc with a radius that is imposed by a procedure (SID, STAR, Approach).
For example, Advanced RNP (see later) requires this functionality.

RF is defined by the radius of the curve, the length of the arc and a point (the terminator.

RNP systems that are able to comply with this requirement guarantee the same navigation accuracy during a turn than during a straight segment, which makes it possible to ensure separation from obstacles, to comply with environmental directives (noise) and therefore optimizes procedures.

Procedures takes into account bank angle limits for different aircraft and altitude winds.



Fixed Radius Transition

Fixed Radius Transition (FRT) is a variant of the RF used in en-route procedures to connect two route segments.
FMS mode is EN ROUTE.
The turn radius is 22,5 NM for routes above FL195 and 15 NM for lower altitude.
It allows reducing the distance between parallel routes in the airspace.

It is extremely important while flying manually to follow the flight director and the speed constraints during these transitions.

Other « Path terminators »

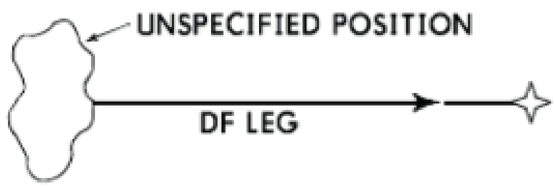
Among the "Path Terminators" frequently encountered in a procedure, there are:

IF (Initial Fix): Defines a point in space.

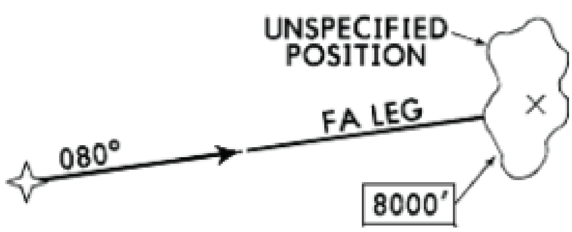
TF (Track to Fix): Direct route between two waypoints (it is an orthodromy).

CF (Course to a Fix): Segment that ends at a waypoint (ex: runway axis until 3000 ft then track 080° to MOCCA, the portion of flight between 3000 ft (previous terminator) and MOCCA is a CF).

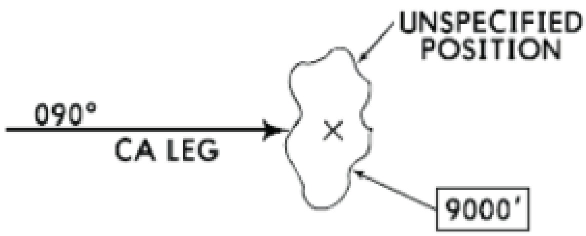
DF (Direct to Fix): Segment of a flight path from an unspecified position to a waypoint.



FA (course from a Fix to an Altitude): Segment of a flight path that starts at a waypoint and ends when the aircraft has reached an assigned altitude. The position is when this happens.

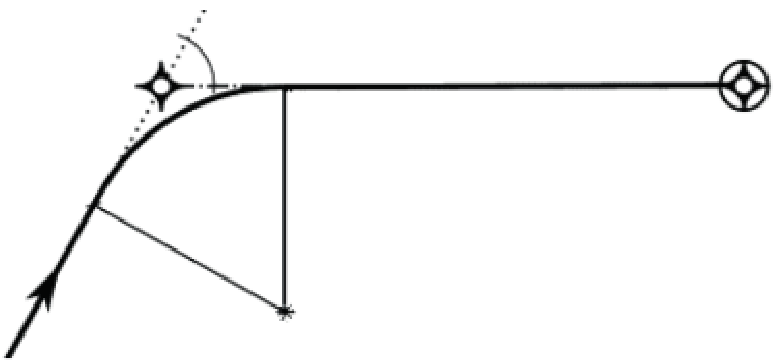


CA (Course to an Altitude) Segment of a flight path that ends at an altitude without a specific position.

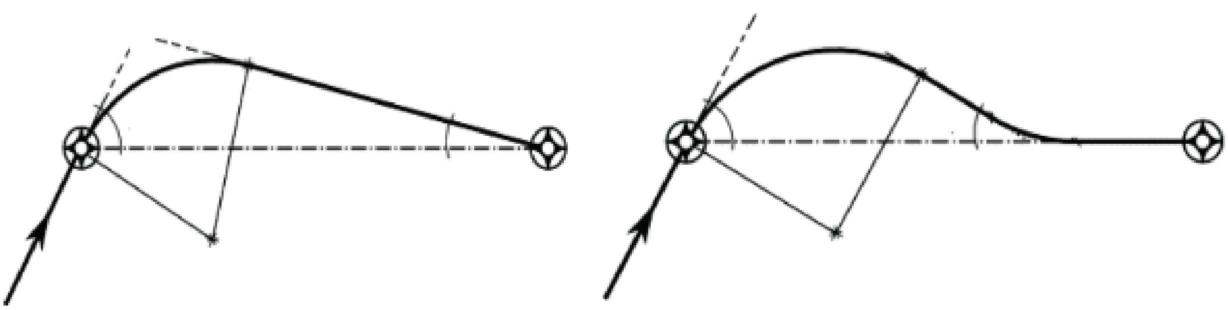


FLY BY TURN and FLY OVER:

A Fly by Turn allows a smooth transition from one flight segment to another.
The aircraft "cuts corners".
The computer anticipates the turn and calculates the required bank angle using speed and wind information.
Turn radius parameters can vary from one aircraft to another.



A Fly Over forces the aircraft to fly over the path terminator before turning to connect to the next flight segment.
The connection flight path will also vary from one aircraft to another.
The flight crew must be very careful during these transitions.



Holding pattern:

A holding pattern is defined by a holding fix, direction of turn, the inbound leg, and the ability to plan exiting from the hold.

PBN allows holding wherever we want.

Route OFFSET:

A route OFFSET is a route parallel to the flight plan with an offset of X NM on the right or on the left side, with offsets generally in increments of 1 NM (from 1 to 20 NM).
OFFSET navigation does not affect the initial route of the active flight plan.
An OFFSET amber alert on the CDU reminds the flight crew that it is flying with an offset.
This function can be used for example by ATC to let a faster aircraft overpass a slower one, and it is an alternative to radar vectoring.

04 PBN OPERATIONS

4.1 - Principles

Path Definition Error:

A navigation area route is defined by segments between waypoints.
The definition of the path depends on precision of the waypoints and the capacity of the navigation system to manage the waypoint data.
The Path Definition Error (PDE) is the difference between the desired theoretical route and the one defined by the area navigation system.
However, waypoints can be defined very precisely and most navigation systems can manage a high level of precision so this error is minimal and generally considered to be zero.

