

~~OFFICIAL~~



AP3456
The Central Flying School (CFS)
Manual of Flying

Volume 7 – Avionics

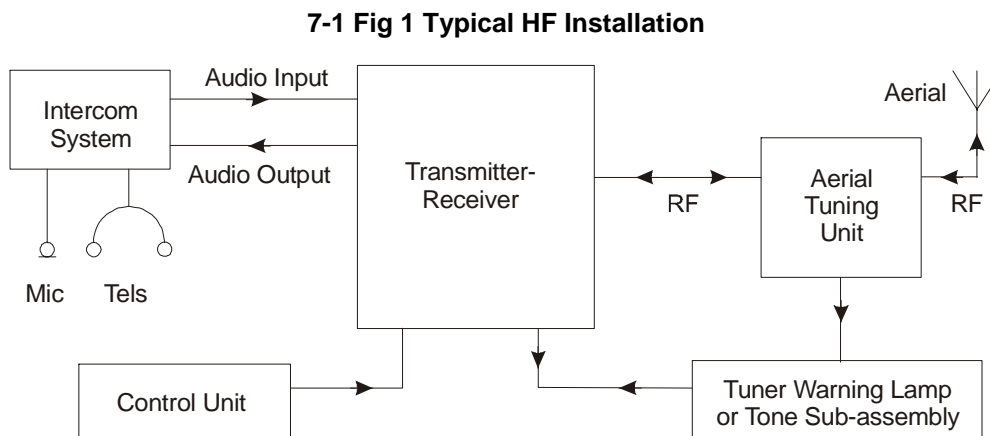
CHAPTER 1 - LF, MF AND HF COMMUNICATIONS

Introduction

1. This chapter describes the installation and operation of a typical HF communications system and briefly mentions the uses of MF and LF systems.

Typical HF Installation

2. Fig 1 illustrates, in block schematic form, a typical aircraft HF installation.



3. **Control Unit.** The control unit enables selection of the frequency and mode of operation of the equipment. Typically an HF set is able to operate in the 2 – 30 MHz range, usually in the single sideband (SSB) mode but may be able to operate in double sideband (DSB) and CW modes. Channel spacing is in the order of 100 Hz intervals (giving 280,000 channels) with some older equipments more widely spaced. The power output of an HF set is between 100 and 400 W in the SSB and 125 to 250 W in the DSB and CW modes. The control unit will normally be installed in the cockpit although in some multi-crew aircraft it may be at a rear station.

4. **Aerial Tuning Unit.** The aerial tuning unit, or coupler antenna, matches the transmitter and antenna to ensure efficient radiation of maximum power from the aerial. Aerial matching is normally automatic when the transmitter is keyed.

5. **Tuning Indications.** During the tuning cycle, coarse matching of transmitter and aerial takes place and, in some systems, this causes the tuner to resonate. Failure to resonate produces a 1 kHz audio warning tone which will continue to operate until the frequency has been reselected. Other systems include a warning light which extinguishes when tuning has been accomplished. Automatic fine tuning takes place during the first 0.5 seconds of each transmission.

Operation of HF Radio

6. **Mode Selection.** Some devices operate only on the upper sideband and have no other modes. However, on other equipments the following modes may be selectable:

- a. Upper Sideband (USB).
- b. Lower Sideband (LSB).

- c. Double Sideband (DSB).
- d. Continuous Wave (CW).
- e. Data Transmission (DATA) (eg RATT).

7. **Frequency Selection.** The required frequency is set by turning selector knobs until the desired frequency appears in the indicator windows. Channels may be preselectable.

8. **Radio Frequency (RF) Sensitivity.** The sensitivity control acts on the RF stage of a receiver. If a received signal is particularly strong it may saturate the RF stage and thus be distorted. Reducing the sensitivity reduces both the signal and the background noise. This allows strong signals to be received without distortion and with low levels of background noise. Conversely, the sensitivity needs to be increased when receiving weak signals, despite this increasing background noise. In the latter case, the clarity of weak signals can be improved by using various filters, such as bandwidth controls.

9. **Bandwidth.** The type of signal being transmitted will determine the bandwidth needed at the receiver. CW transmission may be received with bandwidths as low as 500 Hz whilst voice operation will generally need bandwidths of at least 3 kHz (low fidelity SSB). The bandwidth of data transmissions may vary considerably and may be as high as 100 kHz. Generally, the receiver bandwidth needs to be twice the highest message frequency. Within reason, increasing bandwidth increases clarity, decreasing it decreases interference.

10. **Preparation for Use.** To prepare the HF set for use, the mode is selected and the frequency is then set. If the required frequency was already set before switching on, it may be necessary to offset, then re-select the frequency to allow aerial tuning. The equipment cannot be used until the 1 kHz tone stops or the tuning light has extinguished, indicating that the equipment is ready for use. When aerial tuning is complete the RF sensitivity should be adjusted to an appropriate level.

11. **CW Operation.** Few modern aircraft require a Morse Code CW transmission facility but often make use of CW reception for which the CW mode (if available) should be selected.

12. **Radio Airborne Teletypewriter (RATT) Operation.** (See also Volume 7, Chapter 25). The RATT may transmit data via any selected channel on HF using the upper sideband.

13. **Skywave Disturbances.** Regardless of the efficiency of the equipment, reliable communications on HF are only likely to be achieved if the frequency is close to the optimum working frequency for the time/date/distance determined from radio propagation charts or the Flight Information Handbook.

14. **Safety.** When making HF transmissions, aircrew are to be aware that a radiation hazard exists to personnel and aircraft nearby, and relevant safety procedures must be observed.

The Uses of Medium and Low Frequencies

15. **Medium Frequency (MF).** Medium frequencies using ground waves are very reliable when used for reception of non-directional beacons and broadcast stations up to about 350 km. Sky wave results are unpredictable although occasional reception can be achieved many thousands of km distant. MF is susceptible to atmospheric and electrical interference.

16. **Low Frequency (LF).** LF is used for long-range transmission of digital data to aircraft fitted with teletype installations, and certain global navigation systems (eg LORAN C). Very Low Frequencies

(VLF) are ideal for communicating with submerged submarines. The LF and VLF bands may also be subject to atmospheric interference caused by electrical discharges from clouds and by interference from unsuppressed electrical apparatus.

CHAPTER 2- SATELLITE COMMUNICATIONS

Introduction

1. The continued increase in long distance communications traffic and the expected growth in worldwide digital transmissions have led to the development of high capacity satellite communications systems.
2. Satellite communication is by line of sight; the satellite acts as a relay between Earth-based communications stations. For military applications, a satellite link has the advantage of providing a reliable and secure communication system between remote theatres of defence interest without the necessity for multiple, en route ground relay stations.

Frequencies

3. Satellite communications systems mostly operate in the frequency range 1 GHz to 32 GHz. Military satellite systems use UHF (Ultra High Frequency; 300 MHz - 3 GHz), SHF (Super High Frequency; 3 GHz - 30 GHz) and EHF (Extra High Frequency; 30 GHz - 300 GHz). Below 1 GHz, cosmic noise is a restrictive factor. Above 15 GHz, cloud and heavy rain can cause signal attenuation, depolarization and increased noise. Signal strength is also reduced by low elevation angles of the Earth station antennae. However, 40 GHz transmissions are feasible, provided that antenna angles are above 10° to the local horizon, and the weather is clear.

Satellite Orbits

4. Depending on the Earth coverage requirements, satellite orbits may be circular or elliptical, with equatorial, polar or inclined planes, two examples of which are illustrated in Fig 1.

7-2 Fig 1 Circular and Elliptical Orbits

Fig 1a Circular Equatorial Orbit

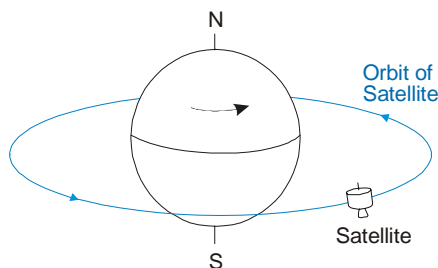
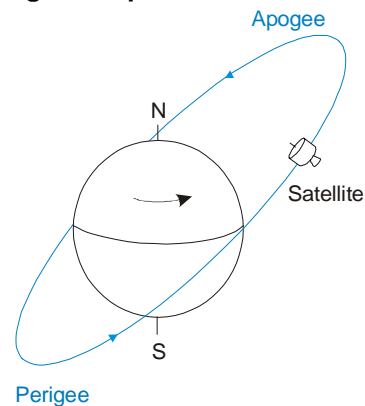


Fig 1b Elliptical Inclined Orbit

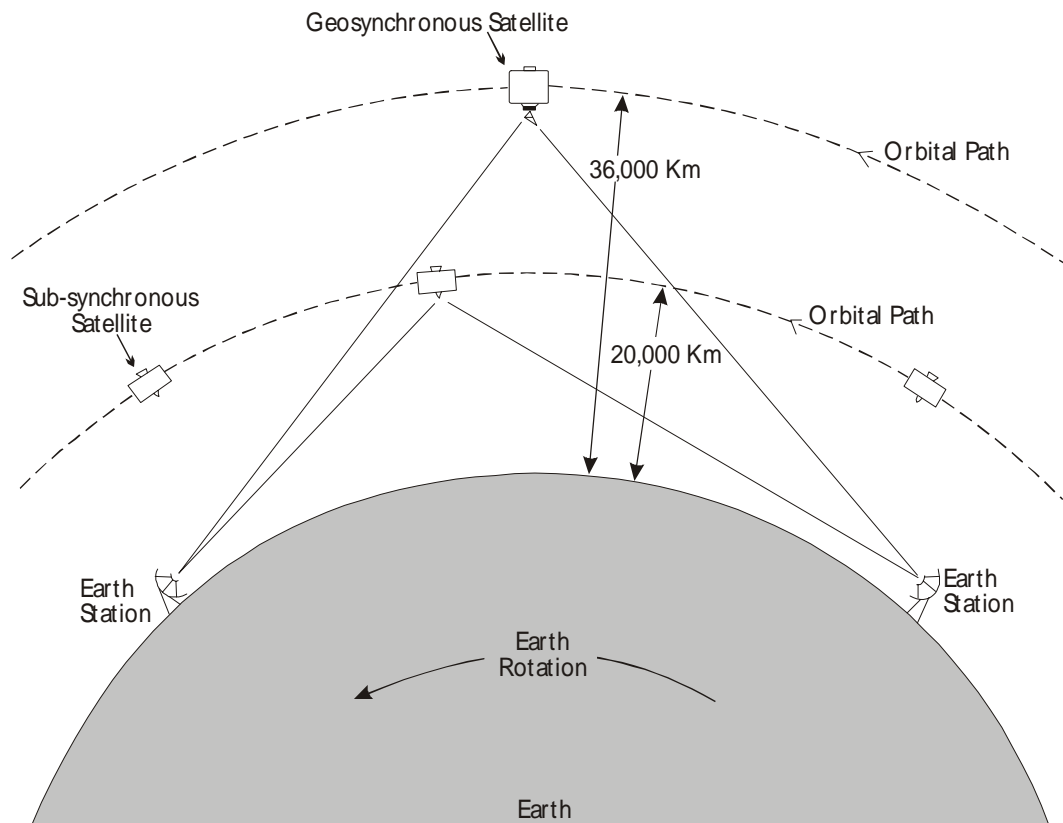


5. These orbital categories may be further classified by reference to the orbital period, as depicted in Fig 2:

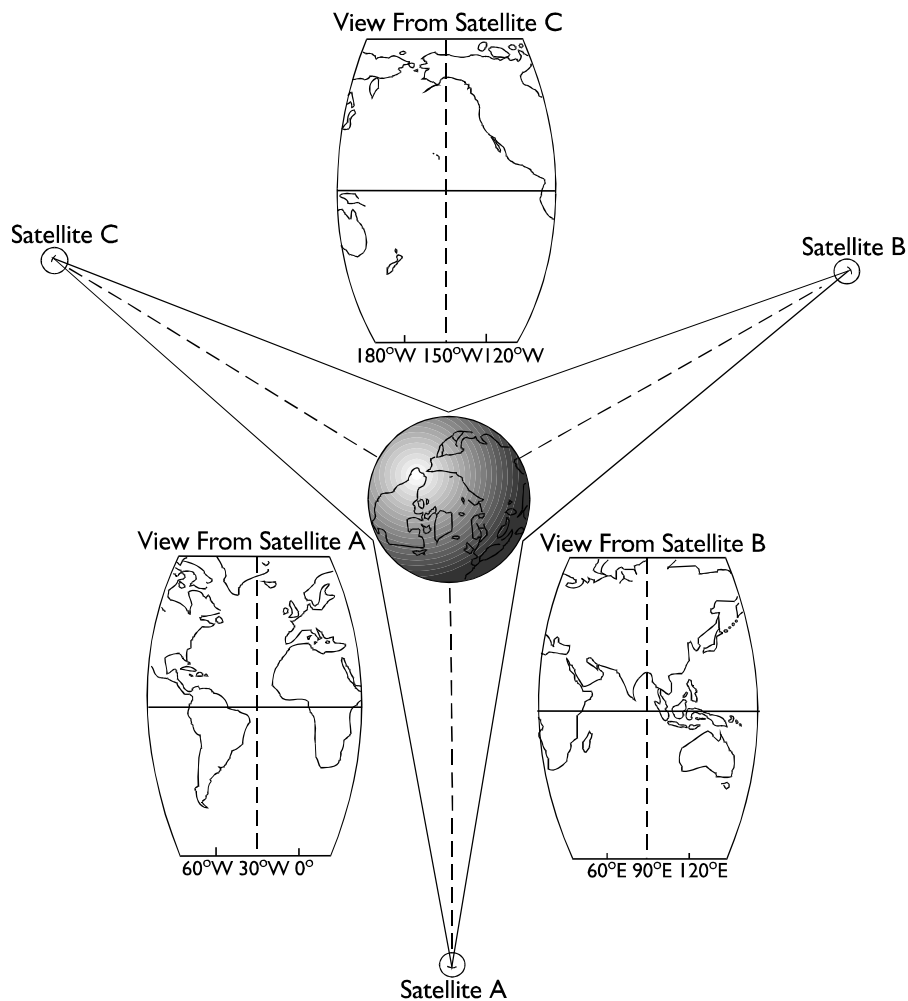
- a. **Sub-synchronous Orbit.** A satellite in a sub-synchronous orbit has an orbital period of up to 12 hours depending on its height (2,000 to 20,000 km). A communication system based on this type of satellite would employ a number of satellites, spaced apart. Low Earth Orbit (LEO) satellite systems use constellations of up to 66 satellites to give full Earth coverage. Earth stations communicate with each other via a satellite mutually in view.

b. **Geosynchronous Orbit.** A geosynchronous orbit is one in which a satellite stationed in an equatorial plane at a height of approximately 36,000 km has an orbital period of 24 hours; viewed from an observer on Earth the satellite appears to be stationary.

7-2 Fig 2 Sub-synchronous and Geosynchronous Orbits



6. The geosynchronous satellite configuration requires fewer satellites and launches than a typical LEO constellation, but more complex signal technology is required to provide a comparable service. Worldwide communications coverage, with the exception of the polar regions, can be achieved using three geosynchronous satellites 120° apart, as shown by Fig 3.

7-2 Fig 3 Worldwide Earth Coverage

Satellite Sub-systems

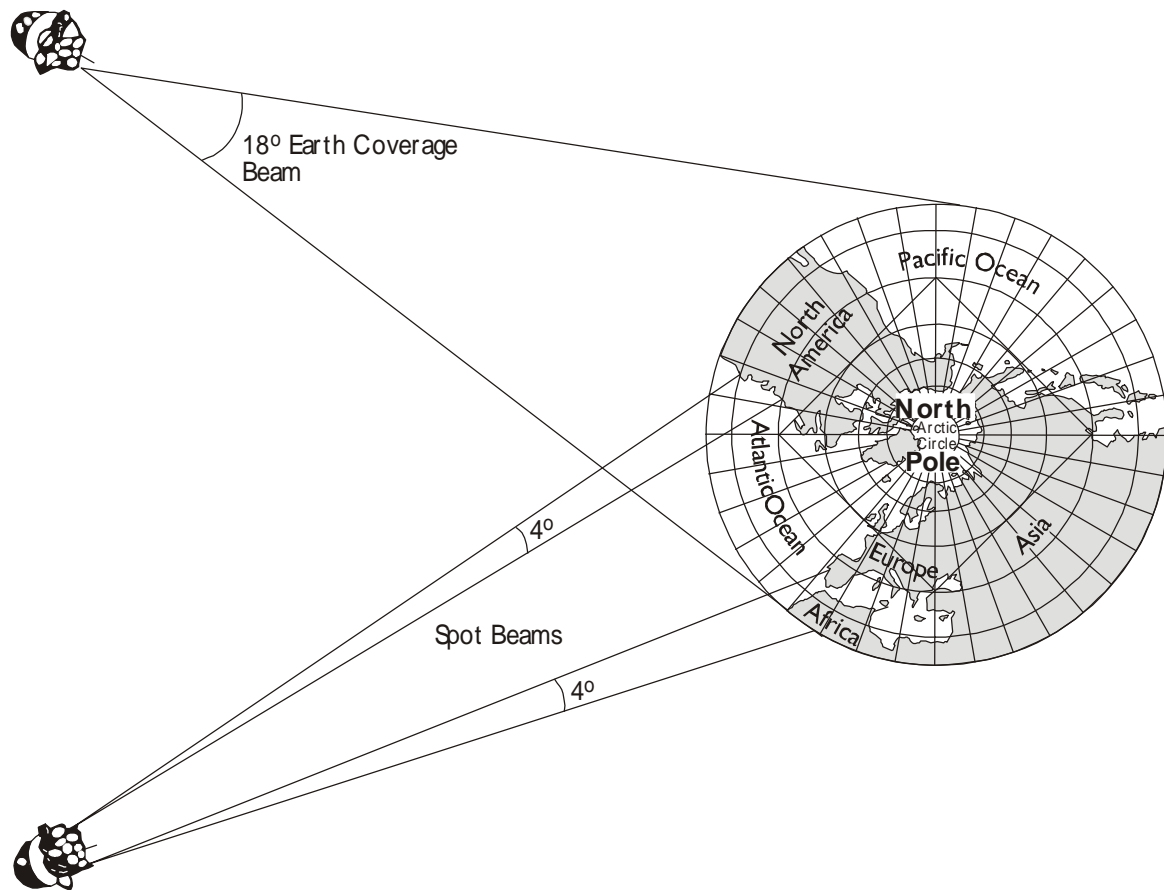
7. Communications satellites contain the following sub-systems:

- a. Transponders.
- b. Antennae.
- c. Telemetry and Command.
- d. Attitude and Orbit Control.
- e. Propulsion.
- f. Electrical Power.

8. **Transponders.** The communications sub-system contains a number of transponders, each capable of handling multiple signals simultaneously. A transponder receives the up-link signal in one frequency band, translates the signal to a down-link band, then amplifies and retransmits the signals within a specified bandwidth and at a set power level. In addition to up-link/down-link frequency separation, use is made of circular polarization to further isolate the transmitter from the receiver. The transmit signals have right-hand polarization; the receive signals have left-hand polarization.

9. **Antennae.** Satellites carry one or more communications antennae. Simple satellites may use only a single antenna to receive and another to transmit. More complex satellites use a mixture of widebeam and narrowbeam antennae. The widebeam, or Earth coverage, antenna enables the transponder to radiate its power over an 18° illuminating area, or 'footprint'. The narrow, or spot beam, antenna concentrates its power into a 4° beam thus focusing the available power into a restricted area (see Fig 4). Current communications satellites use digital technology to form multiple spot beams which are reflected from a single large antenna.

7-2 Fig 4 Satellite Coverage Beams



10. **Telemetry and Command.** Separate telemetry and command antennae are usually carried. The down-link transmissions contain engineering and equipment status information, and the up-link is used for commands to select equipment modes and to pass any information needed to maintain or change the satellite's orbit.

11. **Attitude and Orbit Control.** Communications satellites need to be maintained in the chosen orbit as follows:

- a. **Attitude.** The attitude of a satellite is defined by its aspect relative to Earth, in particular the direction in which its antennae are pointing. Satellites are either spin stabilized or three-axis stabilized. Spin stabilization is achieved by allowing the satellite to spin about its pitch axis whilst the antenna platform is maintained pointing towards Earth. Three-axis stabilization means that the satellite body, including the antenna platform, maintains a fixed attitude relative to Earth whilst the solar panels point towards the sun.
- b. **Orbit Control.** If a satellite is disturbed, or if it deviates from its designated path, its inclination and speed may need to be adjusted. Corrections are made using propulsion units to

manoeuvre the satellite about its North-South or East-West axis (relative to an observer on the surface of the Earth).

12. **Propulsion.** The propulsion system provides the motive power to drive the satellite to the selected orbit, and to maintain it there by commands sent from the ground control station. Most satellites use either hydrazine thrusters, low impulse ion thrusters, or a combination of both.

13. **Electrical Power.** Primary electrical power is derived from solar cells, either mounted on panels around the body of the satellite or in the form of winged arrays. Typical power levels are between 300 W and 16 kW. Stand-by batteries are provided for use when solar power is not available, eg during periods of solar eclipse.

Earth Segment

14. The Earth segment of a military satellite communications system comprises the communications ground stations and their control function. Ground stations may be fixed, semi-static deployable, or mobile tactical, and include terminals located on ships and in aircraft. The Earth segment has to ensure that:

- a. The satellite is not jeopardized by the action of any of its associated ground stations.
- b. Receiving ground stations can obtain, and maintain, links with the terminals having the lowest performance rating.

Military Applications

15. A military satellite communications system is designed to provide adaptable and flexible telecommunications to support operational requirements. It must also be compatible with existing communications networks.

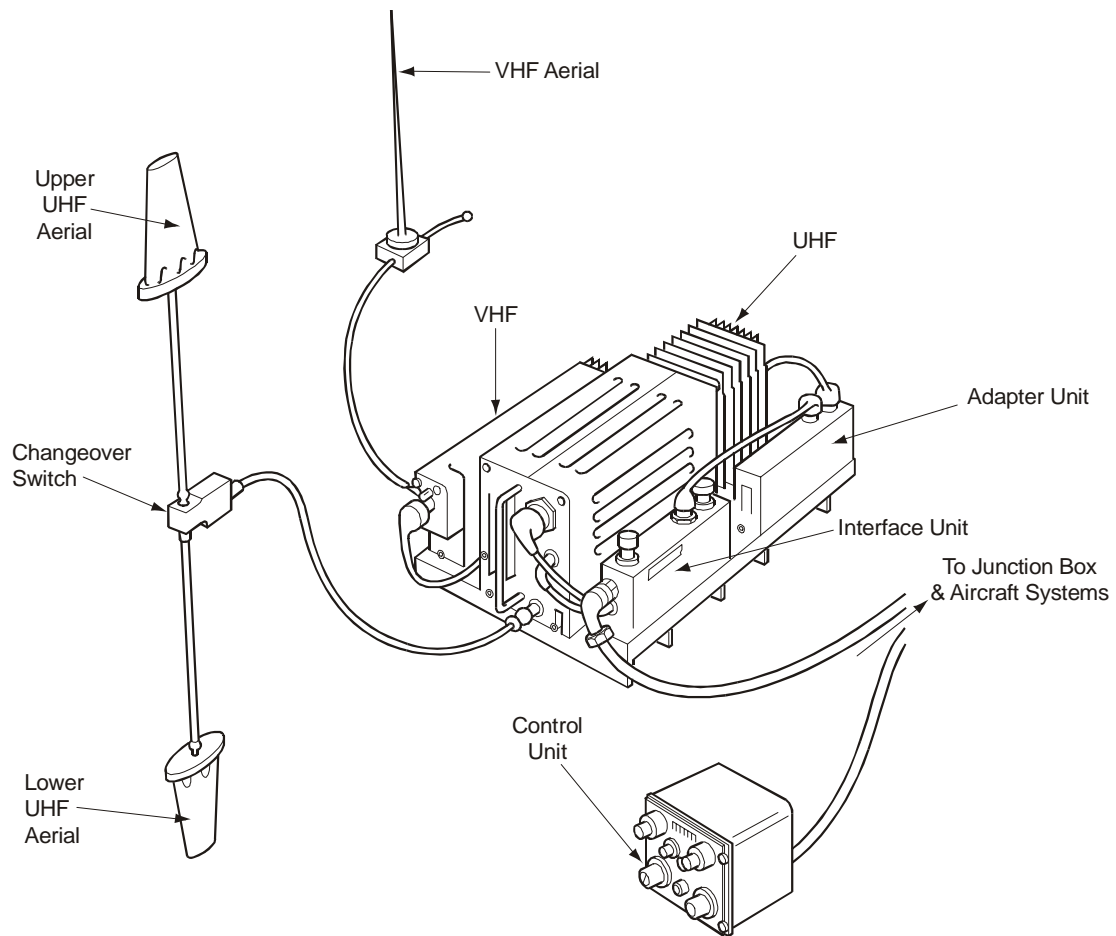
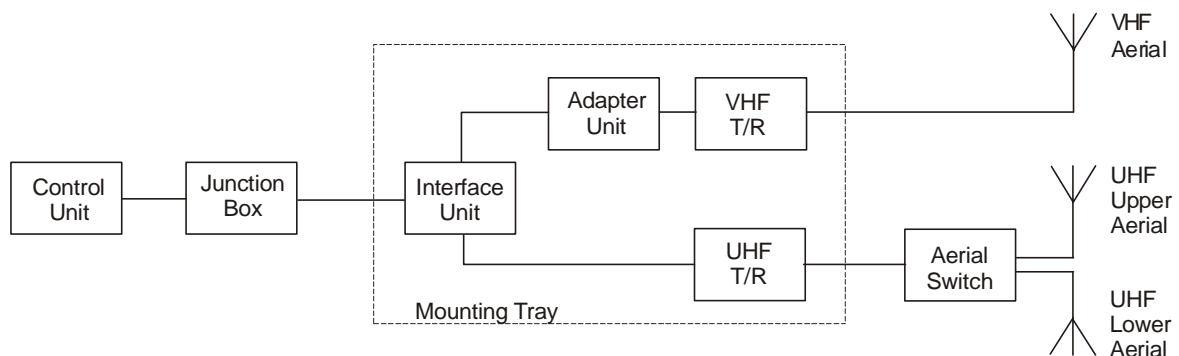
16. The United Kingdom uses the Skynet 4 series of satellites to provide secure, jam-resistant military communications. The satellites are operated under a Private Finance Initiative by Paradigm Services. This organisation will provide satellite communication services to UK military users by utilizing the existing Skynet 4 satellites and by providing new Skynet 5 satellites as the user need for bandwidth increases. From a master ground station, Paradigm Services link ships at sea, a static ground station in Cyprus and various deployed transportable terminals.