Flight Technical Error:

The FTE (Flight Technical Error) is also called flying error (by the autopilot or the pilot flying manually) from the flight path.
There can be longitudinal and vertical FTE.
This value represents the capacity of the aircraft guidance system to follow a calculated flight path. The FTE is usually evaluated by the aircraft manufacturer on the basis of flight tests.
FTE values will generally vary for a specific aircraft depending on the aircraft guidance system used.
For example, a lower FTE can be applied to operations where the autopilot is coupled to the flight director in comparison to manual flying using the flight director.
This variation can in turn lead to overall performance values that differ according to the control method.

Navigation System Error:

The NSE, Navigation System Error, (sometimes called EPE for Estimated Position Error) is the value of the error of the navigation avionics in the determination of the position, compared to the actual position of the aircraft.
NSE depends on the precision of the inputs and the position resolution.
It defines the precision of a navigation system.
In general, for a system with GPS/SBAS, NSE<< FTE.

Total System Error:

The Total System Error (TSE) is the geometric sum of the component errors.

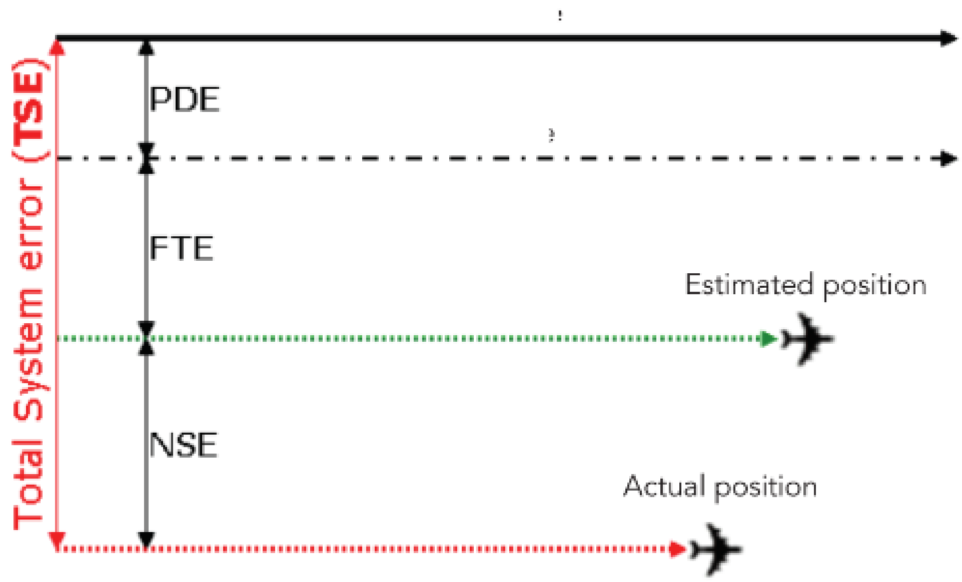
$$TSE = \sqrt{NSE^2 + FTE^2 + PDE^2}$$

The TSE is the achieved navigation accuracy (in NM) in real time, and it must be lower than the required accuracy during at least 95% of the flight time.

FTE is easy to monitor.

NSE is low but difficult to determine.
TSE is always much lower than the required precision.

Example of a TSE survey of many flights on the same runway for an RNP AR APCH:
The observed TSE is around 18 meters for an RNP of 0,3 NM, which is 528 meters.
The TSE is thus 15 times lower than the required precision.



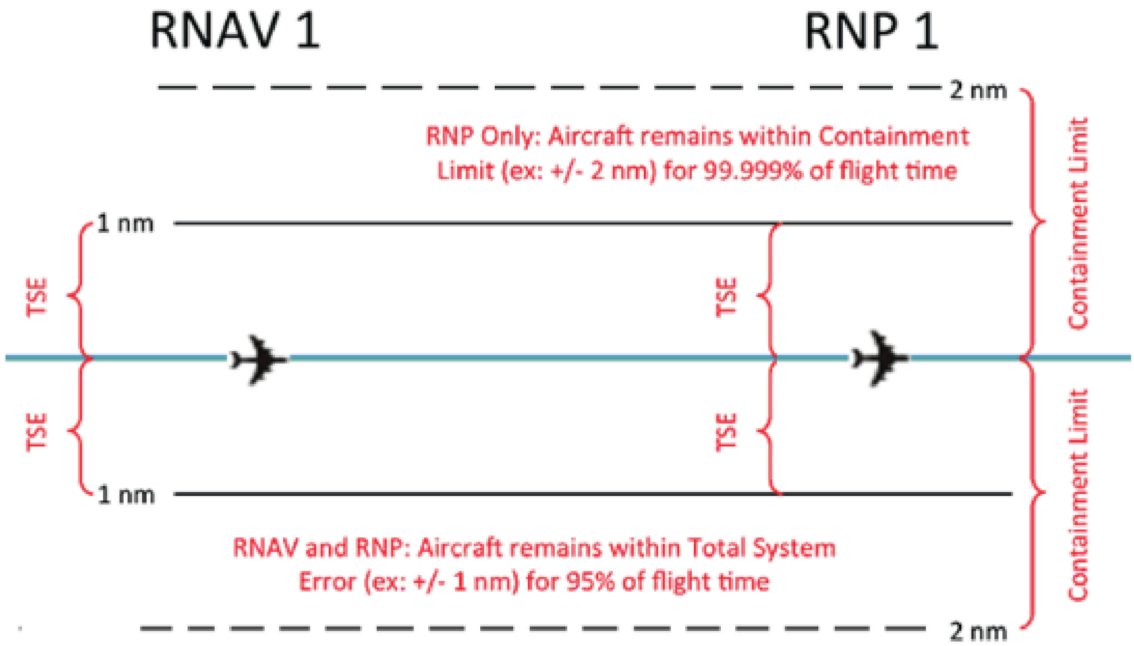
4.2 - On-board performance monitoring and alerting

What differentiates RNP from RNAV?

Let's first look at what they have in common:

Both RNAV and RNP systems define navigation precision as being able to stay within a total system error at least 95% of the total flight time.
In the following example, with RNAV 1 or RNP 1, the aircraft will be within less than 1 NM 95% of the flight time.

However, in a RNP system:
The TSE must be at maximum 1 NM during 95% of the flight time, but as soon as the probability of being at 2 NM (here 2 times the RNP) exceeds 0,001% the system alerts the flight crew.

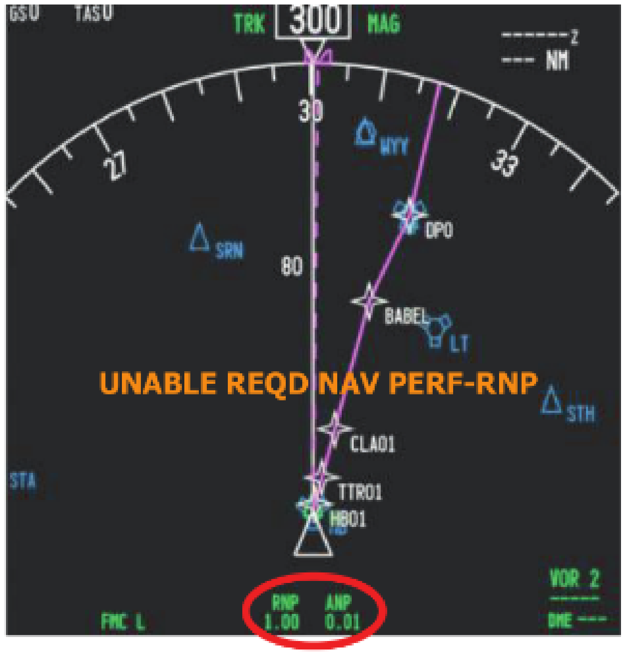


Monitoring of on-board performances must not be considered as error monitoring.
Alerts are issued when the system cannot guarantee with sufficient integrity that the position satisfies precision requirements.
When an alert is issued, the likely reason is the loss of capability to validate the position data (for example an insufficient number of satellites).
In other words, even if the position complies with precision requirements, because the system cannot prove it, an alert will be issued.

The NSE is also referred to as EPE - Estimated Position Error – but the most common terms that appear in the cockpit are ANP (Actual Navigation Performance), EPU (Estimated Position Uncertainty) or ACTUAL.

The EPE (or ANP) is defined as the statistic limit on the NSE and not the TSE.

The Actual Navigation Performance is computed by the FMS.
Many sources of independent navigation data such as the IRS or GNSS can be included in the determination of the ANP.
These data sources are continuously analyzed in order to best the actual position of the aircraft and the ANP.



A320 on PROG page of the FMS

If EPE < RNP the ACCUR field will indicate HIGH

If EPE > RNP the ACCUR field will indicate LOW

ND Boeing 747
RNP and ANP displayed at the bottom.
Crew alerting message.
This is only an example and is not representative of a real flight situation.

If the navigation precision goes from HIGH to LOW, a message warns the flight crew on the Navigation Display (ND) and on the Multipurpose Control Display Unit (MCDU).

Performance monitoring and alerting			
	RNAV specifications	RNP specifications	
		RNP X specifications not requiring RF or FRT capability	RNP X specifications requiring RF or FRT capability
NSE Monitoring and Alerting	No requirement for position alerting or comparison of the NSE.	Alerting based on navigation precision and integrity.	
FTE Monitoring and Alerting	Provided by the on-board system or by the flight crew.	Provided by the on-board system or by flight crew procedures. Display on indicator with appropriate scale.	
PDE Monitoring	Error assumed equal to zero. Desired route is not defined in turn.		Assumed equal to zero. Flight path defined in RF and TRF.

4.3 - Abnormal situations

Abnormal and contingency procedures are to be used in case of loss of the PBN capability.

- Failure of navigation system components, including those affecting the technical flight errors (for example failures of the flight director or the autopilot);
- RAIM warning or loss of the integrity function failures;
- Warning flag or equivalent indication on the lateral and/or vertical navigation display;
- Degradation of the GNSS approach mode during an LVP approach procedure (for example, going from LPV to LNAV);
- Low altitude warning (when applicable).

In Europe, for RNP AR APCH, the RNP performance also takes into account the effect of abnormal events, and different RNP performances can be specified in accordance with operational circumstances.
Different RNP values are published for an aircraft with N and N-1 engines.
The conception of an ICAO approach procedure does not take into account abnormal conditions and RNP for all engines operating is applicable, but limitations set by the manufacturer should be taken into account during the FOSA (Flight Operational Safety Assessment).

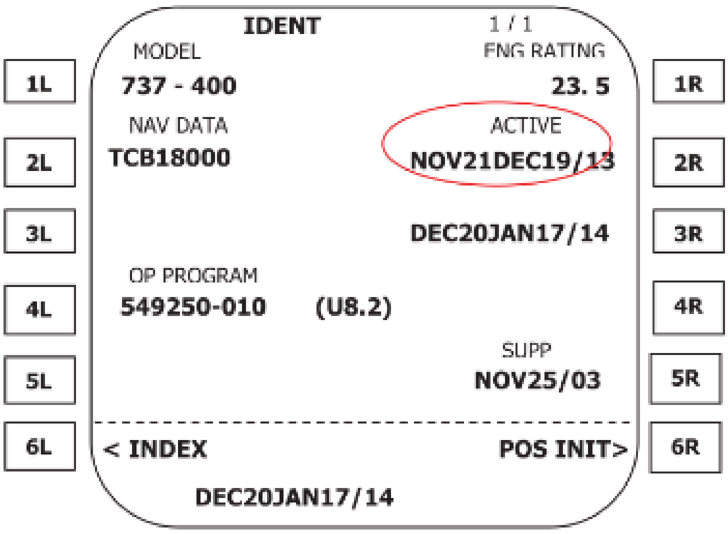
Source EASA Workshop, 20th October 2010
The FOSA should ensure that for each specific set of operating conditions, aircraft and environment, all failure conditions are assessed and, where necessary, mitigations are implemented to meet the safety criteria.
The flight crew must inform ATC of any problem with the navigation system that causes the loss of approach capability.
An UNABLE RNP warning during the approach requires a go around.

4.4 - Database management

The navigation database must include all the data/information necessary to follow the published approach procedures.
Therefore, the onboard navigation data must be valid for the current AIRAC cycle and must include the appropriate flight procedures.
The operator shall implement procedures that ensure the timely distribution and insertion of current and unmodified electronic navigation data to all aircraft.
During the initialization of the navigation system, the flight crew must verify the validity of the database.

Arrival, approach or departure procedures should not be used if the validity of the navigation database has expired.
Navigation databases should be up to date for the duration of the flight. If the AIRAC cycle is due to change during flight, the flight crew should follow procedures established by the operator to ensure the accuracy of navigation data, including the suitability of navigation facilities used to define the routes and procedures for the flight.

The validity of the database is checked by the flight crew before the flight on the IDENT page of the FMS.
The database for cycle N+1 can be loaded.
By pressing the 3R key, the N+1 cycle database is copied in the scratchpad of the CDU, and it is activated by pressing the 2R key.