17. Skynet 4 is also used to provide the NATO Satellite Communications System, which forms part of the NATO Integrated Communications System (NICS), linking member governments and Commanders-in-Chief in each strategic theatre. The system comprises 2 NATO satellites positioned over the Atlantic Ocean illuminating fixed ground stations and transportable terminals located in various countries, ranging from Turkey to Canada. The master ground station is in Belgium.

CHAPTER 3 - V/UHF COMMUNICATIONS

Introduction

1. Short range communication with aircraft normally takes place on frequencies in the very high frequency (VHF) or ultra high frequency (UHF) bands at ranges of up to approximately 100 miles. The UHF frequency band is used exclusively for military aircraft control, while VHF is used for both civil and military control purposes.
2. The theoretical maximum range of V/UHF air-ground communications is the line-of-sight range which varies according to the altitude of the aircraft. In normal operations, the maximum range of V/UHF communications is approximately 100 nm, although this may be extended by the use of remote relay slave sites. The maximum range reduces at lower aircraft altitudes until, at ground level, ranges of a few miles are normal. Because transmitter power has little effect on range at V/UHF frequencies aircraft transmitter output power is typically 20 watts or less.
3. On modern equipment there are over 1300 VHF and 7000 UHF channels available at 25 kHz spacing. The control switches usually select and identify a channel by its frequency in MHz, although a channel number or letter may be used. In order to increase the number of channels available, 8.33 kHz spacing has been introduced.
4. Because V/UHF equipment may be used to communicate over transmission path lengths varying from a few hundred metres to around 100 nm, use is made of automatic gain control (AGC) to maintain a constant receiver output and obviate the need for continuous adjustment of the manual volume control.
5. The aerials for VHF and UHF equipments are quite small and can be placed almost anywhere on the aircraft. They may be built into the contours of the aircraft skin to reduce drag on high performance aircraft. An aircraft fuselage may block radiation from aerials in certain directions and to overcome this problem two sets of aerials may be mounted on the aircraft, one set on top and one set underneath. When two sets of aerials are installed the pilot is able to select which set to use by means of an aerial changeover switch (see Figs 1 and 2).

7-3 Fig 1 V/UHF Installation**7-3 Fig 2 Block Schematic Diagram of a Typical V/UHF Installation**

6. Within the equipment there are separate receiver channels which are tuned to the emergency frequencies 121.5 MHz, VHF or 243 MHz UHF. These channels, known as GUARD channels, may be monitored by superimposing the output from the GUARD receiver on to the frequency or pre-set channel selected (see para 8).

MAIN EQUIPMENT

V/UHF Installation

7. A V/UHF system is shown in Fig 1. The UHF and VHF transmitter/receiver units and interface and adaptor units are mounted together. A single control unit is used to control both transmitter/receivers. The adaptor unit is needed to convert serial data from the control unit to parallel data for the VHF. The adaptor unit also contains a 1 kHz oscillator to provide tone transmissions which, with the aid of a simple code, can be used to communicate information when speech is not possible, ie microphone failure, or, in tactical situations, where the use of speech is not advisable. The interface unit contains controls and links to allow the use of carbon and electro-magnetic microphones and the emergency intercom. In multi-seat aircraft where boom microphones are used the function of the interface unit may be carried out by the intercom system. A block schematic diagram of a typical V/UHF installation is shown in Fig 2.

Control Unit

8. A typical control unit incorporates controls and switches for the selection of function, mode and frequency.

- a. **Function.** The required function may be selected from transmit-receive (T-R), transmit-receive + Guard (T-R+G), transmit-receive + homing (T-R+H), or transmit receive + guard + homing (T-R+G+H). Homing is only available when a UHF frequency is selected.
- b. **Mode.** The mode of operation selected determines whether the equipment will operate on a manually selected frequency (M), or the frequency of a pre-set channel (P). Pre-set channels, or STUDS, may be identified by a letter or number. The VHF and UHF guard frequencies, 121.5 MHz and 243 MHz can be selected by placing the Mode Selector to Gv or Gu as appropriate.
- c. **Manual Frequency Selection.** Frequencies can be selected manually in increments of 100 MHz, 10 MHz, 1 MHz, 100 kHz and 25 kHz. The first five digits of the frequency are displayed. If 8.33 kHz spacing has been installed, frequencies are selected by channel numbers.

There are also controls for volume and lighting dimmer.

EMERGENCY UHF

Introduction

9. There are several types of Emergency UHF transceivers in use. Their prime purpose is to enable communications on the Emergency UHF frequency 243 MHz in the event of power or main radio failure. In addition, some installations allow operation on between 1 and 4 other preset channels.

Leading Particulars

10. The standby radio frequency range spans 238 MHz to 248 MHz. Apart from the guard frequency of 243 MHz, facilities allow operation on an additional channel or channels, depending on radio type, between 241 MHz and 245 MHz.

11. Although some installations include a separate $\frac{1}{2}$ wave aerial, when in use most transceivers are switched to the main UHF antennas. On some installations, where the emergency system can be utilized simultaneously with the main transceiver, the Emergency UHF is switched automatically to the non-selected aerial.

Power Supplies

12. The Emergency UHF requires 24V - 28V DC which is obtained from the aircraft power supply. In the event of its failure, a 24V battery continues to allow emergency operation.

CHAPTER 4- DATA LINK AND ENCRYPTION

Introduction

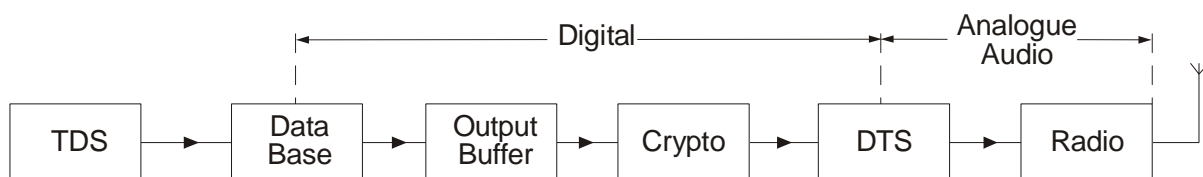
1. Data link is the term used to describe equipment which is used to pass tactical information between units in a manner that can be readily understood and absorbed into the Tactical Data Systems (TDS) of the linked units. Data link is the standard method used by NATO forces to co-ordinate actions, whether anti-air, anti-surface, or anti-submarine, where more than one unit is involved.

2. A Data Link Net consists of a Net Control Station (NCS) and one or more Participating Units (PU), which may be ground stations; reconnaissance, offensive or airborne warning and control system (AWACS) aircraft; or capital warships. The sensor information from each PU is co-ordinated to produce the best possible assessment of the tactical situation which is then made available to all PUs by the NCS.

Operation

3. In a simple system, data is passed between PUs by radio. Fig 1 shows a simplified diagram of a data link transmission system; the reception equipment is essentially the same except that the output buffer is replaced by an input buffer.

7-4 Fig 1 Data Link Transmission Chain – Schematic



4. All sensor inputs are passed to the unit's database by the individual sensor operators for display on the tactical plot and for storage in the structured message formats ready for transmission to other PUs, as and when required. Command and administration messages are also formatted and held ready for transmission. All messages are stored as digital groups. Although it is an operator function to decide which data will be put on the link, from then on the process is automatic.

5. Data is transmitted, typically, as 30-bit frames arranged in predetermined formats, each frame consisting of 24 information bits and 6 error detection and correction bits; two such frames form one link message.

6. Messages for transmission are passed to the output buffer ready to be encrypted by the crypto sub-system. The encryption is completed on-line automatically as part of the normal transmission process. The encrypted information passes to the Data Terminal Set (DTS) where it is converted from its digital form into analogue audio before being used to modulate the transmitter output. Data is passed at high speed, at either 2,250 or 1,364 bits per second, either sequentially, or simultaneously as 15 multiplexed tones (multiplexing is explained in Volume 7, Chapter 27).

7. Analogue audio received signals are converted to digital form and, if multiplexed, rearranged into their original format ready for decryption and subsequent passage via the input buffer to the tactical display; display is generally automatic although inputs may be manually filtered out if required.

8. The control of transmission is vested in the NCS which may interrogate each of the PUs in turn which then automatically release the information stored in their output buffers. Alternatively, in order to

maintain a greater degree of radio silence, PUs may broadcast short bursts of data as necessary to all other units on the net. The NCS promulgates details of the information required on the net together with operational details of, for example, frequencies and crypto to be used.

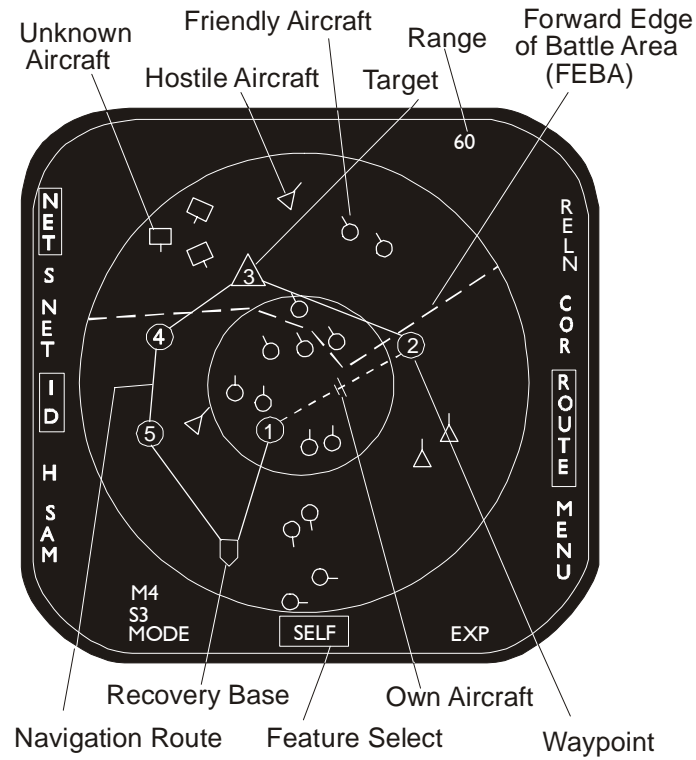
9. Effective data linking demands very precise navigation and any unit in a net must be able to determine its own position accurately and continuously both geographically and relative to other PUs in the net in terms of a common positional grid system. Errors in position can generate false targets on the net and may also result in real targets being missed. Typically, a PU will identify a target by its relative position as range and bearing. Depending on the sensor used there will be some degree of error in this data which may be aggravated when converted into a grid position. Another PU may hold the same target but with different system errors and so may relay a different position to the NCS. One of the principle tasks of the NCS, therefore, is to resolve these inconsistencies and relay the correct target position to the PUs.

Digital Data Link Systems

10. The following paragraphs describe the data link systems in most common military use by NATO forces:

- a. **Link 1.** Link 1 is a non-secure dedicated point-to-point telephone data link for exchange of the Recognized Air Picture (RAP).
- b. **Link 4.** Link 4 is a tactical fighter direction link. It is fitted in a variety of NATO aircraft including the E3D Sentry. This unencrypted link operates at high speed in the UHF band.
- c. **Link 11.** Link 11 is an automatic, high speed, encrypted maritime system using both UHF and HF frequencies. It is installed in the E3D Sentry and certain warships. It provides real-time air, surface and sub-surface track data, track management and EW information.
- d. **Link 14.** Link 14 is a secure link used to disseminate non-real-time data, particularly the Recognized Surface or Sub-surface Picture (RSP/RSSP). It can utilize a broad spectrum of channels covering HF to SHF frequency bands. It is capable of operating in broadcast mode.
- e. **Link 16/Joint Tactical Information Distribution System (JTIDS).** JTIDS and its associated Link 16 is a high speed, ECM resistant, secure system. It has high capacity voice and data capabilities and is used in co-ordination of air, ground and maritime operations. It can transfer data automatically between AWACS assets and command centres, to suitably equipped ships and aircraft and to beyond line of sight by the use of relays. Time Division Multiple Access (TDMA) techniques are employed to enable a large number of users to communicate amongst themselves using formatted or free text messages. To that end, an Interim JTIDS Message Standard, (IJMS) was introduced to allow early development of a message library. The sophisticated JTIDS system uses spread spectrum techniques combined with variable carrier frequencies making it difficult to detect or jam. Fig 2 represents a typical JTIDS display screen showing a variety of features which could be presented in a hostile environment.

7-4 Fig 2 Typical JTIDS Display



CHAPTER 5 - PRINCIPLES OF DIRECTION FINDING

Introduction

1. Radio compasses and ground based radio direction finding equipment use specially designed receiver aerials to measure the direction from which the received signal is coming. Airborne equipment operates in the LF/MF frequency range (usually between 200 kHz and 1800 kHz) and the system normally provides for the automatic determination and display of the relative bearing of the transmitter. The technique of radio direction finding is based on the reception characteristics of a loop aerial.

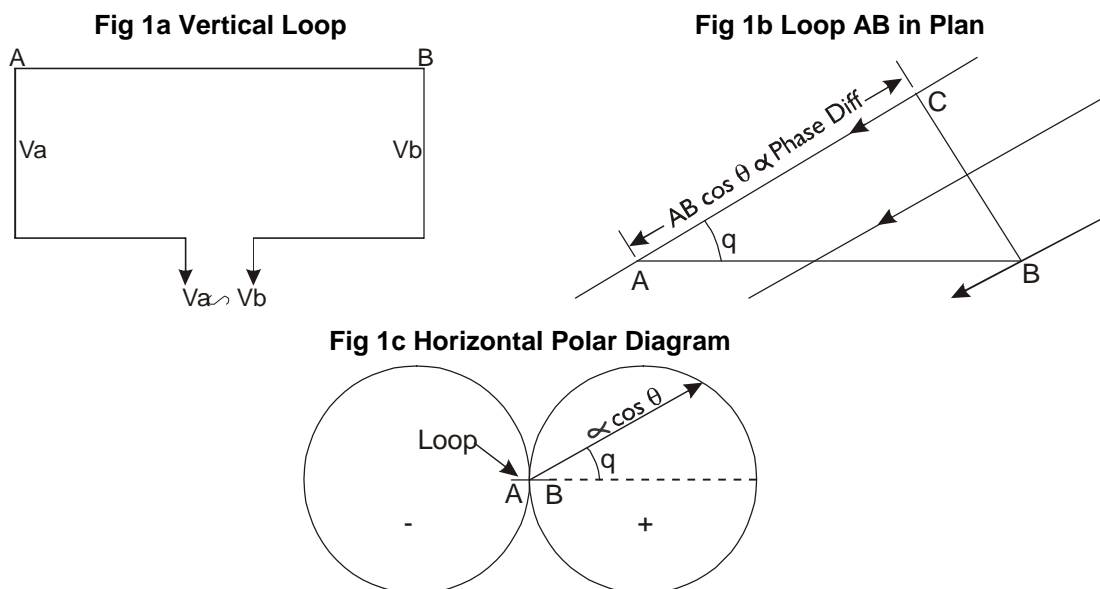
Loop Aerial Theory

2. Fig 1a shows a loop aerial, consisting of two vertical members, A and B, connected in the form of a loop by horizontal members. If a vertically polarized radio wave is incident upon the loop, it will induce voltages in the vertical members of the loop of value V_a and V_b .

3. Consider a wave incident at an angle θ to the plane of the loop (Fig 1b). Distance AB is insignificant compared with the range from the transmitter to the loop, so both A and B receive the same signal strength. However, as the signal travels a different distance to each, there is a phase difference at A and B given by $AB \cos \theta$.

4. Since AB is constant, the value of the resultant voltage in the loop is proportional to θ , giving the 'figure of eight' polar diagram for a loop aerial shown in Fig 1c. The plus and minus signs show the sign of $\cos \theta$, and hence the resultant voltage, V_r , in both lobes of the loop. The horizontal polar diagram has two sharply defined minima at $\theta = 90^\circ$ and $\theta = 270^\circ$, and two poorly defined maxima at $\theta = 0^\circ$ and $\theta = 180^\circ$.

7-5 Fig 1 Simple Vertical Loop



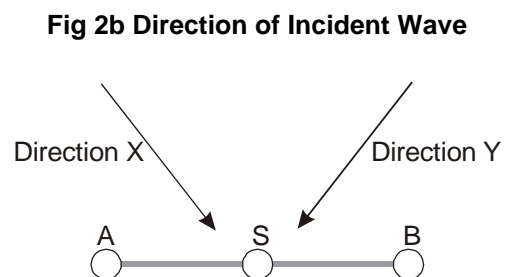
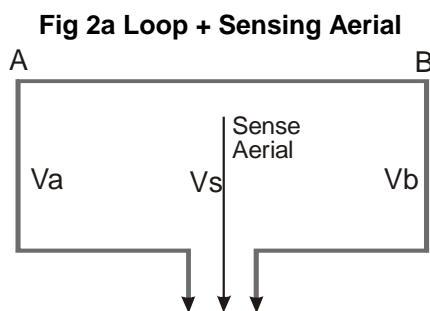
5. If a loop aerial which is receiving a wave from a transmitter is rotated, the resultant voltage in the loop will vary as θ varies. When $\theta = 90^\circ$ or 270° the resultant voltage is zero. When $\theta = 0^\circ$ or 180° the resultant voltage is a maximum. As the minima are the more sharply defined, these are used for direction finding. To take a manual loop bearing, the loop is rotated until a minimum signal (or 'null') is found. At this point, the transmitter must be on the line normal to the plane of the loop (subject to certain errors mentioned later). However, there is no indication on which side of the loop the transmitter is sited. The process of resolving this ambiguity is known as 'sensing'.

Sensing

6. If a vertical omni-directional aerial is placed midway between the two vertical members of the loop, the voltage induced in it by an incident wave will be midway in phase between the voltages induced in the vertical members of the loop (Fig 2a). It can be shown that the phase of the voltage in the sense aerial, V_s , is always 90° removed from the phase of the resultant voltage, V_r , in the loop.

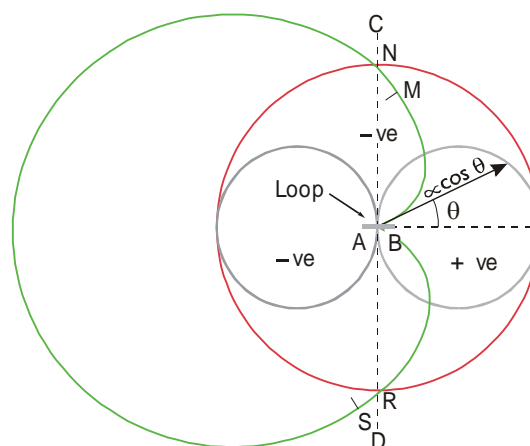
7. If the incident wave comes from the left of the normal to the plane of the loop (direction X in Fig 2b), V_s leads V_r by 90° , while if it comes from the right (direction Y), V_r leads V_s by 90° . By permanently incorporating suitable components in the sense aerial circuit, the phase of V_s can be retarded by 90° . If this is done, V_s will be in phase with V_r if the wave comes from the left of the normal and in antiphase if the wave comes from the right. The aerials are designed so that the value of V_s is equal to the maximum value of V_r .

7-5 Fig 2 Sensing



8. If V_s and V_r are combined, the polar diagram shown in Fig 3 will result. The figure of eight is the polar diagram for the loop alone, the circular polar diagram is for the sense aerial alone, and the heart-shaped or cardioid polar diagram is for the loop and sense aerials combined. (To the right of the normal to the loop, V_s and V_r are in antiphase, and cancel each other; to the left they are additive). In many installations the sense aerial is located away from the mid-point of the loop, but the principle is unchanged.

7-5 Fig 3 Generation of Cardioid Polar Diagram



Manual Sensing

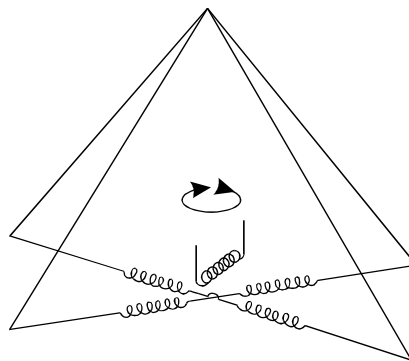
9. With the sense aerial switched out, the loop is rotated until a direction of minimum signal is found. The transmitter must now lie on the line CD in Fig 3 but can be either in direction C or in direction D. To resolve the ambiguity the sense aerial is switched in to combine with the loop to provide the cardioid polar diagram. A signal of medium strength will be received proportional to the distance from the loop to point R or point N on the polar diagram. Rotating the loop anti-clockwise by $10^\circ - 15^\circ$ changes the signal strength due to the eccentricity of the polar diagram. If the beacon is in direction D, point S on

the diagram will now be on the line CD so that a greater signal strength will be received. If the beacon had been in direction C, a weaker signal would have resulted as point M would now be on the line CD. Thus, if an anti-clockwise rotation produces a stronger signal, the assumed bearing of the transmitter is correct, and if a weaker signal is produced, the bearing is a reciprocal.

Automatic Direction Finding (ADF)

10. The loop aerial principle is used in practical radio compasses which automatically determine and display the transmitter direction. Instead of the single loop aerial, a 'Bellini-Tosi' aerial system is used. This consists of two loops at right angles to one another and each loop has a primary coil connected within the loop such that the coils cross at the centre (Fig 4).

7-5 Fig 4 Bellini-Tosi System



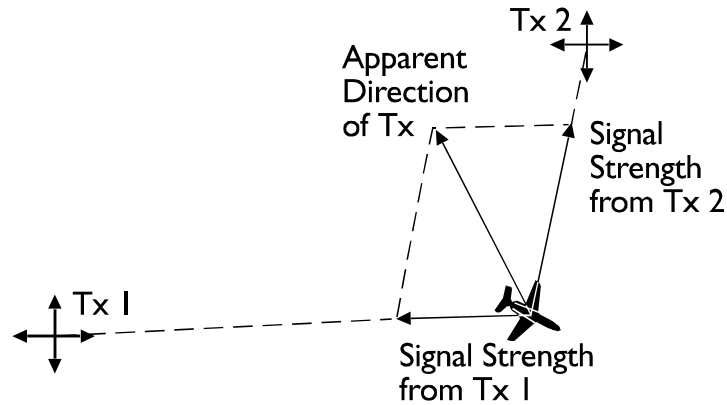
11. The alternating current flowing in each loop is proportional to the angle the incident wave makes with the loops. The field set up by the coil in each loop is proportional to the current flowing in its loop, thus the strength of each field depends on the angle of the incident wave on the respective loop. The resultant field produced by the combined effect of both coils will lie in the same direction as the incident wave. A coil can be rotated within this resultant field and the current induced in it will be dependent on the angle it makes with the field. If it is parallel to the field, maximum current (and therefore maximum signal) will be produced. If it is perpendicular, no signal is produced. The combination of two static perpendicular coils and a rotatable coil is known as a goniometer.

Direction Finding Errors

12. **Night Effect.** A loop aerial is designed to use vertically polarized waves for direction finding. If the incoming wave has a horizontal component of polarization it will induce a current in the horizontal members of the loop which will tend to cancel. However, if the loop is not at right angles to the travelling wave, the currents in the top and bottom will not cancel completely. A resultant current will flow in the loop which will degrade the nulls of the polar diagram thus making it impossible to take accurate bearings. By day MF transmissions propagate by surface waves and their plane of polarization does not change during propagation. However, at night, sky wave propagation becomes significant and, on being refracted at the ionosphere a vertically polarized wave becomes elliptically polarized, ie it has a changing horizontal component of polarization. On arriving back at the surface it is travelling downwards and has a horizontal component of polarization – the two features necessary to interfere with the directional properties of the loop. At night therefore, the nulls of the loop polar diagram tend to be wandering and indistinct. In order to reduce this effect, the horizontal members should be kept as close as possible, so that any phase difference in induced currents is minimized; in addition, the top of the loop should be screened.

13. **Synchronous Transmission.** If two or more beacons are transmitting on frequencies that are the same (or within the pass band of the receiver) and within effective range, the measured bearing will be in the direction of the resultant of the signal strength vectors (Fig 5). The act of deliberately creating a false bearing as part of electronic warfare is termed 'meaconing'.

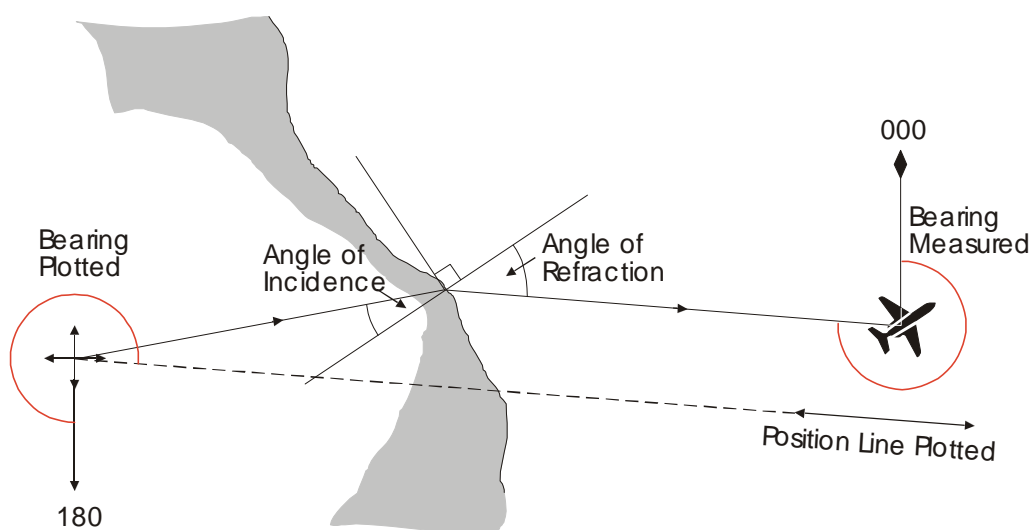
7-5 Fig 5 Synchronous Transmission



14. **Bank Error.** When an aircraft is banked, the aerial is tilted and some of the signal is detected by the horizontal elements of the aerial. This will produce an error in the ADF bearing until the wings are levelled again.

15. **Coastal Refraction.** The velocity of propagation of a surface wave is affected by the nature of the surface over which the wave is travelling. A radio wave travels slower over a surface of poor conductivity than over a surface of good conductivity. Thus, for example, radio waves travel faster over sea than over land. Fig 6 shows the case where an aircraft flying over the sea is receiving transmissions from an inland beacon. When the wave-front reaches the coast it will speed up but, due to the angle of the coast, elements of the wave-front to the north will reach the coast before elements to the south and hence they will speed up earlier. This causes a tilting of the wave-front at the coast away from the normal. The wave is thus refracted at the coast so that the angle of refraction is greater than the angle of incidence. The greater the angle of incidence the more the wave is refracted to a maximum of 3° or 4° . The error can be minimized by using coastal beacons or by taking bearings from an inland beacon when a line joining the DR position to the beacon crosses the coast at 90° .

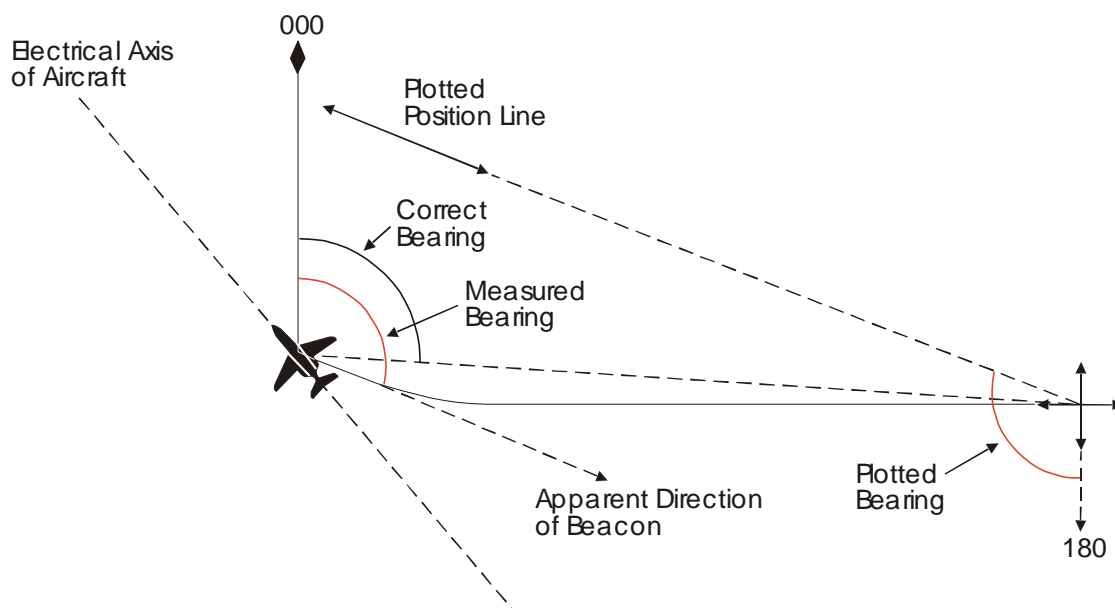
7-5 Fig 6 Coastal Refraction



16. **Quadrantal Error.** The electrical axis of an aircraft usually coincides with its fore-and-aft axis. Incoming radio waves cause re-radiation from metallic parts of the aircraft and the wave is refracted towards this axis (Fig 7). Since the error tends to be greatest on relative quadrantal points it is called

quadrantal error. Quadrantal error can be virtually eliminated by incorporating a quadrantal error corrector in the radio compass.

7-5 Fig 7 Quadrantal Error



CHAPTER 6 - PRINCIPLES OF VOR

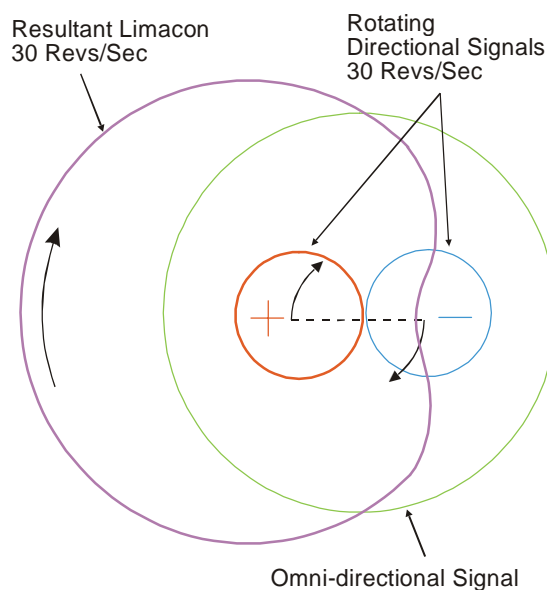
Introduction

1. Very High Frequency Omni-Directional Radio Range (VOR) comprising a ground beacon and an airborne installation, automatically and continuously provides the airborne operator with the magnetic bearing of the aircraft from the beacon. The system is used extensively as an en-route navigation and terminal approach aid.
2. VOR, operating in the frequency band 108.00 MHz to 117.95 MHz, has a great advantage over any MF system offering similar facilities, for its performance is not affected by static or night effect. However the line of sight properties of VHF transmissions limit the VOR coverage provided for low flying aircraft.
3. In 1960 ICAO adopted VOR as the international standard short-range navigation aid, and a large number of Service aircraft now carry the necessary receiver equipment enabling them to use the numerous beacons throughout the world.

Ground Beacon Operation

4. The principle of the VOR system is bearing by phase comparison: the aircraft's equipment derives the bearing from the phase difference between two 30 Hz modulations associated with the radio frequency transmission of the ground beacon.
5. Although the aerial systems of VOR beacons vary in construction, all transmit omni-directional and rotating directional signals. The polar diagrams of these transmissions, a circle and a figure of eight respectively, are shown in Fig 1 together with their resultant, a limaçon pattern. (A limaçon differs from a cardioid in that there is some signal strength present at the minimum).

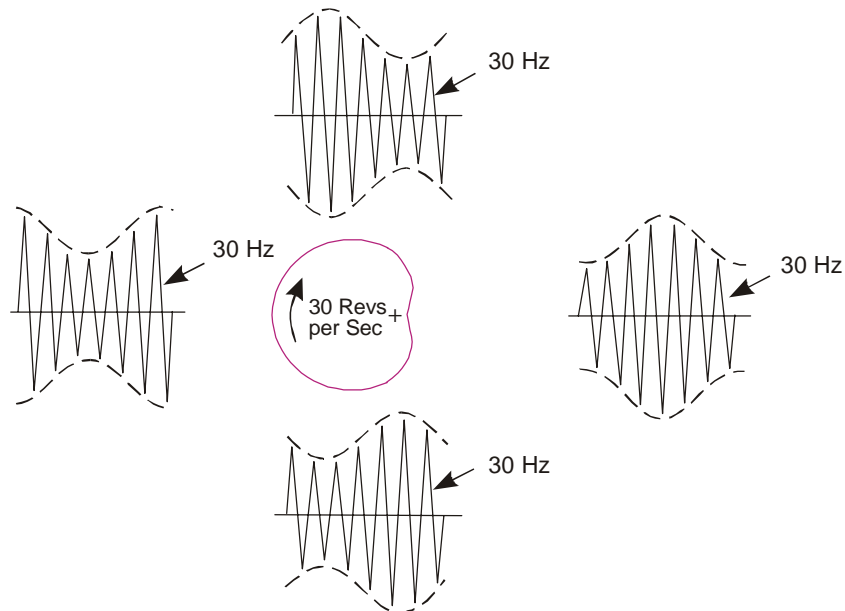
7-6 Fig 1 VOR Polar Diagram



6. The figure of eight diagram, and consequently the limaçon, is made to rotate at 30 revs/sec by rotating the horizontal dipole within a resonant cavity.

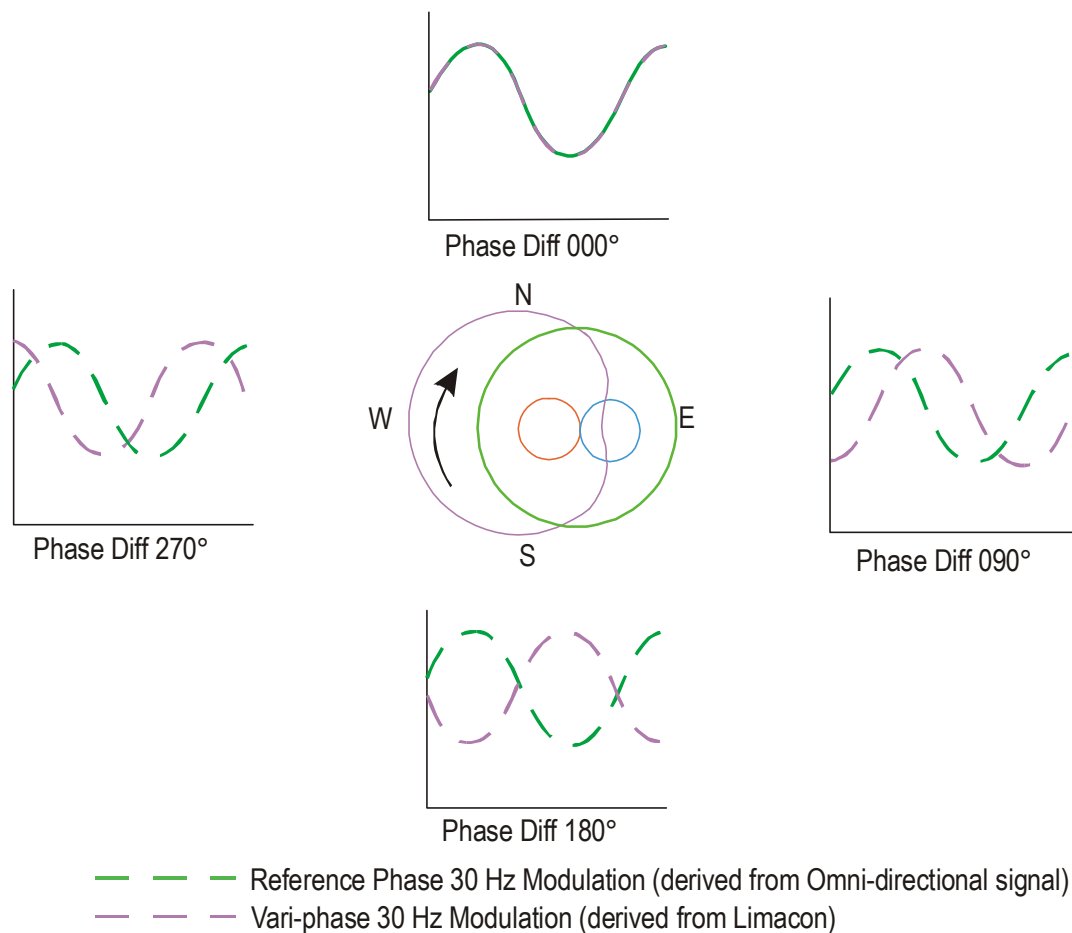
7. If the limaçon polar diagram starts rotating from the initial position shown in Fig 2, then the characteristics of the signal received at the four cardinal points during one revolution are as shown in Fig 2. It can be seen that the signal is amplitude modulated at 30 Hz, one full cycle for each complete rotation of the polar diagram, and that the phase of the modulation varies with the position of the receiver relative to the beacon. The modulation depth is maintained at 30 per cent.

7-6 Fig 2 Amplitude Modulation of Variable Signal



8. To compute the aircraft's bearing from this information it is necessary to have a reference from which the phase of the directional signal can be measured. This is provided by the omni-directional or reference signal which is frequency modulated at 30 Hz by the transmitter.

9. The two 30 Hz modulations are arranged to appear in phase to a receiver due North (Magnetic) of the beacon, hence any difference in phase is a direct measure of the bearing of the receiver from the beacon, see Fig 3.

7-6 Fig 3 Bearing by Phase Comparison

Airborne Installation

10. The heart of the VOR airborne installation is the Navigation Unit: this contains the circuits that detect and phase compare the variable and reference signals and provide outputs to the VOR displays.

11. The input arrives at the phase detector via two paths which are designed to extract the two 30 Hz components and present them in a suitable form for phase comparison within the detector. One path allows only the 30 Hz component of the AM signal to pass; the second path extracts the 30 Hz FM reference signal and converts it to a 30 Hz AM signal. The output from the phase detector is a signal proportional to the measured phase difference.

Frequency Allocation

12. The allocation of frequencies in the 108.00 MHz to 117.95 MHz band is as shown in the Flight Information Handbook, but can be summarized as follows:

- a. From 108.00 MHz to 111.85 MHz, those frequencies with an EVEN first decimal place (eg 109.20 MHz, 111.65 MHz) are VOR channels.
- b. All frequencies in the range 112.00 MHz to 117.95 MHz are VOR channels.
- c. From 108.10 MHz to 111.95 MHz, those frequencies with an ODD first decimal place (eg 109.10 MHz, 111.35 MHz) are ILS localizer channels.

13. The majority of VOR receivers also act as ILS localizer receivers, and the selection of a frequency automatically puts the equipment into the appropriate mode, VOR or ILS.

VOR Performance

14. As VOR is a VHF aid, the ranges that can be expected depend largely upon aircraft height. VHF range is approximately line of sight, plus 15% to 20% to allow for atmospheric refraction. The following table lists the approximate ranges to be expected.

| Height (ft) | Range (nm) |
|-------------|------------|
| 1,000 | 40 |
| 5,000 | 90 |
| 20,000 | 185 |
| 30,000 | 230 |

Note: There is a cone of confusion overhead the beacon, the radius of which is about 4 nm at 30,000 ft.

15. It is difficult to lay down specific accuracy figures for a complete VOR system, the final accuracy being the result of errors induced by the ground beacon, the airborne equipment and the crew's interpretation of the display. Beacon errors can be caused by misalignment with magnetic North (which is itself subject to diurnal and longer term change), siting effects (reflections of the signals from nearby objects or terrain) and minor inaccuracies within the manufacturing tolerances of the equipment. Airborne equipment error is considered to be the total error due to the receiver and display, ie the difference between the received and displayed bearing. Errors due to crew interpretation depend on the way in which the information is used, ie whether the display is used by the pilot to fly a radial or whether the bearing is plotted as a position line.

16. At the 95% probability level these error components are considered to be:

- a. Ground Beacon Error - $\pm 3^\circ$ (easily achieved in practice).
- b. Airborne Equipment Error - $\pm 3^\circ$.
- c. Pilot Error - $\pm 2.5^\circ$ (for radial tracking).

Since the errors are independent of each other the aggregate error can be taken as the root-mean-square of the individual components. This results in an error of $\pm 5^\circ$ in the case of an aircraft attempting to maintain a selected radial and $\pm 4^\circ$ in the case of a plotted position line.

Protected Range

17. To overcome the problem of mutual interference from adjacent beacons a 'protected range' is quoted in the form of two figures, eg 30/50. This describes a cylinder of radius 30 nm and height 50,000 ft within which the transmission is protected from interference from other beacons on the same frequency.

Types of VOR Beacons

18. Various prefix and suffix letters may be seen against VOR beacon information in En-Route Supplements. The most common are:

- a. **BVOR - Weather Broadcast VOR.** Weather Broadcast VORs, as their name suggests, broadcast weather information for selected airfields and areas.
- b. **DVOR - Doppler VOR Beacon.** With Doppler VOR beacons, the modulation type of the reference and variphase signals is reversed, ie FM instead of AM and vice versa. The aerial system is fixed and the horizontal dipole rotation is synthesized by feeding phase differences to many small dipoles in the form of a ring. Provided that the aircraft receives both signals there is no difference in the operation of the airborne equipment. The advantage of the system is that it does not have any moving parts and is less susceptible to siting effects. The majority of VORs worldwide are now Doppler VORs.
- c. **TVOR - Terminal VOR.** Terminal VORs are located at major terminal airfields and are low powered, hence limited range.
- d. **AFIS and ATIS.** AFIS (Aerodrome Flight Information Service) may be seen as a suffix to a TVOR; in this case the beacon will be broadcasting information about that particular airfield. ATIS (Automatic Terminal Information Service) can appear as a suffix to any VOR beacon, and the entry in the En-Route Supplement will indicate for which terminal airfield(s) the information is valid.
- e. **VORTAC.** A VORTAC installation consists of collocated VOR and TACAN beacons operating on the same frequency.
- f. **VOT - Test VOR Transmitter.** Test VORs are NOT for navigational use. They broadcast a fixed omni-directional 180° radial signal for testing VOR receivers.
- g. **Others.** Any combinations of the types mentioned in sub-paras a to e may be found, eg DBVORTAC, a weather broadcasting Doppler VOR beacon collocated with a TACAN operating on the same frequency.

VOR Displays

19. Many forms of VOR displays may be encountered. The type of display will be determined to an extent by whether the information is to be presented for pilot or navigator interpretation. In addition the display may be integrated into other equipments, eg HSI. Some of the more common displays are briefly outlined below but users should refer to the appropriate Aircrew Manual for detailed information and operating procedures.

20. **Radio Magnetic Indicator (RMI).** An RMI (Fig 4) simultaneously displays aircraft heading and the magnetic bearing to a radio navigation aid. The radio bearing is given by a pointer moving over a rotatable compass card which is driven by the aircraft's gyro-magnetic compass. An RMI is usually capable of accepting two radio inputs at a time. An RMI therefore usually has two separate pointers of different shape or colour or both, with changeover switches to control their functions. Before display, the VOR magnetic bearing has to be converted to a relative bearing by subtraction of the aircraft's magnetic heading. The relative bearing is then fed to the display, and as the compass card rotates with aircraft heading changes, the result is a display of magnetic bearing.

7-6 Fig 4 RMI Display

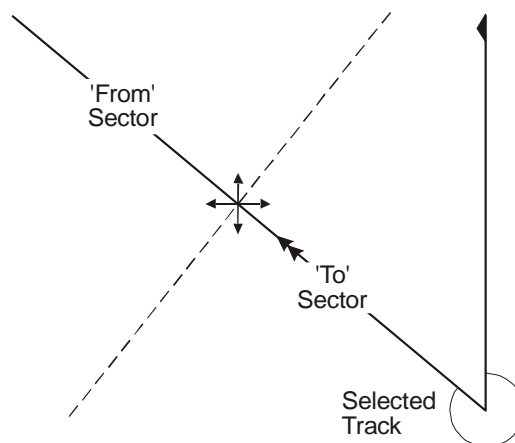


21. **Relative Bearing Indicator (RBI).** The RBI is similar to the RMI in mechanization except that the compass card is not driven by the gyro-magnetic compass. The card is fixed with 000° aligned with the aircraft fore-and-aft axis, ie at the top of the display. The pointer thus indicates a relative VOR bearing.

22. **Omni-Bearing Selector (OBS).** The OBS is the basis of many types of VOR display, which will normally offer the following features:

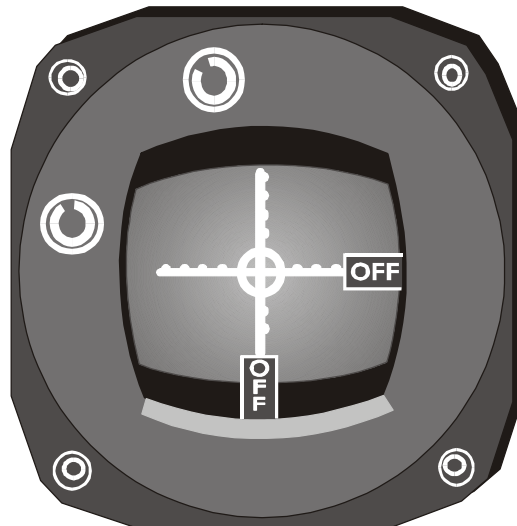
- a. **Track Selector.** The track selector can be used to either set up a desired track on which to approach or leave a beacon, or to determine the current bearing to, or radial from, a beacon. To fly a desired track the figure is set on the selector and the aircraft is then flown to zero the deviation indicator. The current radial or magnetic bearing to the beacon can be determined by adjusting the track selector such that the deviation indicator is zeroed.
- b. **TO/FROM Indicator.** The TO/FROM indicator shows the operator which sector the aircraft is in relative to the perpendicular to the selected track through the beacon (Fig 5). It does not indicate that the aircraft is heading towards or away from the beacon.

7-6 Fig 5 Operation of TO/FROM Indicator



- c. **Deviation Indicator.** The deviation indicator can be separate from or integral with the OBS. It may have either a VOR/ILS localizer needle alone or may be combined with an ILS deviation indicator (Fig 6). The deviation indicator will only be central when the aircraft is actually on the selected track. Full-scale deflection is approximately 10° off the selected track. The indicator is not a demand indicator in that it does not show the direction or amount by which the aircraft should be turned, as it takes no account of aircraft heading. It does indicate the direction relative to the selected track that the aircraft must be flown in order to achieve that track.

7-6 Fig 6 Deviation Indicator



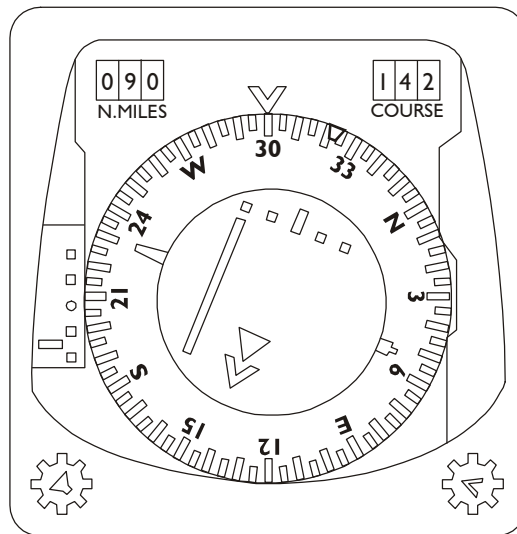
23. **Radial Magnetic Selector (RMS).** The RMS (Fig 7) is a combination of the RMI and OBS. The RMI features are the twin pointers and the gyro-magnetic compass driven compass card. The OBS features are the selection and numeric display of selected track, and the output to a remote deviation indicator. There is no TO/FROM indicator but a radial index always brackets the track selected.

7-6 Fig 7 Radial Magnetic Selector



24. **Horizontal Situation Indicator (HSI).** The HSI (Fig 8) is another OBS derivative and is a multi-purpose navigation display which, in addition to heading, can display VOR, ILS, DME and TACAN information with track selection and deviation indication, depending on the aircraft fit. The instrument is covered more fully in Volume 5, Chapter 13.

7-6 Fig 8 HSI

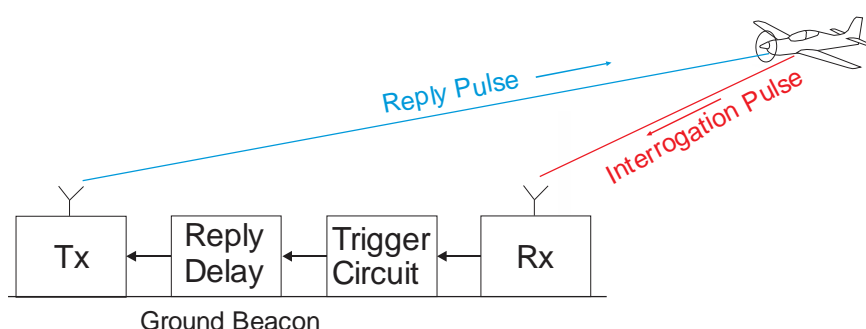


CHAPTER 7 - PRINCIPLES OF TACAN AND DME

Introduction

1. TACAN (TACTical Air Navigation System) and DME (Distance Measuring Equipment) are navigation aids which consist of ground beacons and airborne installations. They both operate in the UHF band from 962 to 1213 MHz. With a channel separation of 1 MHz, this provides for a maximum of 252 channels. Both TACAN and DME provide range information using the 'transponder' principle, which involves the ground beacon receiving interrogation pulses on one frequency and, after a set delay time, replying to the interrogating aircraft on a different frequency (see Fig 1). In addition, TACAN makes use of amplitude modulation to provide magnetic bearing.

7-7 Fig 1 TACAN/DME System



2. The main advantages of TACAN and DME are worldwide beacon coverage, high accuracy and ease of use without the need for special charts. TACAN can also be used in air-to-air mode (see para 41), which has tactical applications, eg in-flight refuelling.

3. The ground beacons are identified by two or three-letter Morse Code aural callsigns and have unique receiving and transmitting frequencies.

4. The airborne equipments differ in terms of information provided and operation as shown in Table 1.

Table 1 Airborne TACAN and DME Equipment Capabilities

| | TACAN | DME |
|-------------------|-----------|---|
| Bearing | Yes | No |
| Slant Range | Yes | Yes |
| Channel Selection | Numerical | By a VHF frequency |
| Channel Capacity | 126 ‡ | 100 (all equipments) ‡ 126 (some equipments) ‡ |
| Offset Capability | Yes † | No |
| Range Rate | No | Yes * |

* Not always mechanized in the airborne display.

† Capability provided by a computer external to actual TACAN equipment.

‡ Can be doubled if equipment has a Y channel capability.

PRINCIPLES OF OPERATION

5. As mentioned in the Introduction, range measurement is achieved by use of the transponder principle, whereby the ground beacon replies to interrogations by airborne equipment. The airborne interrogator transmits a continuous series of pulses on a given carrier frequency. The pulses are transmitted in pairs with a fixed time interval between the two pulses in the pair. The repetition interval between these 'pulse pairs' is varied – a technique known as random or 'jittered' Pulse Repetition Frequency (PRF). At the same time as the interrogations are transmitted, the airborne equipment commences timing and begins searching for replies from the ground beacon.

6. The ground beacon receives the interrogation and replies to it by sending out pulse pairs on a carrier frequency 63 MHz removed from the interrogation frequency. The airborne receiver, tuned to the ground station's transmitter frequency, receives all of the responses sent out but recognises its own replies by the unique jittered PRF. It calculates the slant range between the aircraft and the ground beacon by using the time interval between transmission of the interrogation and receipt of the reply.

7. There are two types of TACAN and DME beacons, X and Y. Both operate on exactly the same principle, but there are differences in detail. In the following paragraphs, which explain the principles of operation, figures used relate to X beacons. The differences between X and Y beacons are discussed further in paras 29 - 35.

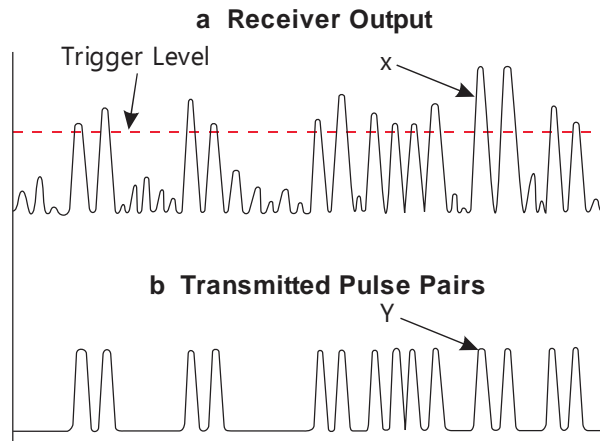
Ground Beacon Operating Cycle

8. The ground beacon consists of a separate transmitter and receiver which operate on frequencies separated by 63 MHz. For technical reasons, the transmitter operates continuously, sending out 2,700 pulse pairs per second (pp/sec). These pulse pairs are generated in response to a trigger circuit which is fed by the output from the ground beacon's receiver.