AMC2 CAT.OP.MPA.175 - Flight preparation

An expired database may only be used if the following conditions are satisfied:
(1) the operator has confirmed that the parts of the database which are intended to be used during the flight and any contingencies that are reasonable to expect are not changed in the current version;
(2) any NOTAMs associated with the navigational data are taken into account;
(3) maps and charts corresponding to those parts of the flight are up to date and have not been amended since the last cycle;
(4) any MEL limitations are observed; and
(5) the database has not expired by more than 28 days.

05 REQUIREMENTS OF SPECIFIC RNAV AND RNP SPECIFICATIONS

5.1 - RNAV 10

Note:
RNAV 10 having often been implemented after the arrival of the RNP concept, there are airspace, routes and approvals called RNP 10 that are actually RNAV (no monitoring and alerting system).
The RNP 10 designation was maintained to avoid redoing the manuals (for cost savings) and the approvals, but the PBN applications are RNAV 10.

RNAV 10 (RNP 10) in oceanic or remote areas cannot rely on ground-based navigation aids (except in rare cases). Aircraft navigation must be based on long-range navigation capability with inertial or global positioning navigation systems.

Aircraft must be equipped with at least two independent and serviceable long range navigation systems (LRNS), comprising either an Inertial Navigation System (INS), an Inertial Reference System / Flight Management System (IRS/ FMS) or a Global Navigation Satellite System (GNSS) so that the system does not present an unacceptable probability of misleading information.

Time limits for aircraft equipped only with INS or IRS:

Inertial systems approved in accordance with FAA 14 CFR, Part 121 Annex G or equivalent documents satisfy the requirements of RNP 10 but are limited to a flight time of 6.2 hours.
Flight time begins when the system is set to Navigation and ends at the last point when it was updated.
Flight time of 6.2 hours is based on an inertial system with a radial position error of 2 NM/hour during 95% of the navigation, which is to say an orthogonal error rate (longitudinal and lateral) of 1,6015 NM/hour. (1,6015NM/hour x 6,2h=10NM).

Source: https://www.faa.gov/documentLibrary/media/Order/8400.12.pdf
Paragraph 13 d) of document 8400.12 also specifies that for flights that allow the navigation position to be updated, the operator can define a higher time limit if the effects of the update on the precision of the position and the associated time limits can be demonstrated.
For example, an applicant can establish an extended time limit by showing that the use of multiple navigation sensors and that mix or average navigation position errors justify such an extension (triple-mixed INS).

5.2 - RNAV 5

Manual data entry:

A navigation database is not a required functionality for the RNAV 5 application.
The absence of a navigation database necessitates manual waypoint entry which significantly increases the potential for waypoint errors.
Paragraph 7.6 of AC 91-002 specifies that in this case the system must be able to store a minimum of 4 waypoints.

5.3 - RNAV 1/RNAV 2/RNP 1/RNP 2

Specificities:

Aircraft approved for RNAV 1 or RNAV 2 SID or STAR can follow these applications only if they are retrievable by their route name from the onboard navigation database and are conform to the charted route.
Before flight, the flight crew should identify which parts of the flight will be conducted in RNAV 1 / RNAV 2 airspace and verify entry and exit waypoints.
The flight crew should cross-check the cleared flight plan by comparing charts or other applicable resources with the aircraft FMS and map display, if applicable.
If required, the exclusion of specific navigation aids should be confirmed.

Note:
Pilots may notice a slight difference between the navigation information portrayed on the chart and their primary navigation display.

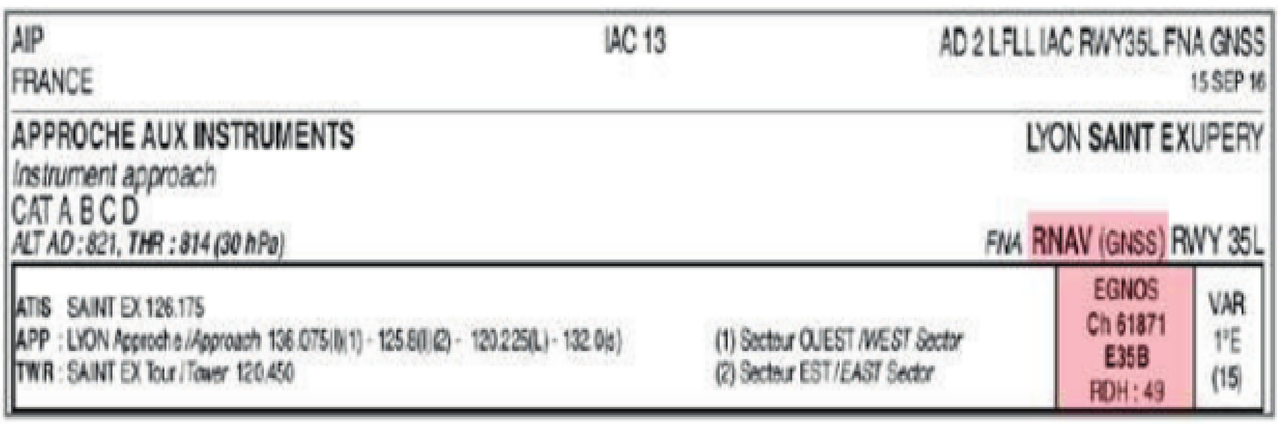
Differences of 3 degrees or less may result from the equipment manufacturer's application of magnetic variation and are operationally acceptable (AMC 036).
The route may subsequently be modified through the insertion or deletion of specific waypoints included in the database in response to ATC clearances.
The manual entry or creation of new waypoints in the database (by latitude and longitude or by radial/distance values from a navigation aid) is not permitted.
Additionally, pilots must not change any waypoint type from the RNAV SID or STAR database (Ex: change a fly-by to a flyover or vice versa).

5.4 - *Intentionally left blank*

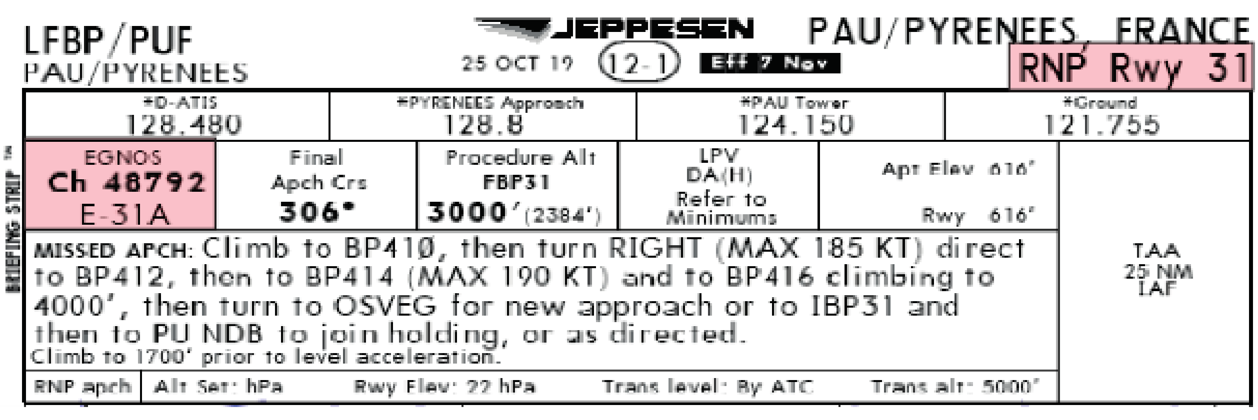
5.5 - Required navigation performance approach (RNP APCH)

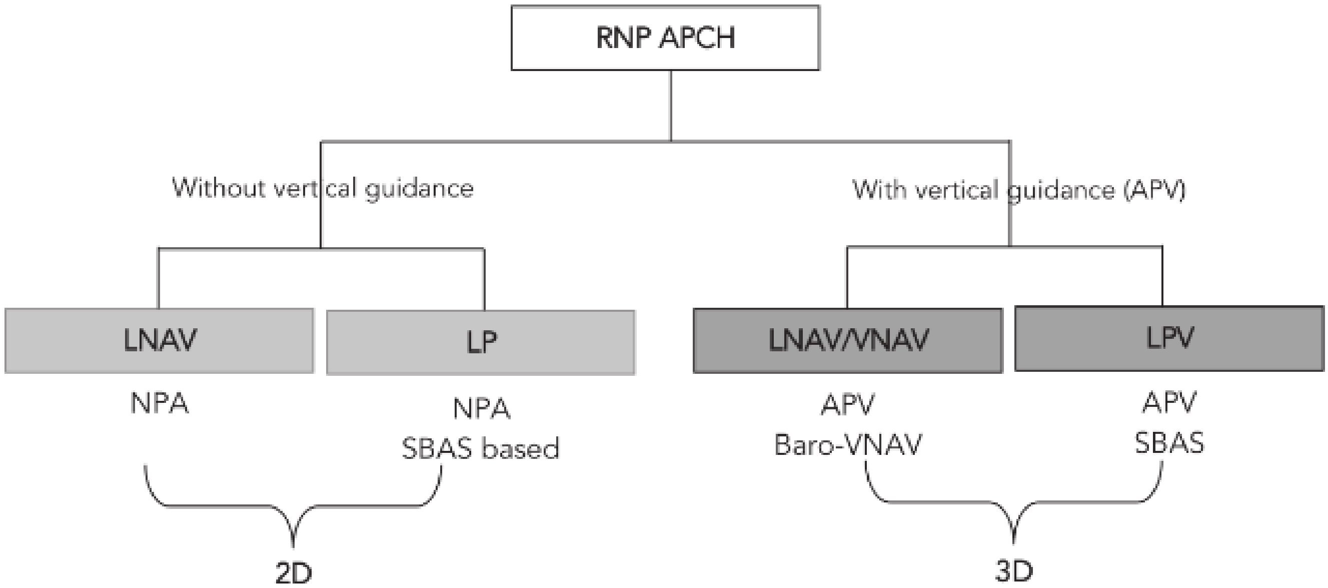
This is the terminology used in the ICAO PBN Manual to describe four approach types shown on the figure below.

The main difference between the various types of RNP APCH operations is the provision of vertical guidance.
All RNP APCH operations are based on the use of basic GNSS.
However, if local safety assessment requires it, conventional navigation aids (VOR, DME, NDB) can be maintained locally.
At present (2019-2020) we can find procedures published with the RNAV (GNSS) RWY xx title.
RNAV is a generic name for any kind of approach designed to be flown using the onboard navigation system using waypoints to describe the path to be flown instead of distances, altitudes and radials from ground-based navigation aids.
RNP APCH navigation specification is synonym of RNAV approach.



Gradually, charts are being renamed RNP.





NPA:	Non Precision Approach
SBAS:	Satellite Based Augmentation System
APV:	Approach with Vertical guidance
Baro-VNAV:	Barometric vertical navigation
APV Baro-VNAV:	RNP APCH with LNAV/VNAV minima
APV SBAS:	RNP APCH with LPV minima
LPV:	Localizer Performance with vertical guidance

Table 1: RNP APCH terminology as per PBN Manual - ICAO State Letter [2]

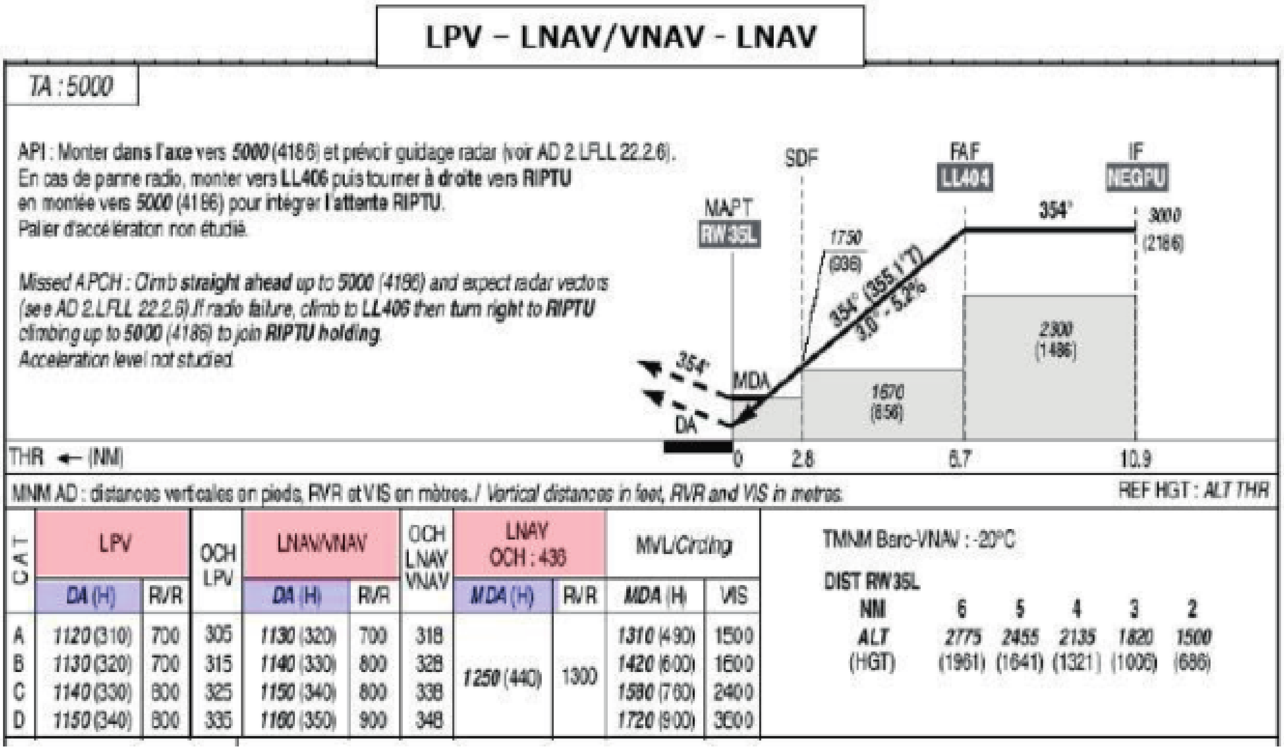
PANS-OPS Terminology	PBN Terminology	Chart Minima	Minimum Sensor
NPA	RNP APCH down to	LNAV (MDA)	Basic GNSS[2]
APV Baro-VNAV	RNP APCH down to	LNAV/VNAV (DA)	Basic GNSS + Baro-VNAV
-No criteria available	RNP APCH down to	LP (MDA)	SBAS
APV SBAS	RNP APCH down to	LPV (DA)	SBAS