9. If the ground beacon does not receive any interrogations from airborne equipments, the output from the receiver will be random noise. This noise is fed to the trigger circuit which will fire on receipt of a pair of pulses above a set amplitude which are 12 μ sec apart (known as a '12- μ sec pulse pair'). The random noise from the receiver will contain a number of pulses which satisfy the firing conditions of the trigger circuit (see Fig 2a). When the trigger circuit fires, it prompts the beacon to transmit its own 12- μ sec pulse pair of a fixed amplitude.

10. If the receiver gain is increased, more pulse pairs will exceed the pre-set amplitude level and the trigger circuit will fire more frequently; the converse applies when the receiver gain is decreased. Therefore, the number of pulse pairs transmitted by the beacon varies directly with the receiver gain which is continuously and automatically adjusted to produce 2,700 pp/sec. These pulse pairs are of constant amplitude and shape, but have a random recurrence frequency (see Fig 2b).

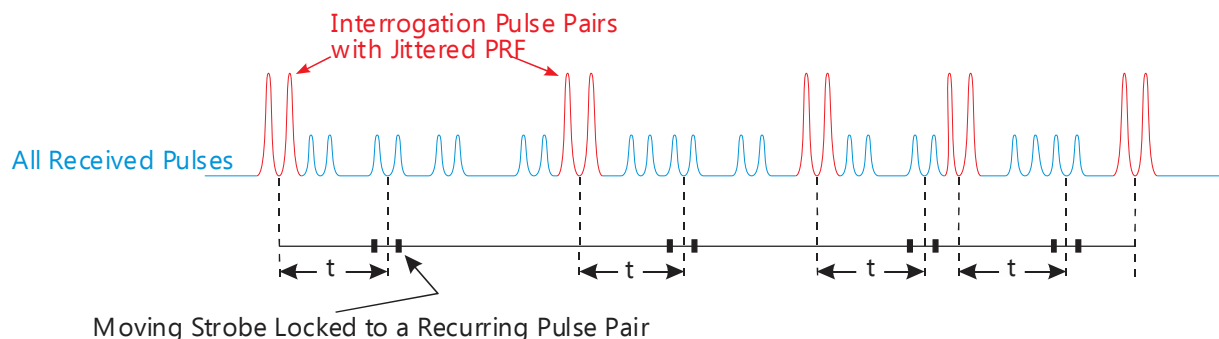
11. If the ground beacon is interrogated by an airborne equipment, a 12- μ sec pulse pair with an amplitude well above the pre-set level will appear at the receiver output, as at X in Fig 2a. The trigger circuit will fire and a reply will be transmitted to the aircraft amidst the pulse pairs produced irregularly by the noise from the receiver, as at Y in Fig 2b.

7-7 Fig 2 Generation of Ground Beacon Pulses

12. Since the ground beacon transmits a constant number of pulses per second (2,700), as the number of interrogation pulses received increases, so the number of random noise-generated pulses decreases. This is because the receiver gain will be reduced until the noise pulses with their smaller amplitudes fail to satisfy the requirements of the trigger circuit. As the standard rate of the interrogation by aircraft is 24 to 30 pp/sec (average 27), the trigger circuit will be saturated by the interrogations from about 100 aircraft and the receiver gain will be gradually reduced until replies are generated only by the strongest 100 interrogations.

Distance Measurement by TACAN and DME

13. **Identification of Correct Transponder Signals.** The airborne equipment will receive replies to its interrogations amidst all the other pulse pairs transmitted by the beacon. However, the airborne receiver must recognize its own replies in order to calculate the distance from the beacon. It can do this because it transmits interrogations at irregular intervals, i.e. the PRF is jittered, therefore the associated replies will always occur at a regular time (t) after each interrogation (see Fig 3). All other signals, whether they are random transmissions or replies to other aircraft, will be received at varying times with respect to each interrogation. Once the correct replies have been identified, the strobe follows small variations in position of the reply pulses as the aircraft's range changes – known as 'lock and follow'.

7-7 Fig 3 Airborne Matching of Replies with Interrogations

14. **Interrogation Pulses.** The initial rate of the interrogation used by the airborne transmitter is 150 pp/sec. This will generate a high reply rate from the ground beacon which is used to 'lock-on' the airborne equipment. In the airborne receiver, an automatic search strobe progressively examines small portions of the time base and counts the number of pulses occurring within a certain time. If the number of pulses is low, the strobe advances to examine the next portion of the time base. The search is stopped when the number of pulses received is

large (corresponding to the 150 pp/sec transmitted). The distance calculation is made using the standard pulse radar technique and the PRF of the interrogation drops from 150 pp/sec to between 24 and 30; the system is now 'locked-on'.

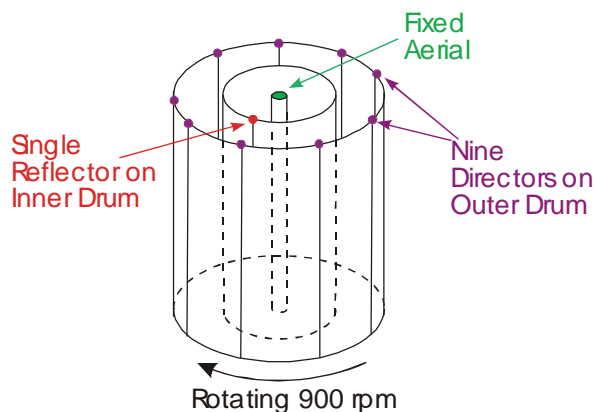
15. **Memory Function.** The memory function of the airborne equipment applies a 'search inhibit' time of approximately 10 seconds. If replies are interrupted for less than the search inhibit time, the distance display will be frozen by the memory circuit, and the search process will be inhibited. If no replies are received after the search inhibit time has elapsed, the search cycle will recommence and the distance counters will rotate (or go blank on digital equipment) until new lock-on is achieved.

Bearing Measurement by TACAN

16. The system described so far applies equally to TACAN and DME equipment. However, the ground TACAN beacon provides bearing information as well as range. The range pulses generated by the ground beacon are of constant shape and amplitude. In order to convey bearing information they are amplitude modulated by the aerial system. As the beacon transmits 2,700 pp/sec, the modulation is carried as accurately as if a continuous wave (CW) carrier was used.

17. The beacon's aerial system comprises a fixed central vertical aerial around which two concentric fibreglass drums rotate at 15 revolutions per second. Embodied in the surface of the inner and outer drums respectively, are one vertical reflector and nine equally spaced directors (see Fig 4).

7-7 Fig 4 TACAN Ground Beacon Aerial System



18. The effect of the reflector is to offset the centre of the resultant polar diagram, as shown in Fig 5.

7-7 Fig 5 Polar Diagram for Fixed Aerial and Reflector

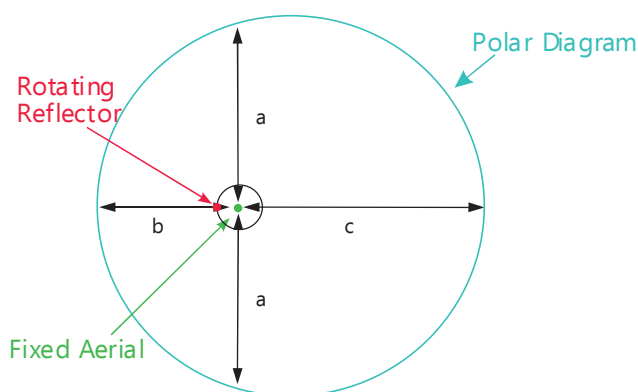
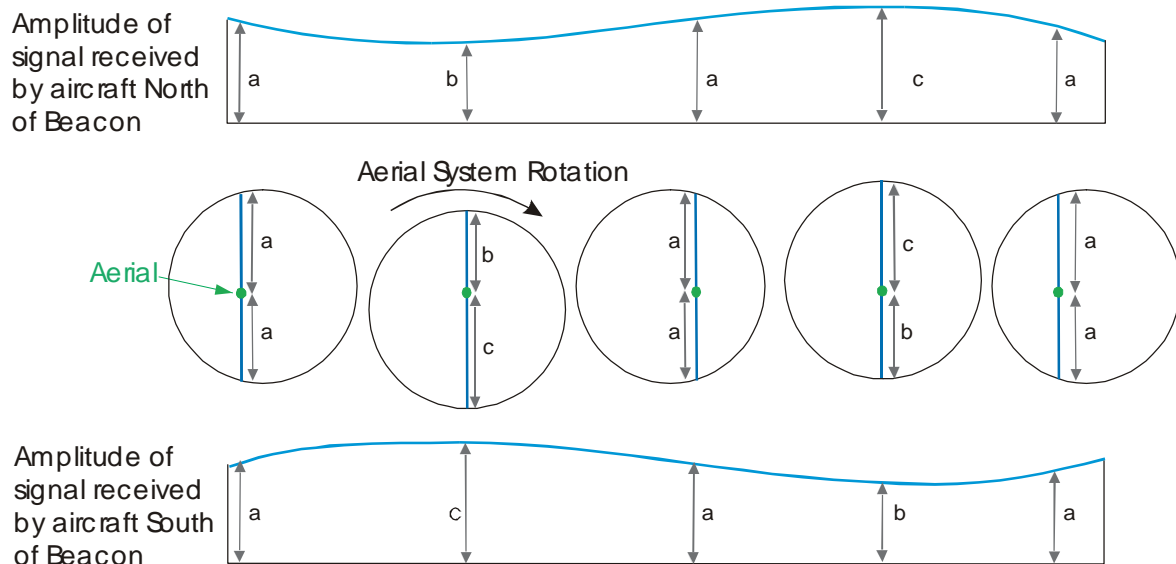


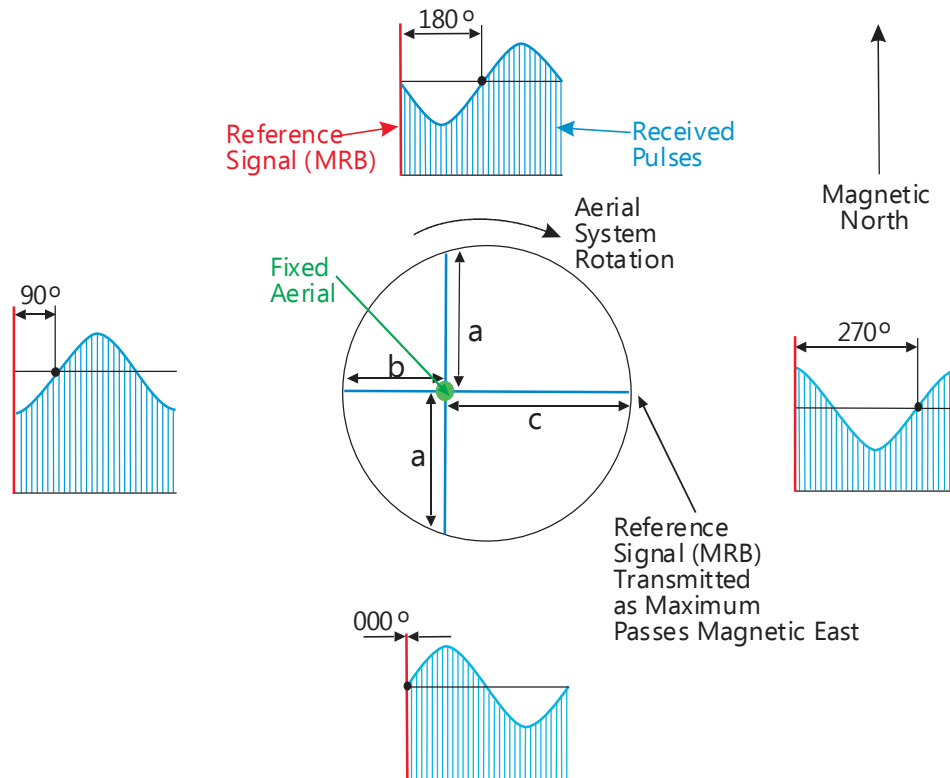
Fig 6 shows the effect of rotating the inner drum through 360°. It can be seen that, at any one point, the amplitude of the received pulse varies through one sinusoidal cycle for each rotation of the polar diagram. The phase of the received signal will depend on the relative location of the receiver from the transmitter. Since the reflector is rotating at 15 revolutions per second (900 rpm), the polar diagram will also rotate at the same rate, generating an output signal amplitude modulated at 15 Hz.

7-7 Fig 6 Amplitude Modulation due to Aerial Rotation



19. In order to calculate a bearing, by utilizing the difference in the phase of the received signal due to location of the receiver, a fixed reference signal is needed. This reference is supplied by the ground beacon transmitting a short train of 12 evenly-spaced 12-μsec pulse pairs each time the maximum of the polar-diagram passes through magnetic East.

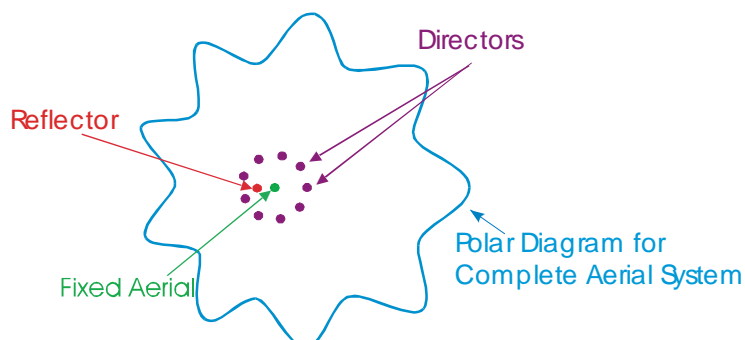
20. The airborne receiver recognizes this distinctive pulse train, known as the Main Reference Burst (MRB), and measures the phase of the modulated signal relative to it, thereby computing the bearing (see Fig 7). As the MRB is transmitted when the polar diagram peak passes magnetic East, if the received reference signal coincides with the peak of the received modulated signals, then the aircraft is to the East of the beacon. In practice, it is easier to compare the position of the reference signal relative to the zero point on the modulated signal (shown by a black dot on the graphs in Fig 7). Relative bearings between the two signals for aircraft on the cardinal bearings from the beacon are shown in Fig 7. It can be seen that, for example, the position to the East of the beacon shows a bearing of 270° **TO** the beacon.

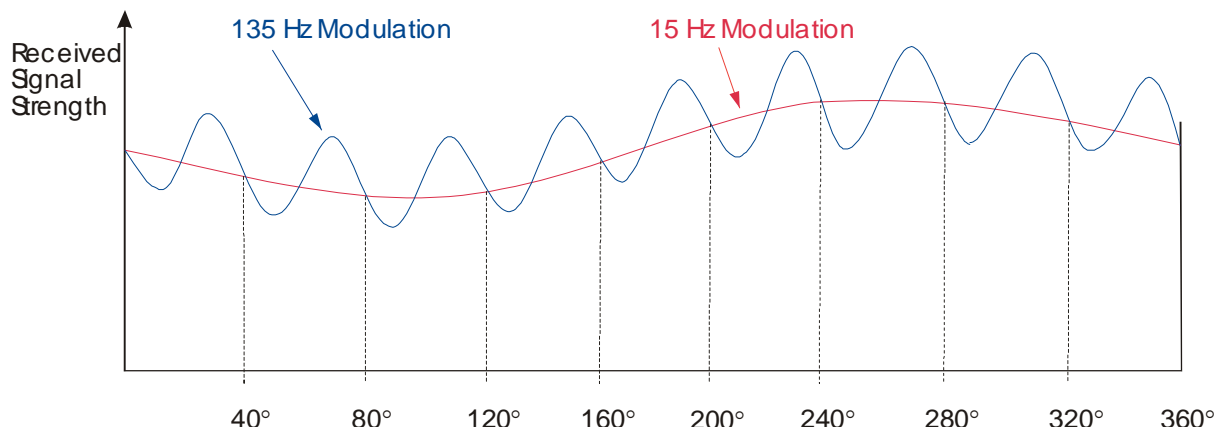
7-7 Fig 7 Determination of Bearing by Phase Comparison

Note: There are two zero points on the sinusoidal wave for each complete cycle. Electronic circuitry within the airborne equipment ensures that the zero point which occurs as the signal amplitude is increasing is the one selected.

21. In the system described so far, there is a 1:1 relationship between the aircraft bearing and the phase of the received signal, ie 1° of phase difference corresponds to 1° of bearing. Therefore, any errors in phase difference, caused either by measurement inaccuracies or phase shifts due to reflections of the signals, would have a direct, and major, effect on the computed bearing.

22. The nine directors in the aerial's outer rotating drum introduce a fine bearing facility which improves the accuracy of the measured bearing. As can be seen from Fig 8, the effect on the polar diagram of introducing the directors is to generate nine corresponding lobes. When this complete aerial system is rotated 15 times per second, a further 135 Hz (9×15) modulation is superimposed on the 15 Hz modulation as shown in Fig 9.

7-7 Fig 8 TACAN Aerial Polar Diagram

7-7 Fig 9 Modulation Envelope

23. As can be seen from Fig 9, one complete cycle of the 135 Hz component is contained within a 40° phase change in the 15 Hz component, ie there are 9 cycles of the 135 Hz component for each cycle of the 15 Hz component. This means that there is now a 1:9 relationship between the aircraft's bearing and the phase of the received 135 Hz signal; a change of 1° in phase of the signal representing only 1/9° bearing change. Therefore, by using the 135 Hz signal, the effect on the accuracy of the measured bearing due to distortion of the received signal is reduced.

24. **The Auxiliary Reference Burst.** A secondary reference signal, of six 12-μsec pulse pairs, is transmitted when each of the nine maxima of the complete polar diagram passes through magnetic East; this is known as the Auxiliary Reference Burst (ARB). These ARBs are used to determine the aircraft's bearing, since the position of the 135 Hz component relative to the reference signals is also a measure of the aircraft's bearing from the ground beacon.

25. **Bearing Ambiguity.** Because the phase of the 135 Hz component changes through one complete cycle for each 40° change of bearing, the same phase difference could occur in any one of the nine 40° sectors around the beacon. This ambiguity is resolved by the bearing obtained from the 15 Hz component, ie the coarse bearing obtained from the MRB.

26. **Fine Bearing Calculation.** Having resolved the bearing ambiguity, the fine bearing could be taken from any of the 9 detector-generated lobes, as they should all be the same. However, to increase accuracy further, an average of the phase differences of the nine lobes is used. This method is more accurate than relying upon one measurement.

Beacon Identification

27. Coded pulses giving a two or three-letter aural callsign in Morse Code are transmitted every 37.5 sec to identify the beacon.

28. **Unreliability Coding.** When an unserviceable beacon is required to transmit, eg for calibration, the aural identification signal is followed by a series of dots (normally five). Bearing and distance information from a beacon coding in this manner should be regarded as unreliable.

X and Y TACAN and DME Beacons

29. The ground equipment described so far is classified as an 'X' beacon (TACAN or DME). In order to increase the number of channels from 126 to 252 within the same frequency allocation, modified beacons known as 'Y' beacons are employed.

30. A Y-type ground beacon receiver operates on the same frequency as an X-type ground beacon receiver of the same channel number, but its transmitter frequency differs by 126 MHz. To determine the receiver frequency for a given TACAN ground installation, add 1024 to the channel number; this will give the frequency in MHz. For example, a channel 27 ground beacon receiver is tuned to 1051 MHz and channel 70 receives on 1094 MHz. Setting the channel number on the aircraft equipment determines the frequency of the airborne transmission, thereby selecting the ground station which will respond. The ground station transmitter frequencies are determined as follows:

- a. **Low Band.** The low band covers the frequency range 962 MHz to 1087 MHz (Channel Nos 1 to 63). X beacons in this band transmit on frequencies 63 MHz lower than the corresponding receiver frequencies whereas Y beacons transmit on frequencies 63 MHz higher, eg channel 27X would transmit on 988 MHz (1051 minus 63) while channel 27Y would transmit on 1114 MHz (1051 plus 63).
- b. **High Band.** The high band covers the frequency range 1088 MHz to 1213 MHz (Channel Nos 64 to 126). In this band, the differences between X and Y types are reversed in that X beacons transmit on frequencies which are 63 MHz higher than the receiver frequencies while Y beacons transmit on frequencies 63 MHz lower, eg channel 70X transmits on 1157 MHz (1094 plus 63), whereas channel 70Y transmits on 1031 MHz (1094 minus 63).

31. Normally, TACAN beacons have the suffix X or Y quoted in addition to the actual channel number, although it can be assumed that any beacon without a suffix is an X beacon. DME beacons are distinguished by the final digit of the VHF frequency, a '5' signifying a Y beacon, and a '0' an X beacon, eg 114.15 is a Y beacon frequency while 114.20 is for an X beacon. The other major differences between X and Y beacons are:

- a. **Pulse Pair Time Intervals.** The time interval between the pulses that make up the pulse pairs is different in X and Y beacons. The X spacing is 12 μ sec for both the beacon receiver and transmitter, whilst the Y beacon is designed to receive pulses with an interval of 36 μ secs and to transmit pulses with an interval of 30 μ secs.
- b. **Reply Delay Time.** The X beacon has a reply delay time of 50 μ secs and the Y beacon one of 74 μ secs. The airborne equipment takes this delay into account when calculating slant range.
- c. **Main and Auxiliary Reference Bursts (TACAN only).** The MRB and ARB are different for X and Y TACAN beacons. For X beacons, the MRB consists of 12 pulse pairs at 30 μ sec intervals while Y beacons emit 13 single pulses at 30 μ sec intervals. The ARB for X beacons is 6 pulse pairs at 24 μ sec intervals but for Y beacons it is 13 single pulses at 15 μ sec intervals.

32. In order to operate with a Y beacon the airborne equipment must be modified to:

- a. Produce the 36- μ sec spaced pulse pairs.

- b. Select the appropriate receiver frequency (126 MHz different to the equivalent X channel).
 - c. React to 30-µsec spread pulse pairs and ignore those spaced at 12 µsec from an X beacon.
 - d. Compensate for the different Reply Delay Time.
 - e. React to the different MRB and ARB and ignore those produced by an X beacon (TACAN only).
33. Details of TACAN channel allocation, and the corresponding VHF frequencies for setting on DME controllers, are shown in the 'Frequency and Channel Pairing Tables' in the Flight Information Handbook. It can be seen from these tables that there are 100 X channels and 100 Y channels available between 108.00 MHz and 117.95 MHz. In addition, there are a further 26 X channels on frequencies between 133.30 MHz and 135.90 MHz. As some older VOR/DME controllers are physically limited to selecting frequencies between 108 and 117.95 MHz, they can not access the additional 26 channels. This explains the entry in Table 1 under Channel Capacity.
34. The majority of the TACAN and DME beacons in use within the UK are of the X type; the number of Y types in use worldwide is steadily increasing.
35. Note that to use a TACAN airborne equipment with a beacon that has a channel number with a Y suffix (TACAN) or with a DME beacon that has a frequency ending in a 5, eg 115.35, the Y function must be selected. For an airborne DME equipment that has a Y capability, dialling the appropriate VHF frequency is all that is required. The X or Y function will be selected automatically.

USE OF TACAN AND DME

Navigational Use

36. The maximum range of TACAN and DME primarily depends on the aircraft height since radio waves in the 1,000 MHz band are quasi-optical in character, ie they are 'line-of-sight'.
37. **Bearing.** Bearings generated by TACAN beacons are magnetic. A TACAN ground beacon has a 70° cone above it in which the bearing function is inoperative. When TACAN bearing information is displayed, the arrowhead of the needle points towards the beacon. Therefore, if plotting a position from the beacon, the tail end of the needle should be used, as this will give the reciprocal magnetic bearing required.
38. **Range.** It should be remembered that the distance shown on the airborne equipment is slant range, not ground range, from the beacon.

Interference

39. To overcome the problem of possible mutual interference, the operational requirement for each beacon for coverage in range and altitude has been specified. Each ground installation has been protected against interference up to the stated range and altitude. Outside this 'protected range and altitude', interference may occur and navigational information may be unreliable. The protected ranges and altitudes of beacons are listed in Flight Information Publications.

Accuracy

40. The bearing information from a TACAN system should be accurate to $\pm 0.5^\circ$. Distance measurements for TACAN and DME are accurate to ± 0.1 nm or $\pm 1\%$ of the distance, whichever is the greater.

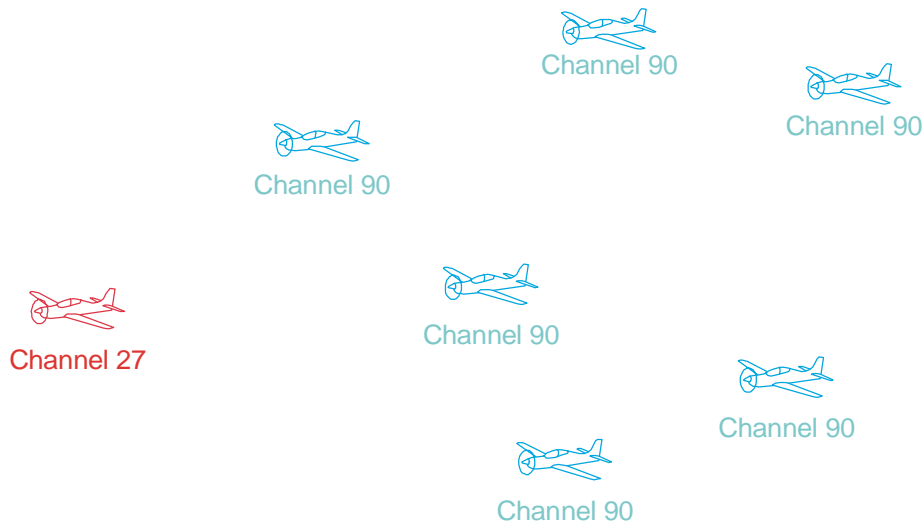
Air-to-Air Mode

41. The air-to-air mode of TACAN enables suitably equipped aircraft to measure the range between each other, up to a maximum of 195 nm; bearing information is not available. The maximum number of interrogating aircraft to which one 'responder' can reply simultaneously is theoretically 33. However, there are many factors which degrade theoretical performance and it is impossible to state precisely the maximum number of aircraft that can lock-on simultaneously in practice.

42. The changeover between air-to-ground and air-to-air operation is controlled by a selector switch on the airborne equipment. This switch affects the way in which the signals are processed internally and also changes the receiving frequency of the airborne equipment; transmitter frequencies are unaltered. For channel numbers in the Low Band, receiver frequencies are increased by 126 MHz compared to the equivalent X channel in Air-to-Ground mode, eg channel 27X in Air-to-Ground Mode would transmit on 1051 MHz and receive on 988 MHz whereas setting channel 27 in Air-to-Air mode changes the receiver frequency to 1114 MHz (988 plus 126). Similarly, for channels in the High Band, the receiver frequencies are lowered by 126 MHz compared to the equivalent X channel in Air-to-Ground Mode, eg channel 90X transmits on 1114 MHz and would receive on 1177 MHz in Air-to-Ground mode, whereas in Air-to-Air mode the receiver frequency becomes 1051 MHz (1177 minus 126).

43. The importance of the changes to receiver frequency discussed in the previous paragraph can be seen when the practical issues of air-to-air operation are considered. For mutual range between two aircraft to be available, the transmitter frequency of one aircraft must be the same as the receiver frequency of the other, and vice versa. Therefore, in practice, the two aircraft must set TACAN channels whose numbers are 63 part. Using the examples in para 42, if one aircraft sets channel 27, its transmission will be on 1051 MHz and reception on 1114 MHz. In the other aircraft, if the TACAN is set to channel 90 it will receive on 1051 MHz and transmit on 1114 MHz. The two aircraft are now matched, and each will receive range information from the other aircraft.

44. This method of mutual ranging has application in formation flying as well as air-to-air refuelling. Fig 10 shows how TACAN could be used in these circumstances. If the formation leader sets channel 27 on his controller, and the other members of the formation select channel 90, then all members of the formation will get a readout of their individual range to the leader. The leader will get an indication of distance to the aircraft whose signal is strongest, probably, although not necessarily, the closest one to him.

7-7 Fig 10 Use of Air-to-Air TACAN in Formation**TACAN Offset Computer**

45. The addition of a TACAN offset computer to an airborne installation enables bearings and distances to be obtained from a point which is offset from the TACAN beacon.

46. The range and bearing of the offset point relative to the TACAN beacon are manually set on the computer control. The computer then converts the range and bearing into electrical signals and transmits them to the navigation display where they are added to the direct TACAN signals. The combination of signals causes the display to indicate the range and bearing to the offset point instead of to the TACAN beacon itself.

Range Rate

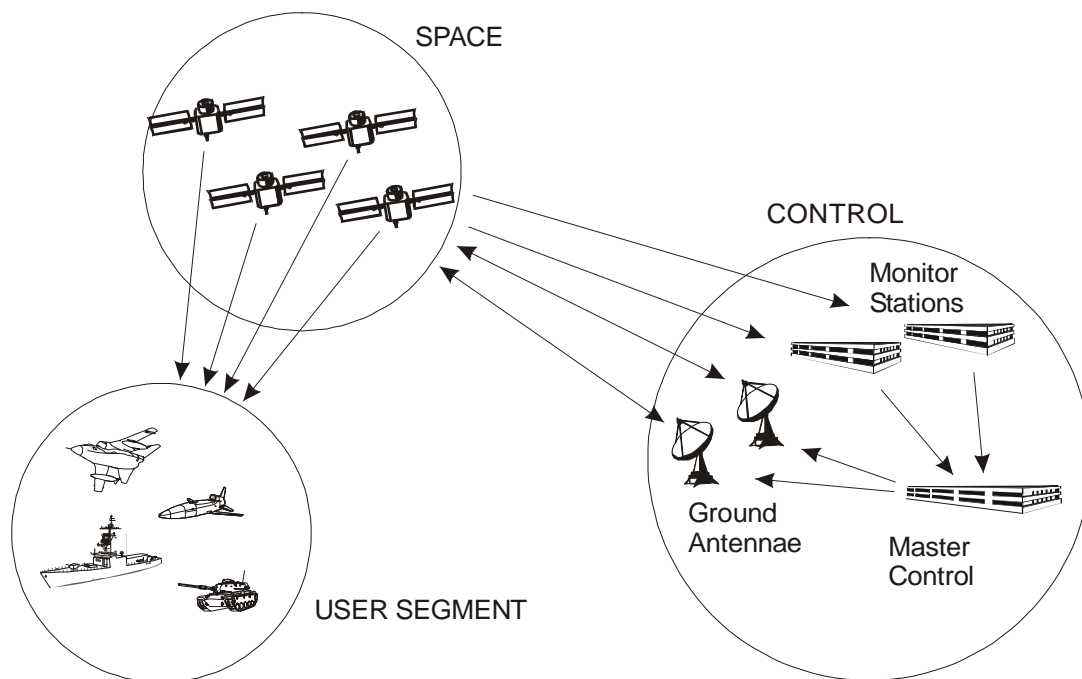
47. The Range Rate capability shown in Table 1 for DME equipment is a facility available on some airborne installations whereby the aircraft's groundspeed is calculated from the rate of change of distance to a DME beacon. The feature only works accurately for beacons along the aircraft's track, and depends on having a dedicated microprocessor to perform the necessary calculations.

CHAPTER 8 - GLOBAL POSITIONING SYSTEM (GPS)

Introduction

1. The NAVSTAR Global Positioning System (GPS) is a US space-based radio positioning system, using a constellation of satellites to provide highly accurate position, velocity, and time (PVT) data. There are three major segments to the system; Space, Control and User (see Fig 1). The system is available globally, continuously, and under all weather conditions to users at or near the Earth's surface. However, the signals are quasi-optical and there must be a direct line of sight between the satellites and the receiver aerial for full operation. As the receivers operate passively there can be an unlimited number of simultaneous users. GPS has features designed to deny an accurate service to unauthorized users, prevent spoofing and to reduce receiver susceptibility to jamming.

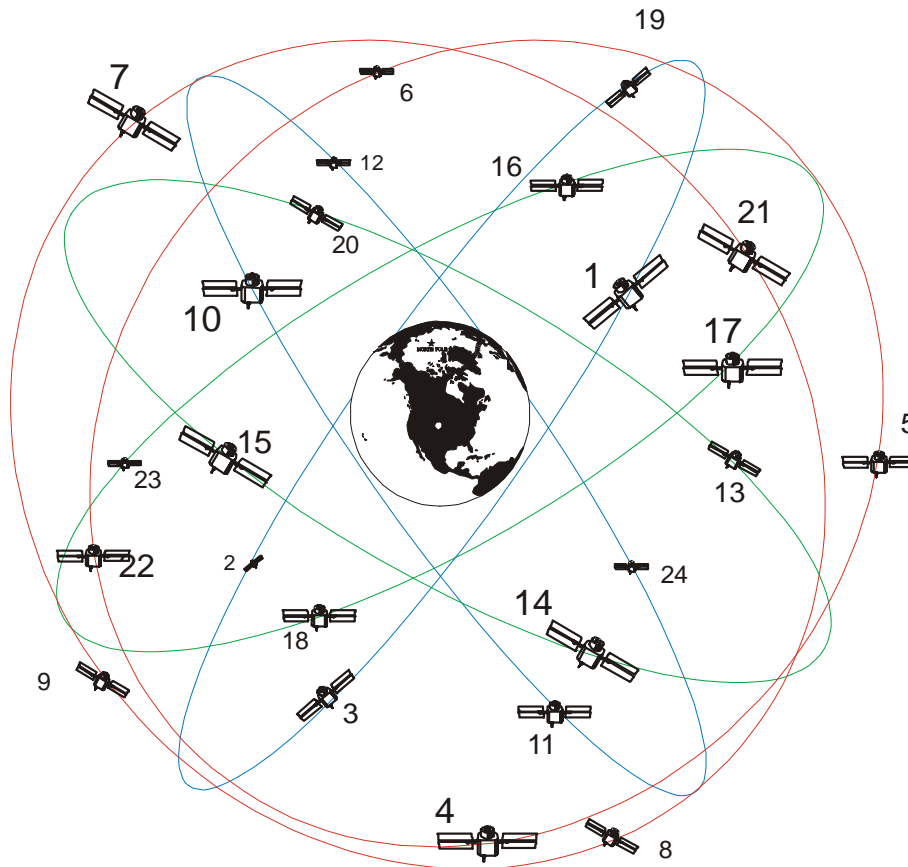
7-8 Fig 1 Major Segments of GPS System



2. Each satellite broadcasts radio frequency ranging codes and a navigational data message. The GPS receiver measures the transit time of the signals and can thereby determine its range from a satellite. The data message enables the receiver to determine the position of the satellite at the time of transmission, and thus by ranging from 3 or more satellites a fix can be obtained.

Space Segment

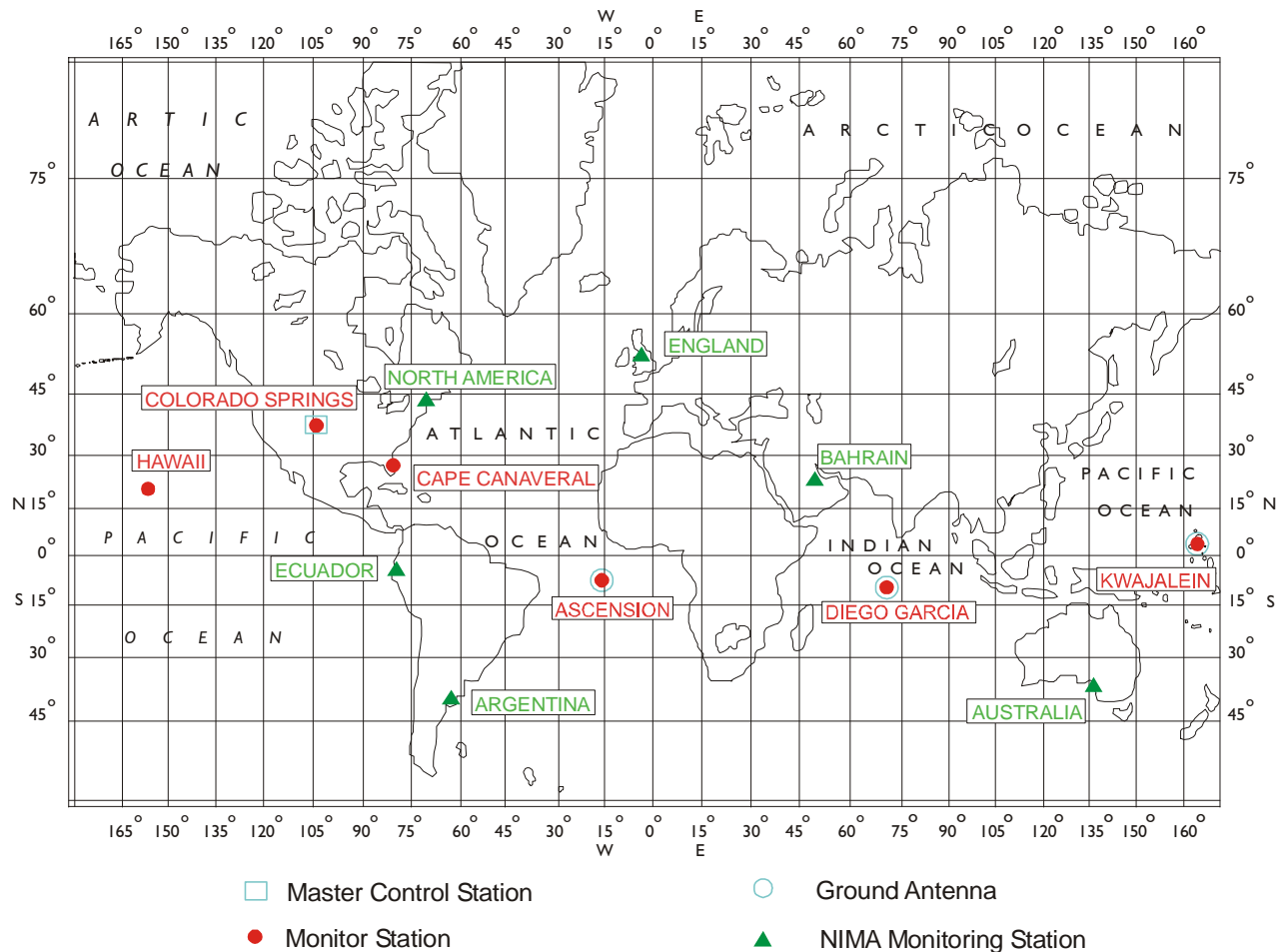
3. The GPS space segment consists of 24 operational satellites (see Fig 2). There is a continuing programme to replace and maintain the space segment to account for technical failures of satellites in service and faults during launch so this figure of 24 may vary slightly as new and failing orbiters are exchanged. The satellites are in six orbital planes with three or four operational satellites in each orbit. The orbit height is 20,200 km and the orbital planes are inclined at an angle of 55° to the equator. A satellite takes approximately 12 hours to complete an orbit. The satellites are positioned such that at least 5 satellites are normally observable by a user anywhere on the Earth.

7-8 Fig 2 GPS Satellite Constellation

4. The satellites transmit on two frequencies: Link 1 (L1) on 1575.42 MHz and Link 2 (L2) on 1227.6 MHz. A coarse/acquisition code (C/A) is transmitted on L1 and a precision code (P) on both L1 and L2; either code can be used by a receiver to determine its range from a satellite. Superimposed on both codes is the NAVIGATION-message (NAV-msg) containing accurate satellite orbit data (known as 'ephemeris data'), ionospheric propagation correction data, and timing information. These aspects are covered more fully in later paragraphs.

Control Segment

5. The control segment consists of one Master Control Station (MCS) in Colorado Springs (USA) together with additional monitor stations at Hawaii, Kwajalein (Western Pacific), Diego Garcia (Indian Ocean), and Ascension (S Atlantic). Additional monitoring stations, operated by the US National Imagery and Mapping Agency (NIMA), provide increased visibility of the GPS satellites (see Fig 3). The monitoring stations on Kwajalein, Diego Garcia, and Ascension are equipped with ground antennae for communication with the satellites. The monitor stations passively track all GPS satellites in view, collecting ranging data which is passed to the MCS where each satellite ephemeris and clock parameters are estimated and predicted. The MCS periodically uploads this data to each satellite for re-transmission in the NAV-msg.

7-8 Fig 3 GPS Control Segment Locations

User Segment

6. The user segment comprises a wide variety of military and civilian GPS receivers which can decode and process the satellite signals. The receivers may be 'stand alone' sets, or part of integrated systems, and can serve a variety of applications such as navigation and surveying. Receivers for different applications can vary significantly in their design and function.

Levels of Service

7. Two levels of navigation accuracy are provided by the GPS: the Precise Positioning Service (PPS) and the Standard Positioning Service (SPS). The PPS is a highly accurate positioning, velocity, and timing service which is made available only to authorized users, whereas the SPS is a less accurate service available to all GPS users.

8. **Precise Positioning Service (PPS).** The PPS is primarily intended for military users and the authorization for use is decided by the US Department of Defense. Authorized users include the US and NATO military and the Australian Defence Forces. The PPS is specified to provide a 16 metre Spherical Error Probable (SEP) (3-D, 50%) positioning accuracy and 100 nanosecond UTC time transfer accuracy. This equates to approximately 37 metres and 200 nanoseconds (3-D, 95%) under typical system operating conditions. Depending on receiver design it is possible to obtain a 3-D velocity accuracy of 0.1 metres per second. A selective availability (SA) feature may be used to reduce

the GPS PVT accuracy to unauthorized users by introducing controlled errors into the signals. The SA level of degradation can be varied and might, for example, be increased in time of crisis or war to deny accuracy to a potential enemy. An anti-spoofing (A-S) feature is invoked to negate potential hostile imitation of PPS signals. The technique alters the P code cryptographically into a code denoted as P(Y) code. Encryption keys and techniques are provided to PPS users which allow them to remove the effects of the SA and A-S features. PPS receivers can use either the P(Y) or the C/A code or both. Maximum accuracy is obtained using the P(Y) code on both L1 and L2. The difference in propagation delay between the two frequencies is used to calculate ionospheric corrections. P(Y) capable receivers commonly use the C/A code initially to acquire GPS satellites and determine the approximate P(Y) code phase although some P(Y) receivers are able to acquire the P(Y) code directly by using a precise clock. Some PPS receivers use only one frequency, L1, and in this case they use an ionospheric model to calculate corrections. Typically this will result in less positioning accuracy than dual frequency P(Y) receivers (approximately 65 metres (3-D, 95%) accuracy).

9. **Standard Positioning Service (SPS).** The SPS is primarily intended for civilian use and is specified to provide a 100 metre (95%) horizontal positioning and 156 metre 3-D (95%) accuracy. UTC time dissemination accuracy is within 340 nanoseconds. The accuracy specification includes the degradation of SA which is the dominant source of error. SPS users can only access the L1 frequency and therefore cannot measure the propagation delays of L1 and L2 signals in order to determine ionospheric corrections. Typically, therefore, an SPS receiver uses only the C/A code and an ionospheric model to calculate corrections, which is a less accurate technique than measuring dual frequency propagation delays. The accuracy specification also includes this modelling error source.

Receiver Operation

10. In order for the GPS receiver to navigate, it has to acquire and track satellite signals to make range and velocity measurements, and to collect the NAV-msg data. Because the clocks in the GPS receivers are not synchronized with system time there is a bias in the ranges actually measured and they are termed pseudo-ranges.

11. **Satellite Signal Acquisition.** The satellite signal level near the Earth is less than the background noise, therefore correlation techniques are used by the GPS receiver to obtain the navigation signals. A typical satellite tracking sequence begins with the receiver determining which satellites are visible for it to track. Satellite visibility is based on the user-entered estimates of present PVT and stored satellite almanac information. If no stored data exists or if only very poor estimates of position and time are available the receiver must search the sky in an attempt to locate randomly and lock on to any satellite in view. If the receiver can estimate satellite availability it will target a satellite to track. Once one satellite has been acquired and tracked the receiver can demodulate the NAV-msg and read almanac information about all of the other satellites in the constellation. The receiver has a carrier-tracking loop which is used to track the carrier frequency and a code-tracking loop which tracks the C/A and P code signals. The two tracking loops work together in an iterative process in order to acquire and track the satellite signals.

12. **Carrier Tracking.** The receiver's carrier-tracking loop generates an L1 carrier frequency which differs from the received carrier signal due to a Doppler shift. This Doppler offset is proportional to the relative velocity along the line of sight between the satellite and the receiver, plus a residual bias in the receiver frequency standard. The carrier signal strength is lower than the background and so has to be code correlated by the code-tracking loop in order to be seen through the noise. The carrier-tracking loop adjusts the frequency of the receiver-generated carrier until it matches the incoming carrier

frequency and thereby determines the relative velocity between the receiver and the satellite being tracked. The receiver uses the velocity relative to four satellites being tracked to determine its velocity in the navigation reference frame. The velocity output of the carrier-tracking loop is used to aid the code-tracking loop.

13. Code Tracking. The code-tracking loop is used to make pseudo-range measurements between the receiver and satellites. The receiver's code tracking loop generates a replica of the targeted satellite's C/A code with estimated ranging delay. The phase of the replica code is compared with the phase of the received signal code and the difference is directly proportional to the pseudo-range between the receiver and satellite. In general, prior to tracking, the receiver-generated code will not correlate with the received code due to the time taken for the satellite signal to reach the receiver and the receiver's clock bias error. The receiver will therefore slew its generated code through a one-millisecond cycle search window to achieve C/A code tracking. However, in the case of the P code, every week each satellite is allocated a different 7-day portion of a 267-day-long code and an approximate P code phase must therefore be known in order to obtain signal lock. Information on the P code phase is contained in the NAV-msg and the receiver uses this data, together with the C/A code-derived pseudo-range to minimize the P code search window. It is feasible for a P code receiver to acquire the P code without first acquiring the C/A code, but this requires a good knowledge of the receiver position and a very good knowledge of the GPS time; an external atomic clock would usually be required.

14. Navigation. When the receiver has acquired the satellite signals from four GPS satellites, achieved carrier and code tracking, and has read the NAV-msg, the GPS receiver is ready to start navigating. The GPS receiver normally updates its pseudo-range and relative velocities once every second. The next step is to calculate the receiver position, velocity, and GPS system time. Each satellite signal includes its time of transmission in GPS time, and the receiver must determine system time very accurately in order to measure the reception of signals in the same time reference. The difference in these times is directly proportional to the actual range between receiver and satellite. However, it is not necessary for the receiver to have a highly accurate clock such as an atomic time standard. Instead a crystal oscillator is used and the receiver computes its offset from GPS system time by making four pseudo-range measurements which are used to solve four simultaneous equations with four unknowns. Once solved the receiver has estimates of its position and GPS system time. The receiver velocity is calculated using similar equations with relative velocities instead of pseudo-ranges. GPS receivers perform most calculations using an Earth centred, Earth fixed co-ordinate system which is subsequently converted to a geographic co-ordinate system using an Earth model (World Geodetic System 1984 (WGS 84)). During periods of high jamming the receiver may not be able to maintain both code and carrier tracking, although it will normally be able to maintain code tracking even when carrier tracking is not possible. If this is the case the receiver will slew the locally generated carrier and code signals based on predicted rather than measured Doppler shifts. These predictions are performed by the navigation processor which may have additional PVT information from other sources.

Types of GPS Receiver

15. There are four types of GPS receiver: sequential tracking, continuous tracking, multiplex and 'All-in View'. The following descriptions apply to P code receivers which use dual frequency measurements to determine ionospheric corrections. C/A code receivers are essentially the same except that they use a model of the ionosphere to determine corrections.

Sequential Tracking Receivers

16. A sequential receiver tracks the necessary satellites by using one or two hardware channels. The set will track one satellite at a time and combine all the four pseudo-range measurements once they have been made. These receivers are amongst the cheapest available but cannot operate under high-speed scenarios and have the slowest time to first fix.

17. A P code, one-channel, sequential receiver makes four pseudo-range measurements on both the L1 and L2 frequencies in order to determine a position and compensate for ionospheric delay. The search for a satellite, code correlation, NAV-msg reading, data demodulation and ionospheric measurement is accomplished for each satellite in turn and the four pseudo-ranges must be corrected to a common time before a navigation solution can be achieved. Any movement of the vehicle during the time taken by the receiver to collect the pseudo-ranges will result in a degradation of the navigation solution. The use of one-channel sequential receivers is therefore limited to low speed or stationary applications.

18. Two-channel sequential receivers have been developed for use on medium speed vehicles such as helicopters. During the initial power-up each channel operates like a one-channel sequential receiver. Once four satellites have been acquired one channel is dedicated to navigation while the other reads the NAV-msg from each satellite. Both channels are used to perform dual frequency measurements to compensate for ionospheric delay. Two-channel sequential receivers decrease the time it takes to start navigating by better than one minute compared to one-channel sequential receivers.

Continuous Tracking Receivers

19. A continuous tracking receiver must have at least four hardware channels in order to track four satellites simultaneously. GPS receivers are available with up to 12 channels, but, due to their greater complexity, multiple channel sets involve proportionally higher costs. Four and five channel sets offer suitable performance and versatility, tracking 4 satellites simultaneously; a five-channel receiver uses the fifth channel to read the NAV-msg of the next satellite to be used when the receiver changes the satellite selections. The fifth channel is also used in conjunction with each of the other four for dual frequency measurements. A continuous tracking receiver is the best for high-speed vehicles such as fighter aircraft and those requiring a short time to first fix such as submarines. It also exhibits a good anti-jamming performance.

Multiplex Receivers

20. A multiplex receiver switches at a fast rate (typically 50 Hz) between the satellites being tracked, continuously collecting sampled data to maintain two to eight signal processing algorithms in software. The NAV-msg data is read continuously from all the satellites. For a receiver tracking four satellites this results in the equivalent of eight channels delivering 20 parameters continuously; four L1 code phases, four L2 code phases, four L1 carrier phases, four L2 carrier phases, and four NAV-msgs. Four pseudo-ranges, four velocities and the ionospheric delay can be derived from these twenty parameters.

'All-in-View' Receivers

21. GPS receivers traditionally choose the four satellites of those available which give the best geometry to perform a position fix. However, in situations where one or more of the satellites are temporarily obscured from the antenna's view, the receiver will have to acquire additional satellite signals to generate a PVT solution. The solution will be degraded until the new satellite is acquired. In order to overcome this problem a receiver can be designed to use all available satellites in view, typically six or seven, to generate the solution. If one or two satellites are then lost from view there will be little or no loss of

accuracy. The receiver will need a channel for each satellite or will have to use a multiplex technique with the attendant penalties of increased hardware, weight, power consumption and cost.

Automatic Dependent Surveillance (ADS)

22. The incorporation of a datalink with a GPS receiver enables the transmission of aircraft location to other aircraft and/or air traffic control. This function, called ADS, is in use in some Oceanic regions. The key benefit is to enable ATC monitoring in regions where radar services are not available. ADS has potential use in many other monitoring situations, including ground control of aircraft and vehicles at airfields.

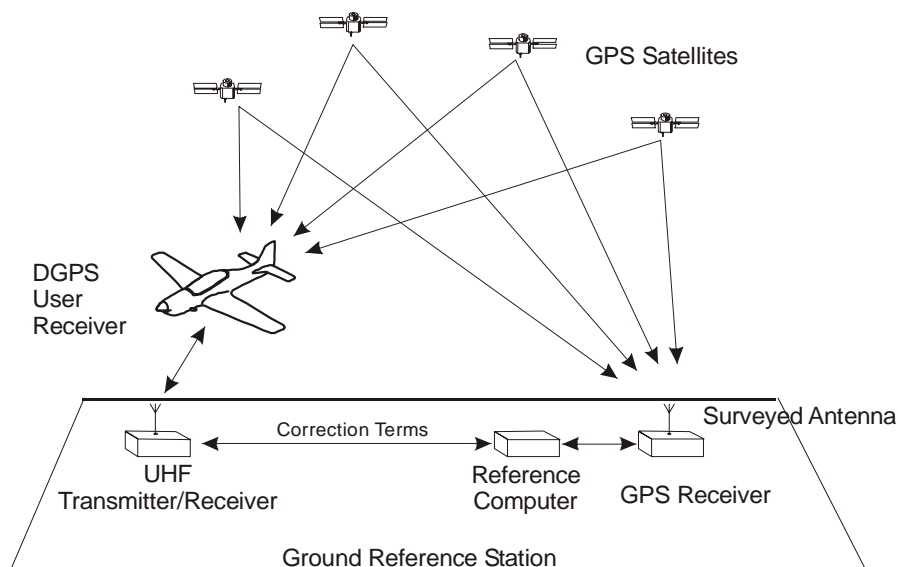
Differential GPS (DGPS)

23. GPS provides a worldwide navigational facility which is extremely accurate when compared to systems previously available to the aviator. However, it remains subject to system errors which will affect the precision at user level. These errors include:

- a. Selective Availability (see para 8).
- b. Ionospheric delay to signal propagation, as a result of solar activity (this is also a function of time of day).
- c. Tropospheric delay to signal propagation, caused by moisture in the lower atmosphere.
- d. Satellite ephemeris error, which is the difference between actual and predicted satellite location.
- e. Satellite Clock Error.