LP minima are added where terrain or obstacles do not permit the publication of LPV with vertical guidance minima. Lateral sensitivity increases when the aircraft gets closer to the runway (or a point in space for helicopters).

Aircraft approved for RNP APCH can follow these applications only if they are retrievable by procedure name from the onboard navigation database and conform to the charted procedure.
The RNAV system can be used for the approach phase of flight, provided RNAV approach procedures are designed and published. RNAV approaches are described by a series of waypoints, legs, altitude and speed constraints published and stored in the onboard navigation database (ICAO EUR DOC 025).

RNP APCH with LNAV minima (MDA) is a non-precision instrument approach that provides lateral guidance only (two-dimension system (2D) management of lateral and longitudinal deviations).

OPS 1.430 (d) (2) specifies that all non-precision instrument approaches must be based on the Continuous Descent Final Approach technique.



CAT	LPV DA (H)	LPV RVR	OCH LPV	LNAV/VNAV DA (H)	LNAV/VNAV RVR	OCH LNAV VNAV	LNAV OCH : 438 MDA (H)	LNAV RVR	MVL/Circling MDA (H)	MVL/Circling VIS
A	1120 (310)	700	305	1130 (320)	700	318	1250 (440)	1300	1310 (490)	1500
B	1130 (320)	700	315	1140 (330)	800	328			1420 (600)	1600
C	1140 (330)	800	325	1150 (340)	800	338			1580 (760)	2400
D	1150 (340)	800	335	1160 (350)	900	348			1720 (900)	3600

NM	6	5	4	3	2
ALT	2775	2455	2135	1820	1500
(HGT)	(1961)	(1641)	(1321)	(1006)	(686)

RNP APCH with LNAV/VNAV minima (DA) are instrument approaches used for lateral guidance but also vertical guidance hence the 3D operation designation.

Note:
Minima also depend on the airport runway classification (ICAO Annex 14).
The following principle applies according to ICAO PANS-OPS:
For RNP APCH to LNAV/VNAV or LPV to a non-precision approach runway the DH shall not be below 300 ft.
For RNP APCH to LNAV/VNAV or LPV to a precision approach runway a DH below 300 ft is permitted.

APV SBAS CAT I with minimum DH 200 ft (precision approach)

Applicable airport operational minima can be found in UE OPS 1.430 Annex 1.

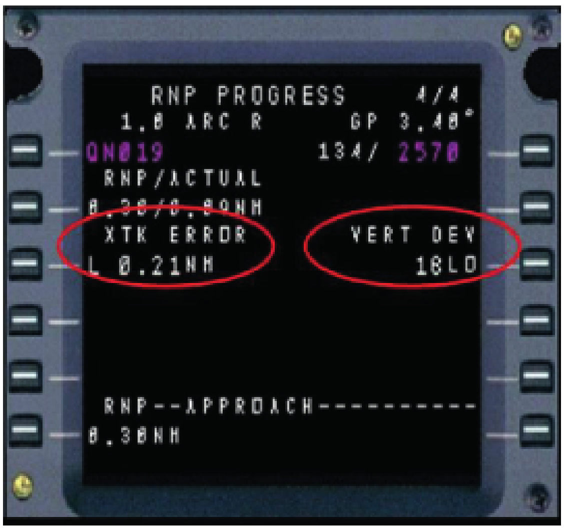
Like LNAV/VNAV, LPV is associated to a vertical guidance down to a decision height (DA) and it is also a 3D operation.

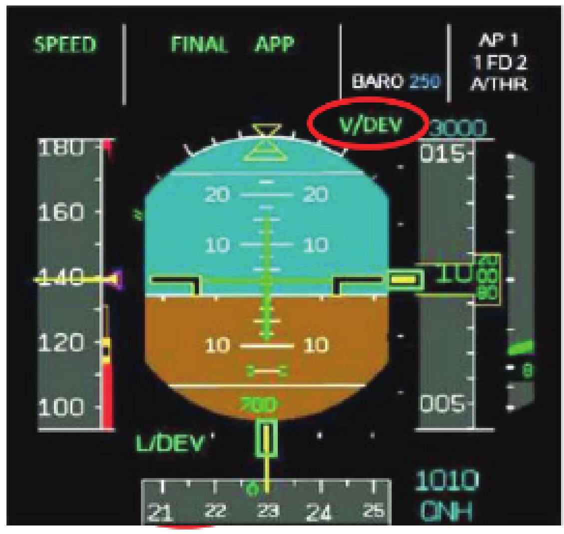
Vertical guidance:

RNP APCH benefit from lateral guidance based on the GNSS and vertical guidance based either on barometric altitude (APV Baro VNAV) provided by the CADC for example, or on SBAS information (APV SBAS = LPV).
Vertical guidance must be approved.

For Baro VNAV approach operations, the following elements are required:

- An area navigation system that can determine the distance to a waypoint that is the origin of the vertical flight path;
- The angle of the vertical path from the origin point (normally the runway threshold) coded in the navigation database;
- A barometric system with sufficient precision;
- A flight guidance system that can provide vertical guidance;
- Control and monitoring displays.