24. The method chosen to minimize some of these errors is through 'differencing'. This concept requires use of a ground reference station, surveyed to a high level of accuracy. The reference station receives the GPS signals, and calculates the difference between GPS and surveyed positions (see Fig 4). The reference station then computes corrections and transmits them to the user via UHF or data link. The corrections are applied to the GPS signal at the user receiver.

7-8 Fig 4 The Principle of Differential GPS



25 Where a single reference site is used, the improved user accuracy will degrade with range from that reference station. DPGS systems can compensate for accuracy degradations over large areas by employing a network of co-ordinated reference receivers.

26. A list of Differential GPS (DGPS) reference stations, with their stated area of cover, is contained in RAF FLIPs. In addition to the permanent stations, DGPS reference stations can be portable, and thus transported to operational areas. Outside of DPGS cover user equipment usually reverts to working as a 'normal' GPS receiver.

Integrated Navigation Systems

27. Although extremely accurate, GPS should not, be regarded as a perfect stand-alone navigation system. The aviator who relies solely on GPS is putting himself in potential danger, therefore traditional cross-checking of position is essential.

GPS Geodetic Datum

28. The US GPS works on the WGS 84 geodetic datum, as mentioned in para 14. Users must be aware that plotting positions obtained from a GPS receiver on a map based on any coordinate system other than WGS 84 will introduce errors of varying magnitude (see Volume 9, Chapter 2).

Global Navigation Satellite Systems (GNSS)

29. ICAO defines a system that contains one or more satellite navigation systems as a global navigation satellite system (GNSS).

30. The US NAVSTAR system remains the most important and reliable GNSS, and improvements have been ongoing since full operational capability was declared in 1995. The system is now capable of much improved accuracies due to better system monitoring of both clocks and ephemeris data. Moreover, GPS satellites are deliberately positioned in the constellation to give improved accuracy in desired locations. PPS accuracies of 3m are now achievable. Table 1 shows revised accuracy figures for the SPS service, as at 2001.

Table 1 SPS Accuracy (Revised 2001)

| | Global Average | Worst Site |
|------------|---------------------------------|-------------------|
| | 2 Sigma Accuracy (m) | |
| Horizontal | 5 | 15 |
| Vertical | 8 | 26 |
| | 2 Sigma Availability (%) | |
| Horizontal | 99.5% at 15m | 92% at 15m |
| Vertical | 99.5% at 26m | 92% at 26m |
| | 2 Sigma Time (ns) | |
| | Not specified | 20 |

31. Russia has a satellite navigation system, Global'naya Navigatsionnaya Sputnikovaya Sistema (GLONASS). GLONASS is similar in concept to the US GPS, but differs in several major technical

aspects. Furthermore, whereas GPS works on the WGS 84 datum, GLONASS works on the PZ-90 datum. The difference between the two geodetic datum systems must be considered (typically, 5 to 15 metres).

32. The GPS and GLONASS systems can be used in combination, with suitable dual-capable receivers, to give the user an improved satellite availability. This first generation of satellite navigation systems, along with performance enhancements such as DGPS, is sometimes referred to as GNSS-1.

33. A planned European augmentation system for GNSS-1, known as the European Geostationary Navigation Overlay Service (EGNOS), will provide differential positioning over wide areas. In addition, the European Commission is working towards the development of a second generation regional GNSS, called Galileo.

CHAPTER 9 - PRINCIPLES OF INERTIAL NAVIGATION

Introduction

1. In an inertial navigation system (INS), velocity and position are obtained by continuously measuring and integrating vehicle acceleration. Inertial navigation systems are self-contained and are capable of all-weather operation. In early systems, the sensors were mounted on gyro-stabilised platforms. Modern systems use 'strapdown' technology where the sensors are mounted directly to the airframe and corrections applied digitally. Both systems are discussed in this chapter.

BASIC PRINCIPLES

Acceleration

2. The basis of inertial navigation is the measurement of a vehicle's (aircraft's) acceleration along known directions. Accelerometers detect and measure accelerations along their sensitive axes (input axes). The accelerometer outputs are integrated, once to obtain velocity along the sensitive axis, and again to obtain distance travelled along the sensitive axis.

3. **Relationship between Acceleration, Velocity and Distance.** The velocity achieved and the distance travelled by a vehicle accelerating from rest at a constant rate are obtained from the following equations:

$$v = at, \text{ and } s = \frac{1}{2} at^2$$

where, a = acceleration

v = velocity

s = distance

t = time

Aircraft accelerations are not constant, and must be integrated to obtain velocity and distance:

$$v = \int a \cdot dt, \text{ and}$$

$$s = \int \int a \cdot dt, \text{ or } s = \int v \cdot dt$$

The basic principle of inertial navigation is, therefore, the double integration of acceleration with respect to time (Fig 1).

7-9 Fig 1 Principle of Inertial Navigation



4. **Measurement Axes.** Acceleration must be measured along two axes, usually orthogonal, if vehicle velocity and displacement are to be defined in a given plane. Since most accelerometers are designed to measure acceleration along one axis only, two accelerometers are required for inertial navigation in a two dimensional plane. In aircraft systems, the accelerometers are usually mounted with their input axes aligned with North and East, and this alignment must be maintained if the correct accelerations are to be measured. Moreover, the sensitive axes must be kept perpendicular to the gravity vertical, otherwise, the accelerometers sense part of the gravity acceleration. The reference frame defined by these directions, i.e. local North, local East and local Vertical, is called the Local

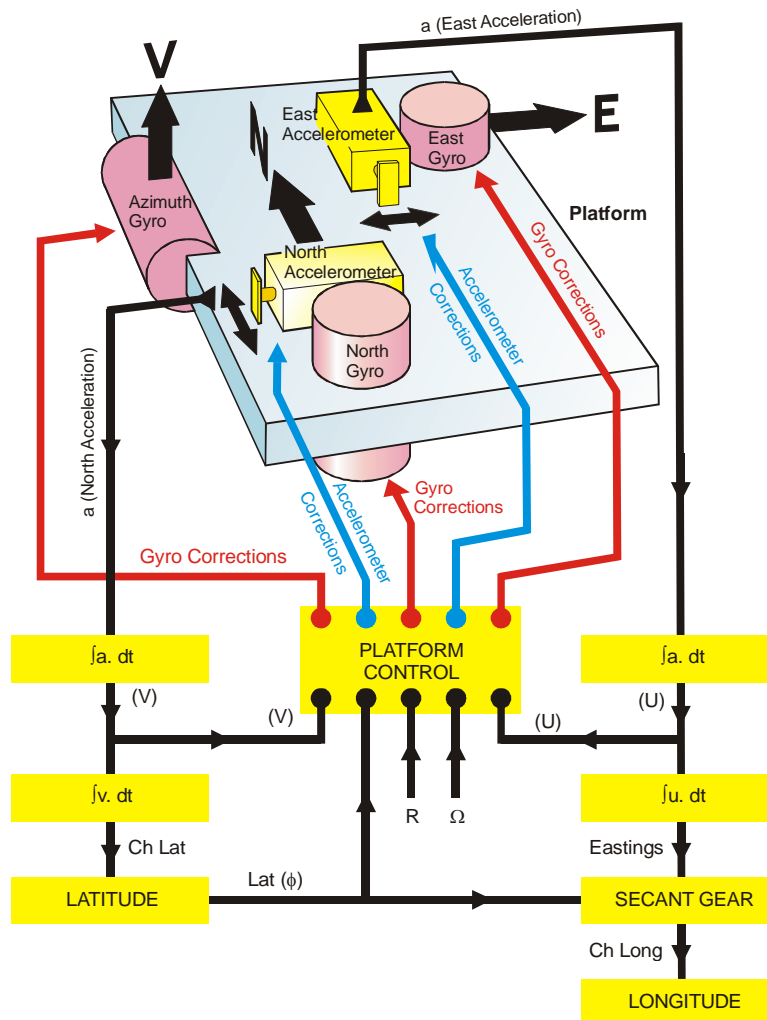
Vertical Reference Frame. Other reference frames can be used, but the local vertical is the fundamental mechanization and is the one primarily considered in this chapter.

5. **Gyro Stabilization.** Once the accelerometers have been aligned in the chosen reference frame, they must be capable of maintaining that orientation during aircraft manoeuvres. The accelerometers are therefore mounted on a platform which is suspended in a gimbal system that isolates the accelerometers from aircraft manoeuvres. However, this platform is not inherently stable, and any tendency for the platform to rotate with the aircraft must be detected and opposed. Gyros are therefore mounted on the platform to detect platform rotation and control platform attitude. Three single degree of freedom gyros are normally used; one gyro detects rotation about the North axis, another rotation about East, and the third rotation about the vertical. The platform rotations detected by the gyros are used to generate error signals, proportional to change in platform attitude, which are used to motor the platform back to its correct orientation.

6. **Effect of Earth Rotation and Vehicle Movement.** An INS operating in the local vertical reference frame must maintain its alignment relative to Earth directions. The gyros used to stabilize the platform are rigid in space and must therefore be corrected for Earth rate and transport wander to make them 'Earth stable'. Additionally, the accelerometers must be corrected for the effects of coriolis acceleration and the central acceleration caused by flying around a rotating spherical Earth.

7. **Platform Control.** The platform control unit computes and applies the gyro and accelerometer correction terms from calculated values of groundspeed and latitude, and stored values of Earth radius and Earth rotation rate.

8. **Simple INS.** A simple INS, capable of solving the navigation problem, is illustrated in Fig 2. A third vertically mounted accelerometer must be added if vertical velocity is required, eg in weapon aiming applications. Conventionally, velocity North is annotated 'V', velocity East 'U', Latitude ' ϕ ', Earth rate ' Ω ' and the radius of the Earth 'R'. Other annotations are self-explanatory and the individual INS components are discussed in detail in the following paragraphs.

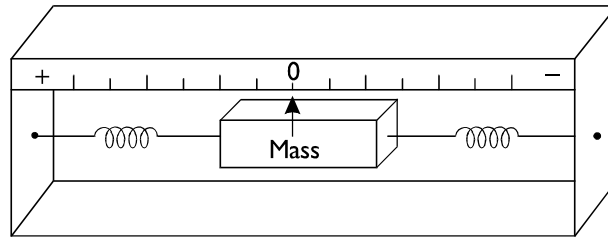
7-9 Fig 2 A Simple Inertial Navigation System

ACCELEROMETERS

Basic Principles

9. The accelerometer is the fundamental component of an INS. Its function is to sense acceleration (a) along its input axis and to provide an electrical output proportional to sensed acceleration. The spring and mass arrangement illustrated in Fig 3 shows the basic principles. If the instrument is accelerated along its longer axis, the mass will move relative to its neutral position until the spring tension balances the displacing force. The deflection of the mass is proportional to the acceleration. A pick-off system could be arranged to provide an electrical output that was the analogue of the acceleration.

10. An inertial grade accelerometer requires a wide dynamic range (typically ± 25 g), a high sensitivity (typically 1×10^{-6} g), and a linear response. These requirements cannot be accommodated in the simple accelerometer of Fig 3. High sensitivity could be achieved by the use of weak springs, but this would necessitate long springs to achieve the required range and the resulting instrument would be too large for use in a practical INS. Alternatively, strong springs could be used to achieve a wide range but this would preclude high sensitivity.

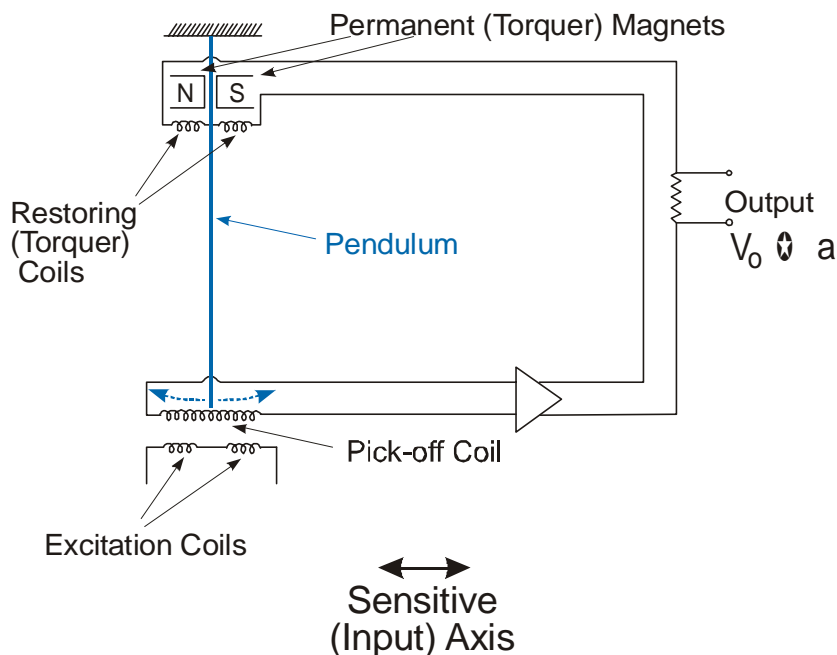
7-9 Fig 3 Simple Spring and Mass Accelerometer

11. A number of accelerometer designs have been developed to overcome these shortcomings. Some of these designs are more suitable for applications other than aircraft, e.g. ballistic missile systems. In aircraft applications, the 'Pendulous Force Balance Accelerometer' is the most common type.

Pendulous Force Balance Accelerometer

12. A basic pendulous force balance accelerometer is shown schematically in Fig 4. With the case horizontal and the instrument at rest or moving at a constant velocity, the pendulous mass is central and no pick-off current flows. When the instrument is accelerated along its sensitive axis the pendulous mass is deflected and the deflection is sensed by the pick-off. A current flows through the restorer coils such that a force is exerted on the displaced mass to restore it to the central position. The initial deflecting force is proportional to the acceleration experienced since the mass is constant ($F = ma$). The restoring force is proportional to the current through the restorer coil and is equal and opposite to the initial force, ie the restorer current is proportional to the acceleration. The pendulous mass is free to move only along the sensitive axis and accelerations perpendicular to this axis have no effect.

13. Instead of the flexure support system, the pendulous element may be floated. The element has its centre of mass displaced from the centre of buoyancy, thus producing a couple in the presence of a linear acceleration.

7-9 Fig 4 Basic Pendulous Force Balance Accelerometer

Pendulous Accelerometer Errors

14. **Cross-coupling.** The accelerometer is sensitive to accelerations along an axis perpendicular to, and in the plane of, the pendulum. If the pendulum is displaced from the null position, either by an

acceleration or by tilting of the platform, then the sensitive axis no longer coincides with the designed fixed input axis. Fig 5 shows the situation where the pendulum has been displaced through a small angle, θ . The input (sensitive axis) IA is rotated through the same angle. If the instrument is accelerated along the displaced axis, the acceleration will have horizontal and vertical components a_x and a_y and the measured acceleration will be:

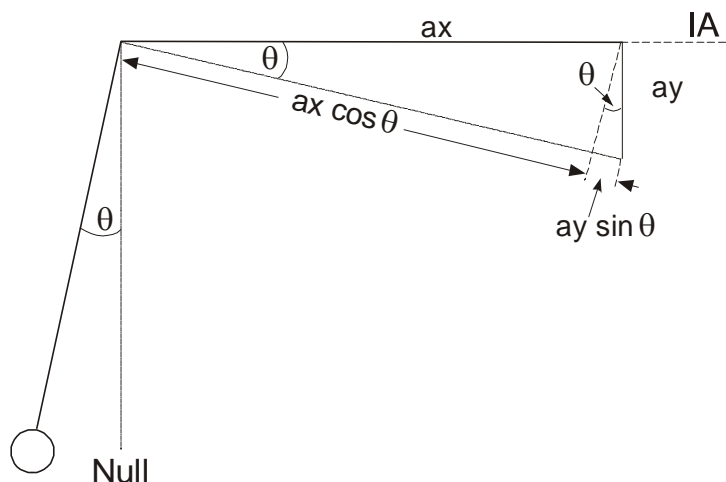
$$a_x \cos \theta + a_y \sin \theta$$

If θ is small and measured in radians this becomes:

$$a_x + a_y \theta$$

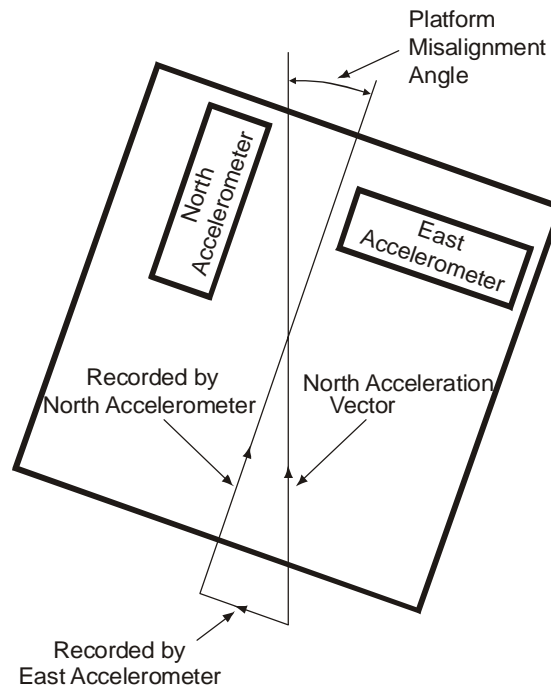
The acceleration that should have been measured is a_x and the term $a_y \theta$ is an error known as cross-coupling error. It should be noted that when the input axis is displaced from the horizontal it will sense a component of the acceleration due to gravity. Cross-coupling error can be minimized by ensuring that the accelerometer platform is maintained horizontal and by using a high gain feedback loop so that the displacement of the pendulum due to accelerations is kept small. Alternatively, the error ($a_y \theta$) can be calculated and corrected.

7-9 Fig 5 Cross-coupling Error



15. **Vibropendulosity.** When an accelerometer is operated in a vibration environment, below the natural frequency of the accelerometer loop, components of the vibration may act along the input axis causing the pendulum to deflect and thus register erroneous accelerations.

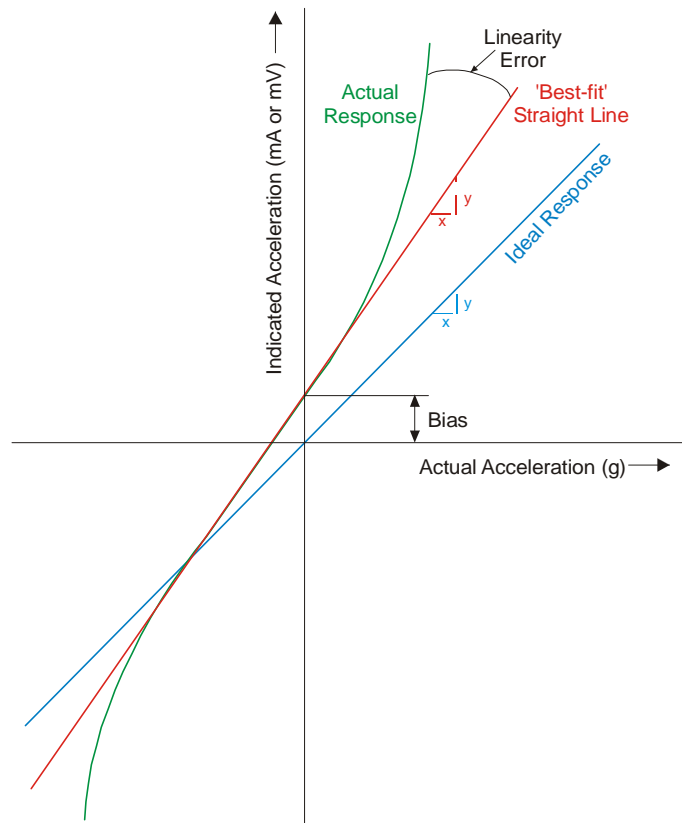
16. **Displaced Orientation.** Accelerometers are arranged, mutually at right angles, to measure accelerations in specific directions, normally North and East. If the platform is misaligned as in Fig 6 and accelerated in a North/South direction, the north sensitive accelerometer will not detect the full acceleration and the east accelerometer will detect an unwanted component.

7-9 Fig 6 Accelerometer Misaligned

Performance Characteristics

17. Accelerometers are required to give an accurate indication of vehicle acceleration over a wide range of accelerations. Fig 7 shows a typical accelerometer response, and indicates the performance parameters usually referred to in technical descriptions. The straight blue line represents the ideal response from the instrument, i.e. a linear relationship between the actual acceleration and the indicated acceleration. The green curved line represents the actual response; the red line is the 'best-fit' straight line to this curve.

- a. **Bias.** Bias is the electrical output under conditions of no acceleration input due to residual internal forces acting on the mass after it has been electrically or mechanically zeroed. Bias is expressed as an equivalent error in 'g's and is usually less than $1 \times 10^{-4}g$. An INS can be designed to compensate for known accelerometer bias provided the bias is stable.
- b. **Scale Factor.** Scale factor is the ratio of the change in output to a unit change in input (in units of acceleration). It is usually expressed in millivolts or milliamps per g (mV/g or mA/g). In Fig 7, scale factor is the slope or gradient of the straight lines (y/x). In some technical literature, scale factor is referred to as 'sensitivity', although some manufacturers use sensitivity to mean change in output as a result of secondary or unwanted inputs.
- c. **Linearity.** Linearity error is defined as the deviation from the best-fit straight line drawn through a plot of the electrical output in response to a known acceleration input. Typically this would be in the order of $5 \times 10^{-5}g$ up to 1g and less than 0.01% of applied acceleration at higher g levels.
- d. **Threshold.** Threshold is the minimum acceleration input which causes an accelerometer electrical output. It is equivalent to scale factor (or sensitivity) but with an incremental change about a zero input. A typical value is $1 \times 10^{-6}g$.
- e. **Null (Zero) Uncertainty.** Null uncertainty is also known as bias uncertainty and is the variation in accelerometer output under conditions of zero acceleration input. The random drift of the accelerometer output at zero acceleration input is known as null (zero) stability.

7-9 Fig 7 Accelerometer Performance Parameters

INTEGRATORS

Function of the Integrator

18. The accelerometer outputs are integrated to obtain velocity and again to obtain distance. The initial integration may be carried out within the accelerometer or by a separate integrating device. The accelerometer output may be in voltage analogue form if analogue techniques are used, or pulse form if digital techniques are used.

Analogue Integrators

19. Analogue integrators are normally electronic or electro-mechanical. Electronic integrators are more accurate, but are capable of integrating continuously for only limited periods of time. The electro-mechanical integrators are less accurate, but can integrate indefinitely.

Digital Integrators

20. Many inertial systems use digital computers and therefore digital integration techniques. A digital computer integrates by adding small increments of the quantity to be integrated. As the computer will be dealing with discrete quantities instead of continuous values, there will be a certain amount of approximation in the integration process.

GYROSCOPES

Terms

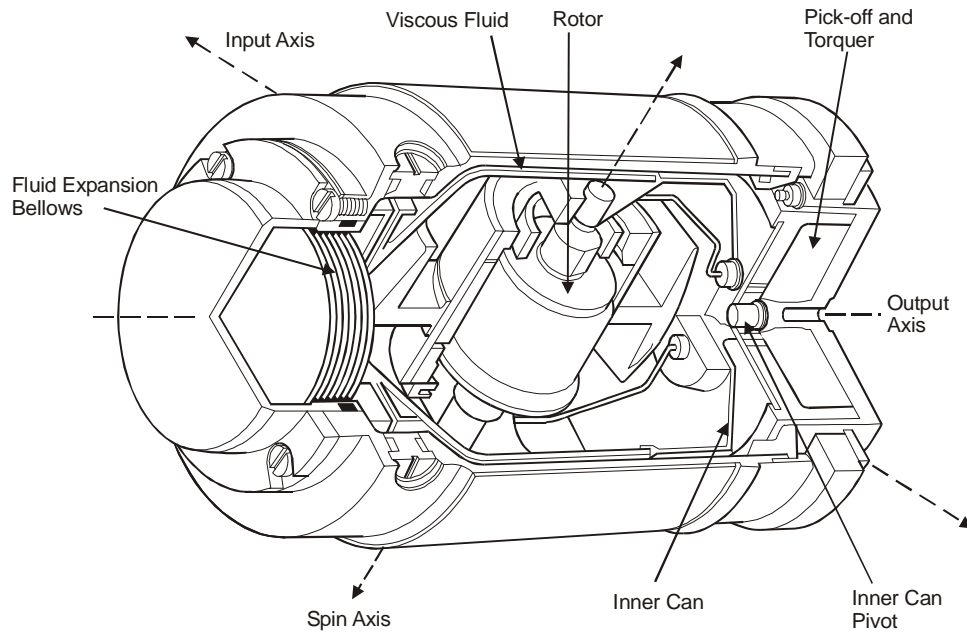
21. The following terms are included for clarification:

- a. **Degrees of Freedom.** In the convention used throughout this chapter, the gyro rotor axis is not counted as a degree of freedom, since it cannot be a sensitive axis. A free or space gyro is therefore defined as a two degree of freedom gyro.
- b. **Gyro Drift.** The term gyro drift describes any movement of the gyro spin axis away from its datum direction.
- c. **Levelling Gyros.** Gyros which control the platform about the horizontal axes are called levelling or vertical gyros, irrespective of the direction of their spin axes.

22. **Inertial Quality.** A gyro is described as being of inertial quality when the real drift rate is 0.01° per hour or less. Such low drift rates were first achieved with single degree of freedom rate-integrating gyros.

Single Degree of Freedom (SDF) Gyros

23. **Rate-integrating Gyro.** The rate-integrating gyro achieves its accuracy by reducing gimbal friction: the gimbal and rotor assemblies are floated in a fluid. A typical floated rate-integrating gyro is illustrated in Fig 8; the rotor is pivoted in an inner can (gimbal), which in turn is floated in an outer can. The outer can contains all the controls, pick-offs, torquers and heaters, etc. Rotation of the gyro about the input (sensitive) axis causes the gyro inner can to precess about the output axis, ie there is relative motion between the inner and outer cans. This precession is sensed by the pick-offs which measure the angular displacement of the inner can relative to the outer can. Thus, the pick-off output is proportional to the time integral of the input turning rate. This output signal is used to drive the platform gimbals to maintain the platform in the required orientation. The ratio of output to input (gimbal gain) is a function of rotor mass, gimbal size and fluid viscosity. A high ratio enables the gyro to detect small input rates. However, the fluid viscosity varies with temperature. Temperature must therefore be controlled to ensure a constant gimbal gain. With this type of gyro, it is also important to limit the inner can precession; as the inner can precesses, the rotor and the input axes are also precessed. Unless this precession is rapidly detected and opposed (the gimbal drives the platform and the gyro in opposition to the input), cross-coupling errors are likely to occur. A cross-coupling error is caused by the gyro sensing a rotation about a displaced input axis.

7-9 Fig 8 Typical SDF Floated Rate-integrating Gyro

Two Degree of Freedom (TDF) Gyros

24. Two degree of freedom gyros are used in some IN applications. SDF and TDF gyros have comparable performances, but the TDF gyro has the advantage of being able to detect movement about two axes. Since the INS monitors motion about three axes, two TDF gyros are not only sufficient, but also supply a redundant axis; the spare axis is normally utilized to monitor azimuth. The two TDF gyros must have their spin axes at right angles to each other; both axes may be horizontal, or alternatively one horizontal and the other vertical.

Comparison of Single and Two Degree of Freedom Gyros

25. The single and two degree of freedom gyros are compared in Table 1.

7-9 Table 1 Comparison of SDF and TDF Gyros

| <i>Property</i> | <i>SDF</i> | <i>TDF</i> |
|---------------------------------------|--|---|
| Number required in IN platform | Three | Two (one redundant axis) |
| Gyro gain | Normally controlled by fluid viscosity | Output = input |
| Cross-coupling | Limited rotor axis movement minimizes cross-coupling | No cross-coupling - angular displacement is measured against fixed input axis |
| Vehicle movement detection capability | Detected by rotor axis movement | Detected by gimbal axis movement |
| Accuracy | 0.003°/hr to 0.1°/hr | As for SDF |

Pick-offs and Torquers

26. Angular movement about a gyro's sensitive axis is detected by pick-offs which generate electrical signals proportional to the movement. The action of the torquers is virtually the reverse; electrical signals proportional to the desired correcting torque are applied to the torquers which cause the gyro to precess at the desired rate. The pick-offs and torquers are usually of the induction type, and may be separate devices or combined in a single unit; in the latter type, the pick-off would use AC and the torquer DC to avoid interaction between the fields.

PLATFORM STABILIZATION

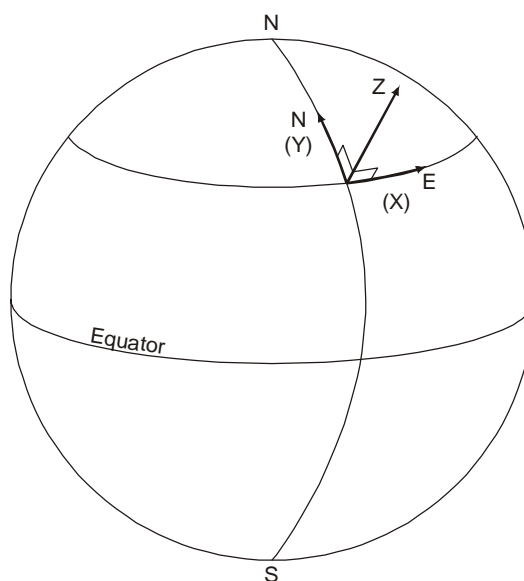
Gyro Control of the Platform

27. **Platform Mounted Accelerometers.** The accelerometers are mounted on a platform which is oriented to a fixed reference frame. The platform is aligned with the desired reference frame and subsequently controlled to maintain its alignment.

28. **Choice of Reference Axes.** A fundamental aircraft INS is aligned in the local vertical reference frame, the axes of which are shown in Fig 9. Basic stabilization procedures are described for this simple system, but in practice, RAF aircraft with a stable platform INS use a modified local vertical reference frame known as a Wander Azimuth System. These systems allow the azimuth gyro to wander, and the IN computer continually transforms position in the wander azimuth frame to the required Earth-fixed co-ordinates.

29. **Platform Alignment.** Inertial platforms are aligned in attitude and azimuth. Any platform misalignment will cause errors.

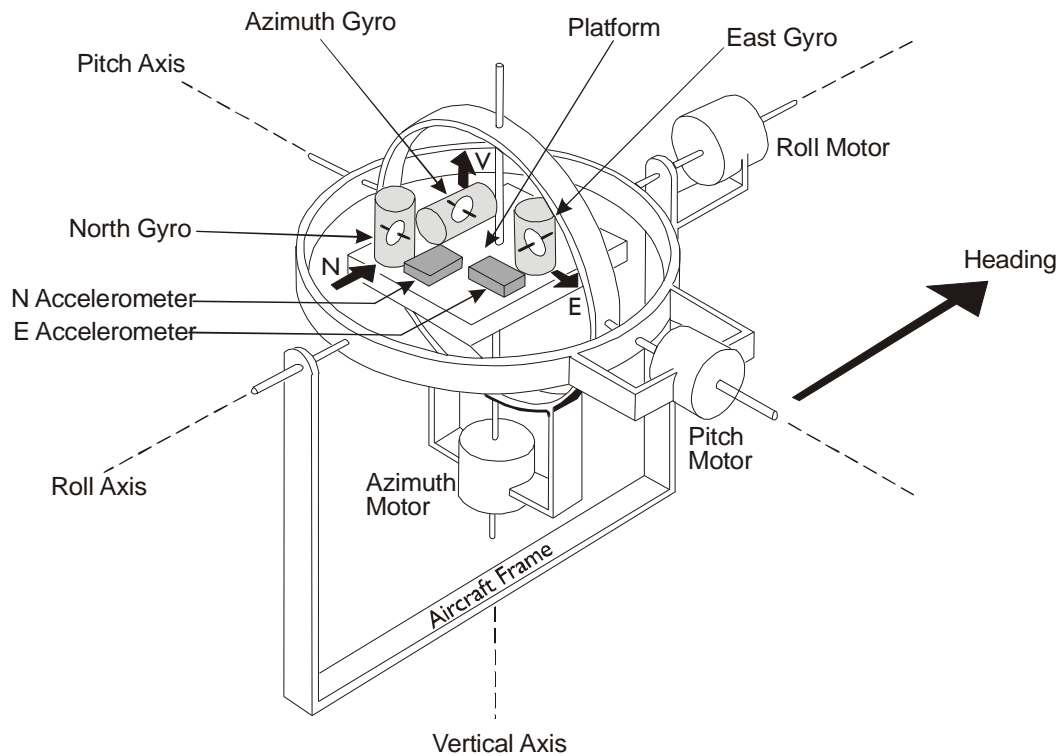
7-9 Fig 9 Local Vertical Reference Frame



30. **Use of Gyros.** The desired platform orientation is maintained by mounting reference gyros on the platform to detect changes in platform alignment. The gyro outputs are used to drive gimbal motors which return the platform to its correct orientation.

31. **Platform Arrangement.** The platform may be arranged as shown in Fig 10. The three gyros have their input axes mutually at right angles and aligned with the local vertical reference frame. The error pick-offs and torquers are built into the gyro cases and are not shown in the diagram. The platform is gimbal mounted to permit the aircraft freedom of manoeuvre without disturbing the platform away from its alignment with the local vertical reference frame. Each gimbal is driven by a servomotor controlled by the error signals from the gyros.

7-9 Fig 10 Platform Arrangement (Aircraft Heading North)



32. **Control on North.** The gyros in Fig 10 are arranged with their sensitive axes pointing in the directions about which rotation is to be detected. The East gyro has its sensitive axis pointing East, and is therefore capable of detecting rotation about East. On northerly headings, pitch manoeuvres are detected by the East gyro which generates an error signal. This error signal activates the pitch gimbal, thereby maintaining the platform's alignment with the reference frame. Similarly, roll is detected by the North gyro, and yaw by the azimuth gyro: the North gyro activates the roll gimbal motor, and the azimuth gyro the yaw gimbal motor. The action is summarized in Table 2.

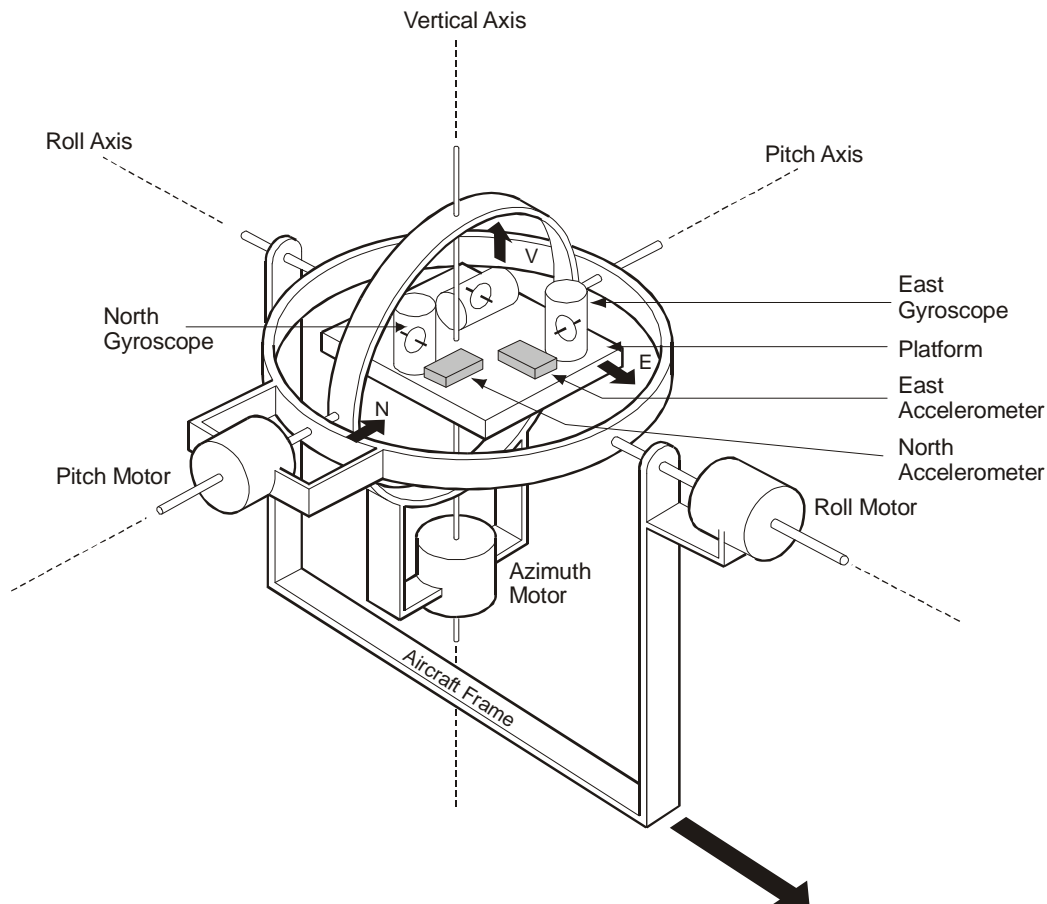
7-9 Table 2 Action on North

| <i>Heading</i> | <i>Manoeuvre</i> | <i>Sensing Gyro</i> | <i>Correcting Servomotor</i> |
|----------------|------------------|---------------------|------------------------------|
| North | Yaw | Azimuth | Azimuth |
| | Pitch | East | Pitch |
| | Roll | North | Roll |

33. **Control on East.** In Fig 11, the same platform is again shown, this time heading East. The action on East is summarized in Table 3.

7-9 Table 3 Action on East

| <i>Heading</i> | <i>Manoeuvre</i> | <i>Sensing Gyro</i> | <i>Correcting Servomotor</i> |
|----------------|------------------|---------------------|------------------------------|
| East | Yaw | Azimuth | Azimuth |
| | Pitch | North | Pitch |
| | Roll | East | Roll |

7-9 Fig 11 Platform Arrangement (Aircraft Heading East)

34. **Conclusions.** Two main conclusions may be drawn from Tables 2 and 3:

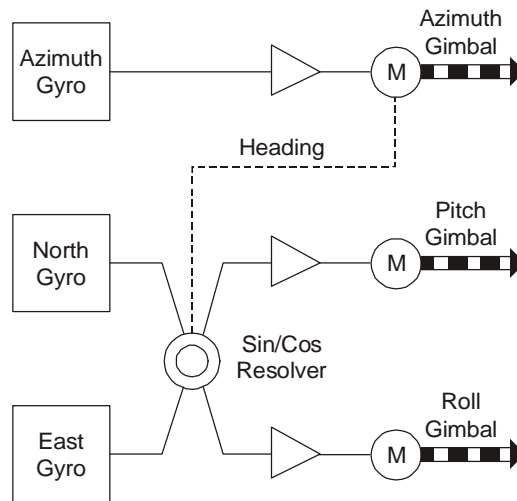
- Yaw, or change of heading, is corrected by the azimuth servomotor which is always controlled by the azimuth gyro.
- Pitch and roll are corrected by the pitch and roll servomotors respectively. However, the control may be exercised by either the North or the East gyros or both, dependent upon aircraft heading.

35. **Change of Heading.** The action of the azimuth gyro and servomotor keeps the platform aligned with the North datum. However, the pitch and roll gimbals remain oriented to the aircraft pitch and roll axes (Figs 10 and 11). Relative motion about the vertical between the platform and the pitch and roll gimbals is yaw, and angular displacement is change of heading. A pick-off of the angular displacement relative to true North as defined by the platform, produces an output of heading.

36. **Control During Manoeuvres.** On northerly headings, the North gyro senses roll and wholly controls the roll servomotor; on easterly headings the East gyro controls the roll servomotor. On

intermediate headings, the control is shared between the North and East gyros, the amount of control exercised being determined by the heading. A sine-cosine resolver, set by the azimuth servomotor, determines the amount of control and transmits the error signal to the appropriate servomotor. The action is shown in Fig 12.

7-9 Fig 12 Gimbal Control Signals



PLATFORM MOUNTING

Gimballed Systems

37. The stable element of the inertial platform is mounted in gimbals to isolate the platform from vehicle manoeuvres. The two types of gimbal system in common use are the three-gimbal system and the four-gimbal system.

38. **Three-gimbal System.** Figs 10 and 11 are diagrams of a three-gimbal system. In such a system, there are three input/output axes, azimuth, pitch and roll. Each gimbal imparts freedom about one particular axis, the particular gimbal being named after that axis.

a. **The Gimbals.**

(1) **Azimuth Gimbal.** The stable element is rigidly attached to the azimuth, or first, gimbal. In allowing relative motion between the stable element and the pitch gimbal, the platform is isolated from vehicle movement about the vertical axis.

(2) **Pitch Gimbal.** The pitch gimbal isolates the platform from pitch manoeuvres.

(3) **Roll Gimbal.** The roll gimbal isolates the platform from roll manoeuvres.

In some installations, the pitch and roll gimbals are reversed in order of position.

b. **Gimbal Lock.** Gimbal lock occurs when two axes of rotation become co-linear and, as a result, one degree of freedom is lost. Fig 13 illustrates how gimbal lock can occur in a three-gimbal system. If the vehicle pitches through 90° the first and third gimbal axes become coincident, and the platform stable element is no longer isolated from yaw.

7-9 Fig 13 Gimbal Lock

Fig 13a Straight and Level Flight

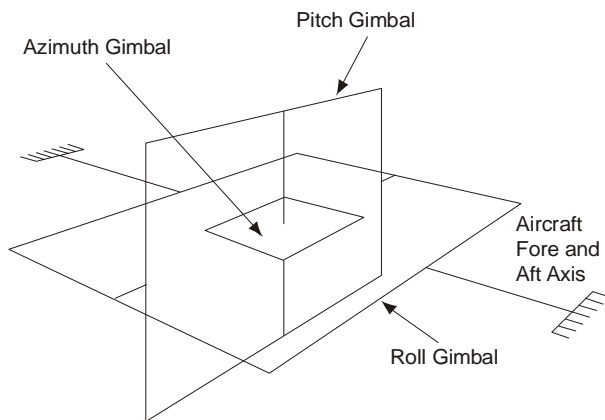
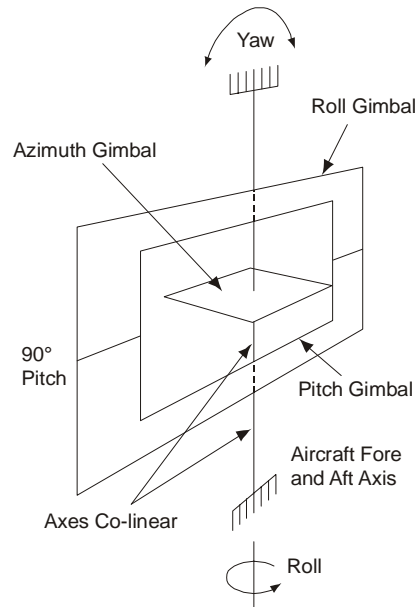


Fig 13b Pitch and Roll Gimbals Coincident



c. **Gimbal Error.** In a three-gimbal system (gimbal order; Azimuth, Pitch and Roll) the roll gimbal axis, which is parallel to the aircraft roll axis, assumes an angle relative to the plane of the platform stable element whenever the aircraft pitches through large angles. When this occurs, the gimbal roll axis and the plane of the levelling gyros' input axes are no longer parallel. Should the aircraft now roll, the gyros sense only a component of roll angle ($\text{roll} \times \cos \text{pitch angle}$), and the roll servo displaces the roll gimbal by an amount ($\text{roll} \times \cos \text{pitch angle}$) instead of the full value of roll angle.

39. **Four-gimbal System.** In a four-gimbal system, the order of the gimbals is azimuth, inner roll, pitch and outer roll.

a. **Gimbals.** The fourth gimbal is introduced to keep the second and third gimbals at right angles, thereby avoiding both gimbal lock and gimbal error. The fourth gimbal is controlled by a pick-off which detects changes in the angle between the second and third gimbals.

b. **Gimbal Flip.** With a four-gimbal system, heading change is picked off from the relative motion between the azimuth and inner roll gimbals. If, however, the aircraft completes a half loop and roll-out manoeuvre, the aircraft heading changes by 180° but there is no motion between the azimuth and inner roll gimbals, and the indicated heading remains unchanged. This problem is overcome by employing gimbal flip. As the pitch angle passes through 90° , the outer gimbal is driven through 180° (ie flips), tending to drive the platform through 180° about the vertical. This tendency is detected by the azimuth gyro which provides an appropriate output signal. This signal keeps the platform correctly orientated by driving the platform in opposition to the flip. One hundred and eighty degree relative motion is produced between the azimuth and inner roll gimbals and the heading output remains correct.

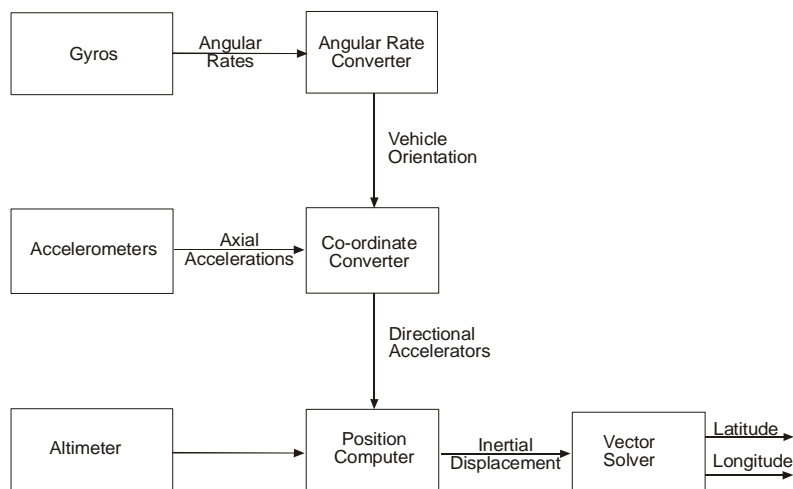
40. **Comparison of Three- and Four-gimbal Systems.** A four-gimbal system is heavier, larger and costs more than a three-gimbal system. However, since the second and the third gimbals of the four-gimbal system are kept at right angles, the aircraft has full freedom of manoeuvre without disturbing the platform.

Non-gimballed Systems - Strapdown Systems

41. In gimballed systems, the accelerometers are mounted on a stable platform which is kept in the correct orientation by torqueing in response to signals from the gyroscopes detecting movement about three orthogonal axes.

42. In a strapdown system (Fig 14), the inertial sensors are fixed to the vehicle and their orientation within the navigation reference frame is computed using the outputs of gyroscopes which detect angular displacement about the aircraft axes. Thus in a strapdown system the gimbals are effectively replaced by a computer. Although a strapdown mechanization is more demanding technically in terms of computing and gyroscope performance, it is potentially cheaper, more reliable and more rugged than a gimballed system.

7-9 Fig 14 Strapdown System Block Diagram



43. **Gyroscopes.** In a strapdown system, the function of the gyroscope is to measure accurately angular changes about a specific axis of rotation. This requires a very wide range of performance, as the gyros may well need the capability to detect rotation rates ranging from $0.001^\circ/\text{hr}$ to $400^\circ/\text{sec}$. Although conventional gyros could be used for strapdown applications, gyros with no moving parts are better suited

44. **Computing Requirements.** The main computing task in a strapdown system is to compute the instantaneous aircraft attitude and to resolve and integrate the accelerometer outputs to obtain velocity information in a useful geographic reference frame. These calculations need to be carried out at very high speed and accuracy. Whereas in a gimballed system the platform reference frame rotates relatively slowly due to transport wander and Earth rate, in a strapdown system the platform reference frame, i.e. the airframe, can be rotating at very high rates. The integration process must therefore be carried out very rapidly to avoid large errors being induced; an iteration rate of 200 Hz would be typical and a dedicated microprocessor may be required.

45. **Reference Frames.** The platform reference frame in a strapdown system is the same as the airframe and is therefore of no use for navigation. However, an advantage of this configuration is that outputs can be used for an automatic flight control system. Strapdown systems commonly use a space-referenced frame for the navigation solutions and then convert to a geographic frame to give the desired outputs of position and velocity. Fig 14 shows the functional layout of a typical system. The outputs of the accelerometers are resolved along the space axes and the cartesian co-ordinates of the aircraft position within the space frame calculated. These are then converted to the geographic frame to give latitude and longitude.

CORRECTIONS TO INERTIAL SENSORS

Introduction

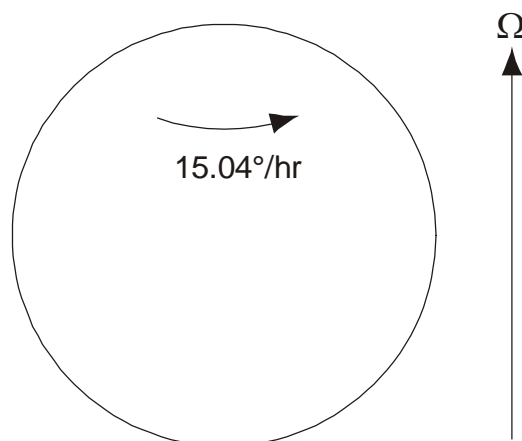
46. It is normal to navigate aircraft with reference to the local Earth co-ordinates of latitude, longitude and height. Aircraft INS are therefore normally aligned as described in para 29; each of the 3 axes of the local vertical reference frame has an accelerometer to detect movement along it and a gyro to provide stabilization against rotation around it. Accelerometers and gyros are both inertial devices in that their sensitive axes extend infinitely in straight lines; in other words, they operate with reference to the constant axes of inertial space. Local vertical axes however are not constant. For an aircraft system using local vertical Earth co-ordinates, it is therefore necessary to change the orientation of the platform axes relative to inertial space in order that the accelerometers are kept aligned with the local vertical axes. This means that the stabilizing effect of the gyros must be adjusted by the rates at which local vertical axes diverge from inertial axes. These rates are due to Earth rotation and vehicle movement as shown in Table 4. The changing orientation of the platform also makes corrections to the accelerometer outputs necessary.

7-9 Table 4 Platform Correction Terms

| | <i>Earth Rate</i> | <i>Vehicle Movement</i> |
|--------------|--------------------|-------------------------|
| North Gyro | $\Omega \cos \phi$ | $\frac{U}{R}$ |
| East Gyro | zero | $-\frac{V}{R}$ |
| Azimuth Gyro | $\Omega \sin \phi$ | $\frac{U}{R} \tan \phi$ |

47. It is now necessary to analyse Earth and transport rates into components affecting the local vertical axes. These are the rates which are applied to the platform's axes to correct it from inertial space stabilization to local vertical stabilization. The method used in the following discussion is that of vector analysis. A rate of rotation is represented by a vector shown parallel to the axis of the rotation. Its length is proportional to the rate of rotation and its direction is the direction an ordinary right hand threaded screw would move if subjected to the rotation in question. This is shown in Fig 15 which shows the Earth's rotation vector. The vector is parallel to the Earth's spin axis, its length represents 15.04°/hr (Ω) and its direction is from South to North.

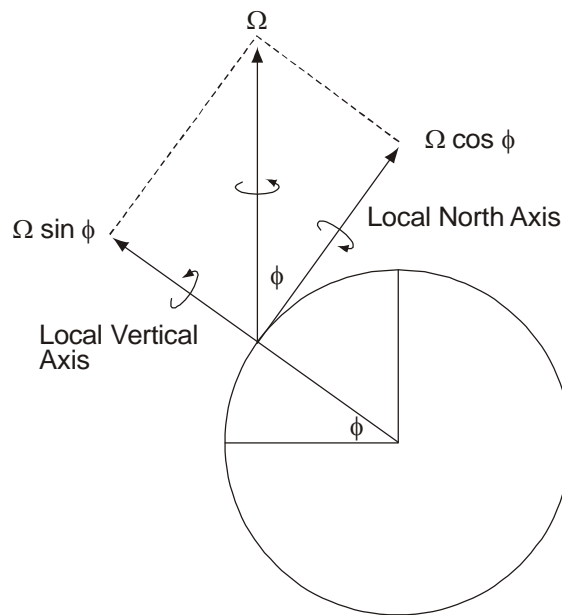
7-9 Fig 15 Earth's Rotation Vector



Gyro Corrections

48. **Earth Rate (Ω).** The Earth's rotation vector may be analysed into components acting about the local vertical axes at any point on the Earth's surface. The component acting about local East is always zero because local East is always at 90° to the rotation vector. At the poles, the rotation vector coincides with the local vertical axis, and at the equator it coincides with the local North axis. This means that an INS not corrected for Earth rotation will appear to drift, but not topple, at the pole; whereas at the Equator it will topple about local North but not drift. Fig 16 shows how the Earth rotation rate is resolved into vector components acting about local North and local vertical axes at intermediate latitudes.

7-9 Fig 16 Earth Rate Vector Components



49. **Transport Rates.** Fig 17 shows that any movement around the circumference of a circle equates to a rotation about the centre of the circle. The angle θ , in radians, is found by dividing the circumferential distance A-B by the radius of the circle. Similarly, the rate of rotation may be found by dividing the rate of movement from A to B by the radius. The axis of the rotation is perpendicular to both the radius and the tangent, ie, normal to the surface of the page.