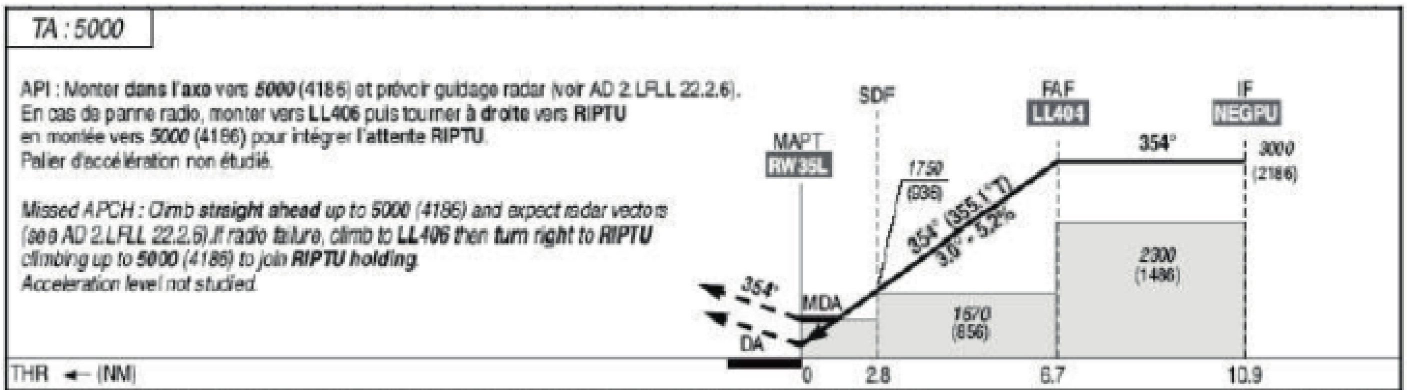




Limits of the Baro VNAV system:

1) Effect of non standard temperature

In the flight instrument course we saw that depending on the temperature (ISA) the altimeter can overestimate or underestimate the altitude of the aircraft.
When there is a negative difference from the standard temperature the profile of descent will differs from the published profile, with consequences on the obstacle clearance margins (MOC = Minimum Obstacles Clearance).
In ISA-10° an altimeter overestimates the altitude by 4%.
Thus, when using Baro-VNAV guidance, pilots must check temperature limitations on published approach charts which can result in approach restrictions (see chart excerpt from Lyon Saint Exupéry) if the aircraft system does not include an automatic barometric correction.



MNM AD : distances verticales en pieds, RVR et VIS en mètres. / *Vertical distances in feet, RVR and VIS in metres.* REF HGT : *ALT THR*

CAT	LPV		OCH LPV	LNAV/VNAV		OCH LNAV VNAV	LNAV OCH : 436		MVL/*Circling*	
	DA (H)	RVR		*DA* (H)	RVR		*MDA* (H)	RVR	*MDA* (H)	VIS
A	*1120* (310)	700	306	*1130* (320)	700	318	*1250* (440)	1300	*1310* (490)	1500
B	*1130* (320)	700	315	*1140* (330)	800	328			*1420* (600)	1600
C	*1140* (330)	800	325	*1150* (340)	800	338			*1580* (760)	2400
D	*1150* (340)	800	335	*1160* (350)	900	348			*1720* (900)	3600

TMNM Baro-VNAV : -20°C

DIST RW35L

NM	6	5	4	3	2
ALT	*2775*	*2455*	*2135*	*1820*	*1500*
(HGT)	(1961)	(1641)	(1321)	(1006)	(686)

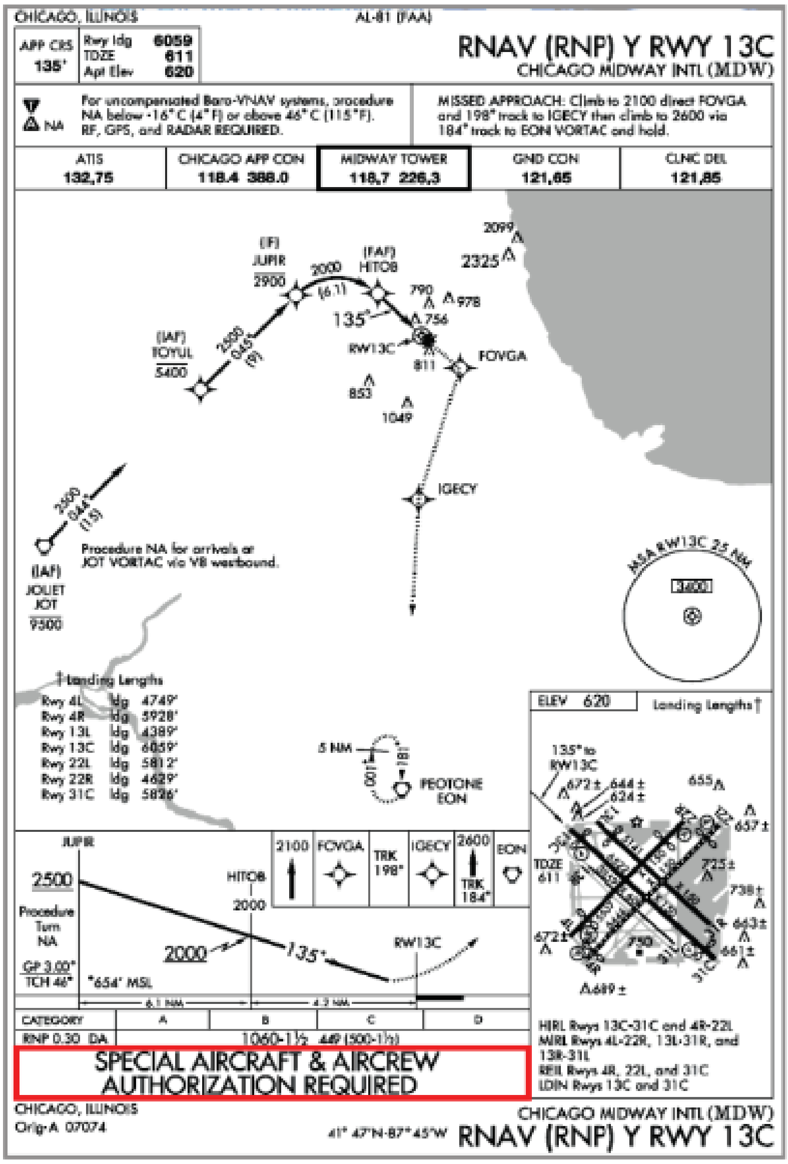
2) Altimeter setting

An erroneous barometric altimeter setting will be a source of error in the true altitude of the aircraft.
But this is not only true for PBN, and approach procedures provide a crosscheck among pilots to minimize the risk of error.
Furthermore, the translation of QNH in hPa is rounded to its lower value (a QNH of 1017,9 will show as 1017).
This allows an acceptably low margin of error (1 hPa = 27 ft).

Final approach segment:

RNP APCH with LPV minima is characterized in the navigation database by an FAS data block (Final Approach Segment).
The FAS data block contains lateral and vertical parameters that define the approach.
Those parameters are computed, validated, and published by the Air Traffic Services.
Additionally, each FAS data block ends with a data validation computer test (CRC – not necessary to know this technique).
Thus, the integrity is guaranteed when the equipment using the data succeeds in passing the data test.

5.6 - Required navigation performance authorization required approach (RNP AR APCH)



RNP AR APCH requires autorization.
AR means Authorization Required.
This is reminded on the approach charts (see chart on the left).

The RNP AR APCH allows:

- To improve safety and facilitate access to mountainous airports.
- To increase access to congested airspace.
- To avoid conflicts between nearby airports (in this case Chicago Midway and Chicago O'Hare).
- Stabilized approaches on the runway axis at airports where conventional approaches cannot be aligned on the runway.

By reducing the margins for lateral and vertical obstacle clearance and by allowing curved approaches by use of Radius Fix when and where necessary.

Note:
It is the operator, not the pilot, who requests the authorization by demonstrating that aircraft and crew can follow the procedure.

5.7 Advanced required navigation performance (A-RNP)

One of the main advantages offered by Advanced RNP is the increased efficiency in the use of airspace (for aircraft, air traffic, or the environment).
This concept allows the positioning of ATS routes, SIDs, STARs where it is most convenient. Predictable turn performance inherent to Advanced RNP, with the use of RF in terminal operations or FRT en-route, also allows for better adherence to predicted turn paths.

It integrates RNAV 5, RNAV 2, RNAV 1, RNP 2, RNP 1 and RNP APCH navigation specifications.

Intentionally left blank

08

ABBREVIATIONS
MEANING

Abbreviations meaning

- A/T Auto Throttle
- AC Alternative Current
- ADC Air Data Computer
- ADF Airborne Direction Finder
- ATC Air Traffic Control
- ATIS Automatic Terminal Information Service
- AWR Airborne Weather Radar
- BFO Beat Frequency Oscillator
- BRNAV Basic RNAV
- CAT Clear Air Turbulence
- CDI Course Deviation Indicator
- CDU Control and Display Unit
- CVOR Conventional VOR
- DDM Difference Depth of Modulation
- DF Direction Finding
- DME Distance Measurement Equipment
- DR Dead Reckoning
- DVOR Doppler VOR
- EADI Electronic Attitude Director Indicator
- EFIS Electronic Flight Instrument System
- EHF Extremely High Frequency
- ETO Estimated Time Overhead
- FCC Flight Control Computer
- FMA Flight Mode Annunciator
- FMC Flight Management Computer
- FMS Flight Management System
- GNSS Global Navigation Satellite System
- GPS Global positioning System
- GS Glide Slope
- HF High Frequency
- HSI Horizontal Situation Indicator
- ICAO International Civil Aviation Organization
- ILS Instrument Landing System
- INS Inertial Navigation System
- IRS Inertial Reference System
- ITU International Telecommunication Union
- LF Low Frequency
- LLZ Localizer
- LNAV Lateral Navigation
- MF Medium Frequency
- MLM Maximum Landing Mass
- MLS Microwave Landing System
- MTI Moving Target Indicator
- MTOM Maximum Take-Off Mass
- MZFM Maximum Zero Fuel Mass
- NDB Non Directional Beacon
- OBS Omni Bearing Selector
- PFD Primary Flight Display

This manual is part of the ATPL (A) collection of course books produced by ENAC and the MERMOZ Institute.

010 :	Air Law
021 volume I :	Airframes and Systems
021 volume II :	Powerplant
021 volume III :	Electrics – Protection and Detection Systems
022 :	Instrumentation
031/032 :	Mass and Balance / Performance (Aeroplane)
033 :	Flight Planning and Monitoring
040 :	Human Performance
050 :	Meteorology
061 :	General Navigation
062 :	Radionavigation
070 :	Operational Procedures
081 :	Principles of Flight (Aeroplane)
090 :	Communications

This collection of manuals complies with the ATPL (A) program and meets all Learning Objectives published by EASA *.

The complete set of manuals has been conceived under the joint direction of the CTKI (Chief Theoretical Knowledge Instructor) of the ENAC and the HT (Head of Training) of the MERMOZ Institute.

The following persons have contributed to the collective works cited above:
Agnès BELLAMY, Alain BELLIARD, Fanny BENAIM, Michael BENHAMED, Christian BEZANGER, Julien BLOYET, Jacques BOURDET, Denis CHAMBELIN, Yanick CHANTAL, Nicolas CHAUMEL, Eric COSTAMAGNO, Gérard DANIEL, Gérald DAVERDIN, Dominique De GEBHARDT, Daniel DUBUIS, Christian DUFOUR, Laurent FOURNIER, Muriel GIZARDIN, Jean Yves GRAU, Matthieu GUALINO, Anne HENRIC, Philippe JEANSON, Harald JOSET, Alexandre LE GOFF, Didier LABYT, Joël LAITSELART, Henri MAROTTE, Nadia MAS, Nadine MATTON, Laurence MORIN, Alain NGUYEN, Catherine ORIA, Jean-Luc PAGESY, Sylvie PANIAGUA, Jean-François PETIT, Serge POUPARD, Jean-Henry ROBRES, Alain ROUGE, Yves ROUILLARD, Régis SEGUIN, Philippe SESTI, Alain TAGLIAVINI, Bruno TALAVERA, Vincent TREIL, Patrick VACHER, Bruno VIDEAU, Quentin VIGNOLLET.

These persons have reread the texts: Fanny BENAIM, Michael BENHAMED, Jacques BOURDET, Jean-Pierre CELTON, Yannick CHANTAL, Eric COSTAMAGNO, Laurent FOURNIER, Didier LABYT, François MARQUINEZ, Nadia MAS, Nadine MATTON, Sylvie PANIAGUA, Kunnawuth PIN, Jean-Henry ROBRES, Yves ROUILLARD.

*EASA : European Aviation Safety Agency

Ecole Nationale de l'Aviation Civile – ENAC® - French Civil Aviation University FR-ATO-0056
7, avenue Edouard Belin BP 54005 31055 TOULOUSE CEDEX 4 – France
www.enac.fr

Institut Aéronautique Jean MERMOZ – Institut MERMOZ® FR-ATO-0020
43, avenue Robert Schuman 94150 RUNGIS – FRANCE
www.institut-mermoz.com

Book covers designed by Jérôme ESPENAN – ENAC – photography credits AIRBUS

ISBN : 978-2-86248-244-6
Legal deposit : 1st Quarter 2021