7-9 Fig 17 Circular Movement

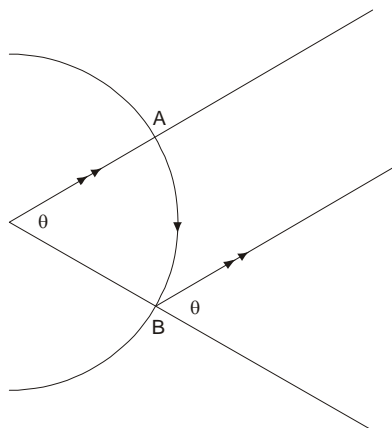
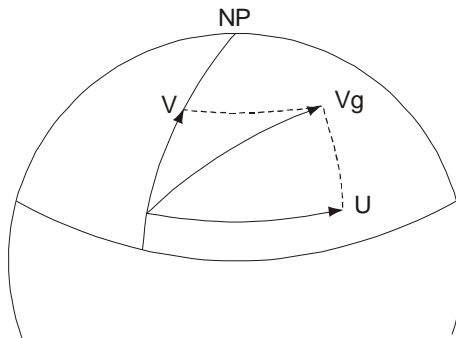


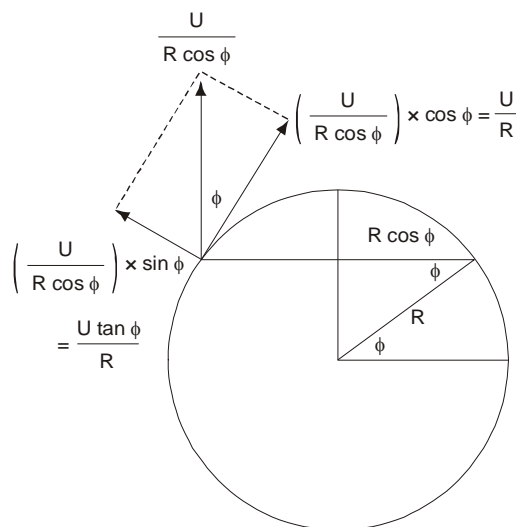
Fig 18 shows how a total aircraft velocity vector V_g may be resolved with North and East components. Component V produces a rotation rate of $\frac{V}{R}$ radians/hr about an axis parallel to the local East axis and through the centre of the Earth; (where V is in knots and R is the radius of the Earth in nautical miles).

7-9 Fig 18 Components of Velocity Vector



Component U , however, acts along a parallel of latitude, ie a small circle; U therefore produces a rotation rate of $\frac{U}{R \cos \phi}$ about the Earth's polar axis as shown in Fig 19.

7-9 Fig 19 Rotation Rates



This rate must be resolved into rates about the local North and local vertical axes, $\frac{U}{R}$ and $\frac{U \tan \phi}{R}$ respectively, before it can be applied to the IN platform. This is achieved using the same analysis by vectors as was used for Earth rate, for the axis of rotation is the same: the Earth's spin axis. The quantities arrived at by this analysis are in radians per hour; they may be approximated to degrees per hour by substituting 60 for R in the final expressions.

50. Correction Method. The drift due to the error rate is eliminated by applying an equal and opposite correction to the gyro output axis. The correction is applied through a torque motor on the gyro output axis, which turns the gyro about its output axis at the same rate but in the opposite direction to the precession caused by the error rate.

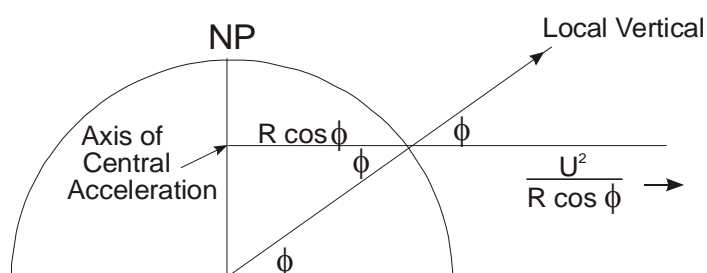
Accelerometer Corrections

51. Stabilizing a platform to local Earth axes requires that it be rotated relative to a spatial reference in order to compensate for the effects of Earth rotations and vehicle movement. The resulting change in the local axes relative to spatial references makes two types of accelerometer corrections necessary:

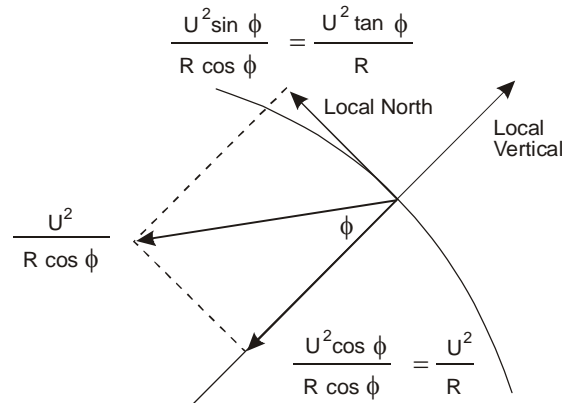
- a. **Central or Centripetal Acceleration.** A body moving at a constant speed v in a circle radius r has a constant acceleration of $\frac{v^2}{r}$ directed towards the centre of the circle. This is a central or centripetal acceleration and affects a local vertical INS because as the platform is transported over a spherical surface it is rotated to maintain its alignment with local North and the local vertical.
- b. **Coriolis Acceleration.** Coriolis acceleration results from the combination of aircraft velocity and the rotation of the Earth over which it flies. A lateral acceleration relative to inertial references is necessary to make good a desired track measured against meridians which are themselves in motion.

52. **Central Accelerations.** At any instant when an INS is moving over the Earth's surface it is moving along an arc of a great circle. An acceleration of $\frac{V_g^2}{R}$ therefore affects the vertical accelerometer where V_g is along track velocity and R is the radius of the Earth. Pythagoras' theorem enables us to convert this term to its component form $\frac{V^2 + U^2}{R}$ and thus make use of the 1st integrals of the North and East channels accelerations. This quantity as a correction must be added to the output of the vertical accelerometer. (See note at end of para 57). Central acceleration corrections must also be applied to the horizontal accelerometers because of meridian convergence. Any East component of velocity acts along a small circle of latitude whose radius is $R \cos \phi$. There is thus a central acceleration of $\frac{U^2}{R \cos \phi}$ along this radius, that is, along an axis inclined at ϕ to the local vertical. This is shown in Fig 20.

7-9 Fig 20 Axis of Central Acceleration



53. **Resolution of Total Acceleration.** Because of this inclination, the total acceleration may be resolved by vector analysis into two components, one affecting the North accelerometer and the other the vertical accelerometer. These are shown in Fig 21. The component $\frac{U^2}{R}$ is contained within the vertical accelerometer correction already discussed. The acceleration component $\frac{U^2 \tan \phi}{R}$ however, must be subtracted from the output of the North accelerometer because it is caused entirely by an Eastward motion. This apparent contradiction arises because while 'East' is a constant direction in terms of navigation over the surface of the Earth, it is a direction which constantly changes with respect to the fixed axes of inertial space. We thus have an Eastward velocity component producing an output from the North accelerometer; this must be removed for purposes of navigation.

7-9 Fig 21 Components of Central Acceleration

54. **Application.** Now consider an aircraft flying a great circle track at a constant groundspeed. The track angle is constantly changing and, therefore, so are U and the North accelerometer correction. In order that a constant total velocity vector results, the output of the East accelerometer must be adjusted in inverse proportion to the North accelerometer correction. The horizontal accelerometer central corrections thus produce varying V and U components of total velocity as track angle changes relative to the converging meridians. Table 5 shows that if there is no East component of velocity there is no central correction to either horizontal axis. If there is no North component, only the North accelerometer correction is applied, as discussed earlier. In addition, the magnitude of the corrections to the horizontal accelerometers increases as latitude increases, ie as meridian convergence increases.

55. **Coriolis Acceleration.** An aircraft flying a constant track over a spherical rotating Earth follows a path which is curved relative to the constant axes of inertial space. As shown in para 46, there is a component of Earth rotation which acts about the local vertical axis. This component, $\Omega \sin \phi$, varies with latitude. An observer may thus be regarded as being at the centre of a rotating disc of Earth's surface; the direction of rotation being anti-clockwise when viewed from above, in the Northern hemisphere. An aircraft flying towards a given point on the horizon is therefore flying to a destination which is moving constantly to the left. A straight track over the ground thus produces a track which is curved relative to a constant spatial direction; this can only be achieved if there is a sideways acceleration. This acceleration is the Coriolis effect and is detected by the horizontal accelerometers. It must, however, be removed if the system is to produce navigation information which is correct relative to Earth co-ordinates. A similar correction is applied to the vertical accelerometer because of the component of Earth rotation acting about the local North horizontal axis. The corrections are given below:

- $-2\Omega U \sin \phi$ applied to the North accelerometer.
- $2\Omega V \sin \phi$ applied to the East accelerometer.
- $2\Omega U \cos \phi$ applied to the Vertical accelerometer.

56. **Gravity Corrections.** When a third accelerometer is used in the vertical channel to measure vertical acceleration for weapon aiming purposes its sensitive axis will necessarily be in line with the gravity vector; the accelerometer will sense the acceleration due to gravity as well as aircraft vertical acceleration. Its output must therefore, be corrected for gravity, in addition to coriolis, and centripetal accelerations. Because the gravity acceleration decreases as the distance from the centre of the Earth increases, the correction is dependent on aircraft altitude. The correction is given by:

$$g_0 \left\{ \frac{2h}{R} - 1 \right\}$$

where g_0 is the standard gravity at the surface of the Earth and h is the aircraft altitude.

Summary

57. The gyro and accelerometer correction terms are summarized in Table 5.

7-9 Table 5 Gyro and Accelerometer Terms

| <i>Axis</i> | <i>Gyros</i> | | <i>Accelerometers</i> | | |
|----------------------|--------------------|-------------------------|----------------------------|------------------------|---|
| | <i>Earth Rate</i> | <i>Transport Wander</i> | <i>Central</i> | <i>Coriolis</i> | <i>Gravity</i> |
| North | $\Omega \cos \phi$ | $\frac{U}{R}$ | $\frac{-U^2 \tan \phi}{R}$ | $-2\Omega U \sin \phi$ | nil |
| East | nil | $-\frac{V}{R}$ | $\frac{UV \tan \phi}{R}$ | $2\Omega V \sin \phi$ | nil |
| Azimuth/ Vertical | $\Omega \sin \phi$ | $\frac{U}{R} \tan \phi$ | $\frac{U^2 + V^2}{R}$ | $2\Omega U \cos \phi$ | $g_o \left\{ \frac{2h}{R} - 1 \right\}$ |

Note: In the Southern Hemisphere, the signs of the azimuth gyro correction terms are reversed. That is, Earth rate (Ω) and velocity East (U) are negative.

CHAPTER 11 - INS ERRORS AND MIXED SYSTEMS

Schuler Tuning

1. When a pendulum is accelerated, the bob lags behind the suspension point in the opposite direction to the acceleration (Newton's First Law). When the acceleration stops, the pendulum oscillates with a period (T) equal to:

$$T = 2\pi \sqrt{\frac{L}{g}}$$

where: T is in seconds,

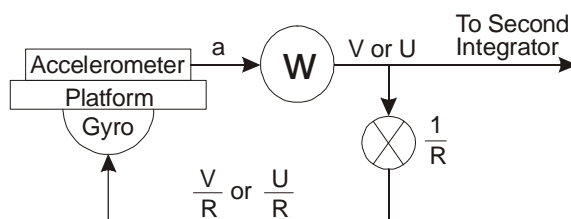
L is the length of the pendulum in feet

g is the gravity acceleration in feet/second²

2. **The Earth Pendulum.** Imagine a pendulum whose bob lies at the Earth's centre. If the suspension point were accelerated around the Earth, the bob would remain vertically below the suspension point because it is at the Earth's centre of gravity. A platform mounted on the suspension point tangential to the Earth's surface, ie horizontal, would therefore remain horizontal irrespective of the acceleration experienced. The vertical defined by the normal to the platform is therefore unaffected by acceleration. If, for any reason, the bob on the Earth pendulum became displaced from the Earth's centre, the pendulum would start to oscillate. The oscillation period would be 84.4 minutes (obtained by substituting the Earth radius in feet for L in the equation in para 1).

3. **The Platform Pendulum.** The INS stable element is maintained normal to the local vertical by feeding back the aircraft's radial velocity as levelling gyro control signals, and in this way the north and east accelerometers are prevented from detecting components of the gravity acceleration. The control signals are the $\frac{V}{R}$ and $\frac{U}{R}$ terms for vehicle movement (transport wander) applied as shown in Fig 1. By mechanizing the platform to remain horizontal, an analogue of the Earth pendulum of period 84.4 minutes is produced. Should the platform be displaced from the horizontal it would oscillate with a period of 84.4 minutes. This period is known as the Schuler period after Dr Maximilian Schuler who discovered the properties of the Earth pendulum. A platform is said to be 'Schuler Tuned' if its oscillation period is 84.4 minutes.

7-11 Fig 1 Schuler Tuning



ERRORS

Error Sources

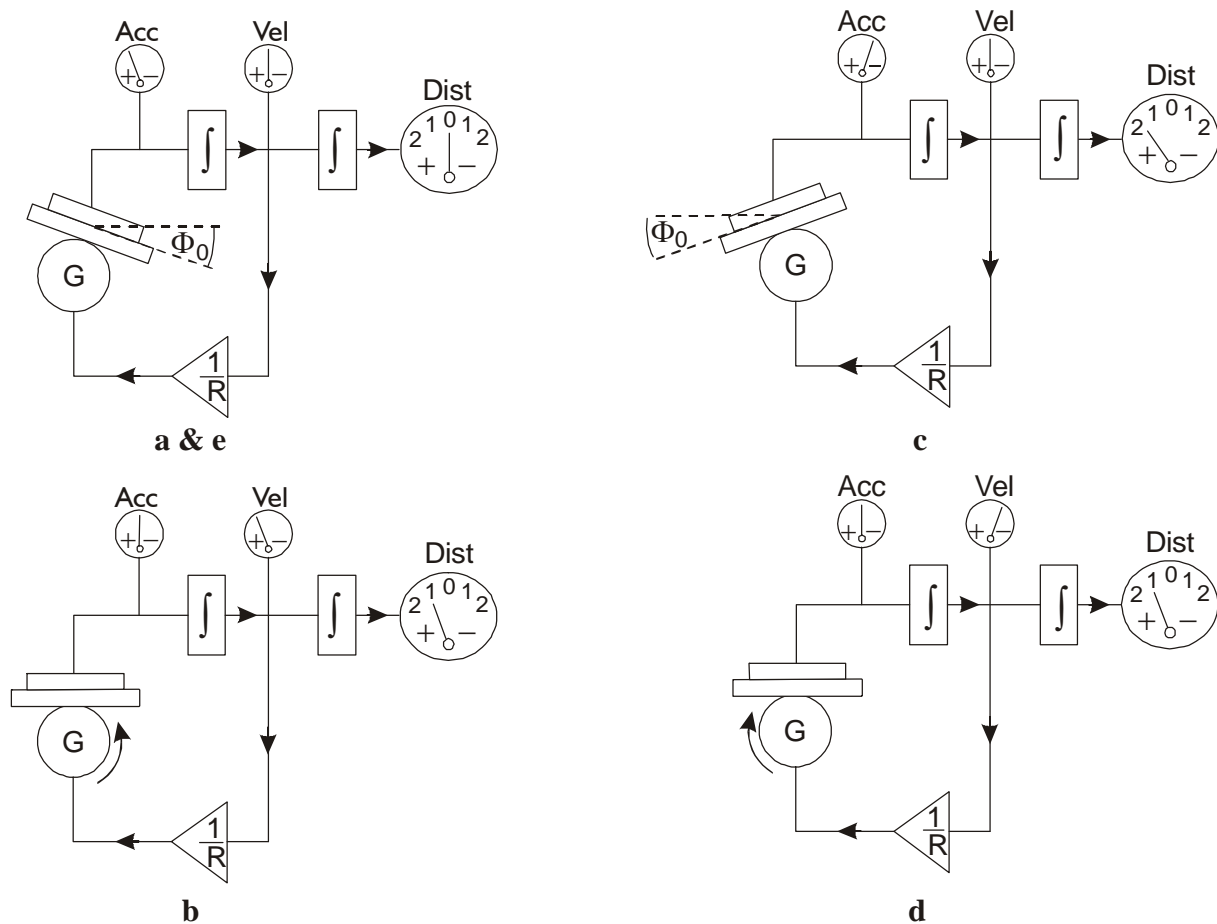
4. The following errors affect inertial navigation systems:
 - a. Initial levelling error.

- b. Accelerometer error.
 - c. Integrator error.
 - d. Levelling gyro drift.
 - e. Initial azimuth misalignment error.
 - f. Azimuth gyro drift.
 - g. Vertical channel errors.
5. **Bounded Errors.** Errors originating in, or effective within, the Schuler loops, are oscillatory and propagate at the Schuler frequency. These errors, which oscillate about a constant mean and therefore do not grow continuously with time, are termed bounded errors.

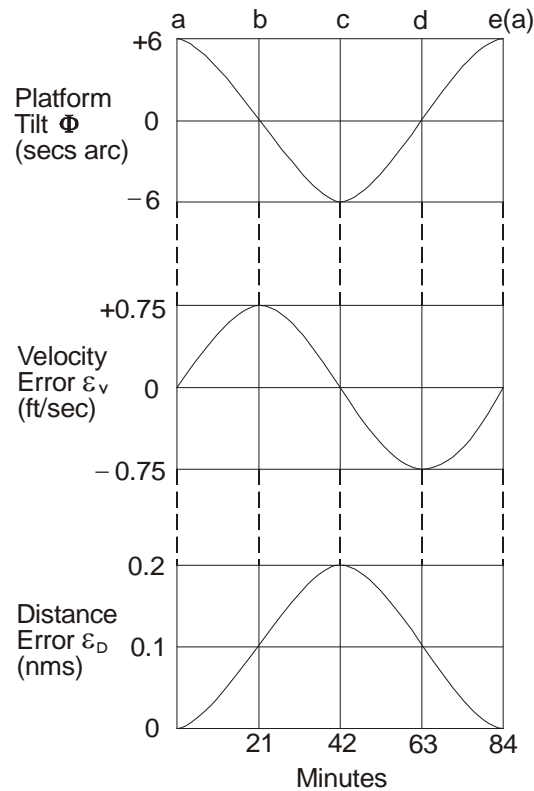
Initial Levelling Error

6. **Oscillation.** No matter how carefully the stable element (platform) and its sensors are aligned, there is always some residual error in the vertical, ie the platform is not completely level. When the 'navigate' mode is selected (at the conclusion of the alignment phase) the following sequence takes place. (Note: The lettering of the sub-paragraphs corresponds with the lettering in Fig, 2):
- a. The accelerometer detects the component of gravity, $g\Phi_0$ (strictly $g\sin \Phi_0$ but the approximation is correct for small angles and Φ_0 expressed in radians). Following the convention that clockwise tilts produce positive acceleration, $g\Phi_0$ is sensed as a positive acceleration. The integration of the accelerometer output takes a finite time, and therefore, velocity and distance are zero at the instant the 'navigate' mode is selected.
 - b. The integration of the detected acceleration produces a positive velocity which drives the platform anti-clockwise to the horizontal. The accelerometer now detects zero acceleration, but the positive velocity continues to drive the platform.
 - c. After the platform passes the horizontal the accelerometer detects the opposite gravity effect, sensed as a negative acceleration. The positive velocity reduces to zero at angle Φ_0 (the original tilt error) and for an instant the platform drive stops. However, the negative acceleration is integrated into negative velocity which drives the platform clockwise.
 - d. The clockwise drive brings the platform once again to the level position, resulting in zero output from the accelerometer. However, the negative velocity continues to drive the platform clockwise.
 - e. After the platform passes the horizontal, the accelerometer detects the gravity effect, sensed as a positive acceleration. This reduces the negative velocity to zero at angle Φ_0 . The cycle is then repeated.

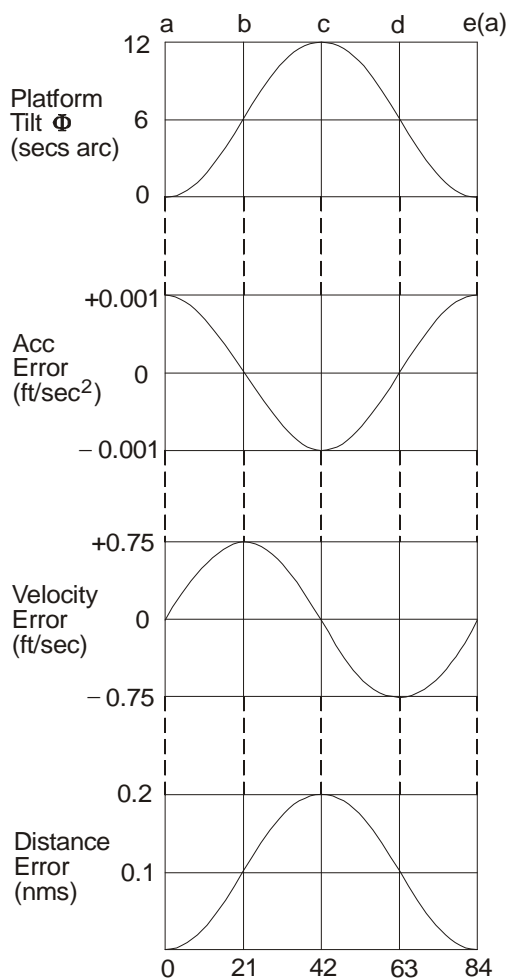
7-11 Fig 2 Initial Levelling Misalignment – Oscillation



7. **Initial Tilt Errors.** The errors caused by an initial tilt are shown in Fig 3. Note: the errors are bounded and do not increase with time. An initial levelling error of 6 seconds of arc is shown to cause a velocity error bounded by ± 0.75 feet per second (0.45 kt) and a mean distance error of 0.1 nm. After one complete Schuler period, both the velocity and distance errors have returned to zero. The a, b, c, d and e positions of the error curve are labelled to correspond to the sub-figure lettering of Fig 2.

7-11 Fig 3 Initial Levelling Misalignment – Errors**Accelerometer Errors**

8. Acceleration errors may be due to bias, cross-coupling, or vibropendulosity (Volume 7 Chapter 9). The error is integrated into an erroneous velocity which torques the platform at an incorrect rate. As with levelling errors, an oscillation is set up because the velocity error is fed back through the Schuler loop. Fig 4 shows the error curves generated by a bias error of 0.001 ft/sec².

7-11 Fig 4 Accelerometer Bias – Errors

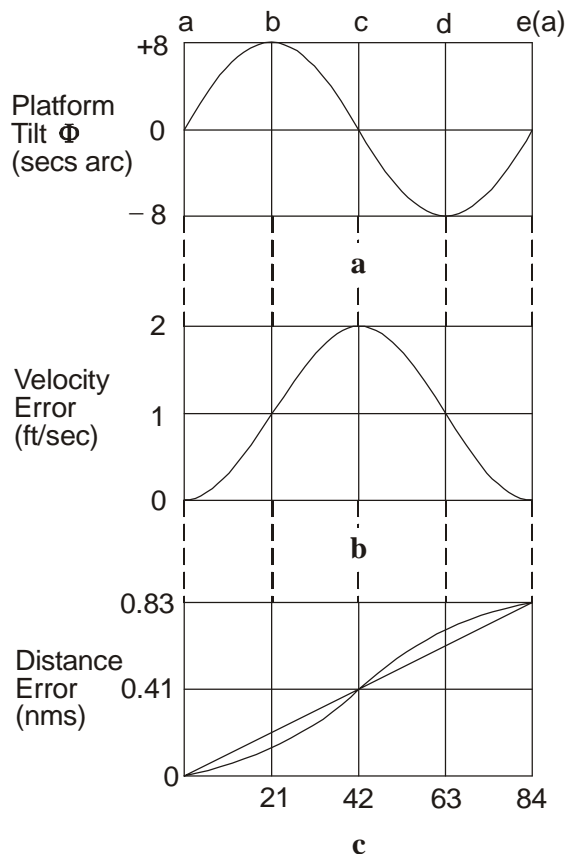
Integrator Errors

9. The first integrator is within the Schuler loop. Any error in the integration results in an incorrect velocity output which produces a platform oscillation and associated error curves similar to those previously discussed. The second integrator is outside the Schuler loop and any errors caused by it produce a position error that increases linearly with time.

Levelling Gyro Drift

10. Although the desired drift for an IN gyro is of the order of 0.001 %/hr it is probable that the drift rate in flight will be greater. A typical figure of 0.01 %/hr is used to illustrate the effect of gyro drift on the platform:

- Oscillation.** The stable element is turned away from the horizontal at the rate of 0.01 %/hr. The rotation is bounded by the Schuler loop, and the platform tilt curve is shown at Fig 5a.
- Velocity Error.** The acceleration error follows the same curve as that shown for platform tilt (Fig 5a). After integration the velocity curve at Fig 5b is obtained, which shows that a mean velocity error develops over the Schuler period.
- Distance Error.** The second integration results in the distance error which grows with time because of the mean velocity error. The growth rate is oscillatory about a mean ramp increase (Fig 5c). The distance error due to levelling gyro drift is unbounded, and for a drift of 0.01 %/hr, the ramp grows at approximately 0.6 nm/hr and has an oscillation of ± 0.13 nm.

7-11 Fig 5 Errors Caused by Gyro Drift**Initial Azimuth Misalignment**

11. If an INS is properly aligned in azimuth the East gyro senses zero component of Earth-rate and the North gyro outputs a signal proportional to $-\Omega \cos \phi$. If the INS is misaligned in azimuth by an angle ψ , the East gyro will output $\Omega \cos \phi \sin \psi$, and the North gyro $\Omega \cos \phi \cos \psi$.

12. The North gyro is torqued for Earth-rate by $\Omega \cos \phi$ and therefore the torquing error will be:

$$\Omega \cos \phi - \Omega \cos \phi \cos \psi = \Omega \cos \phi (1 - \cos \psi) \text{ } ^\circ/\text{hr}$$

Since the magnitude of the misalignment angle is unlikely to exceed 0.5° , this error may be disregarded. eg at Latitude 55° and $\psi = 0.5^\circ$ the error is:

$$15.04 \cos 55^\circ (1 - 0.99996) \text{ } ^\circ/\text{hr} = 0.0003 \text{ } ^\circ/\text{hr}$$

13. The error for the East gyro, given by $\Omega \cos \phi \sin \psi \text{ } ^\circ/\text{hr}$, is appreciable even when ψ is small. At Latitude 55° and ψ of 0.1° , the error is $0.015 \text{ } ^\circ/\text{hr}$. This error appears as East levelling gyro drift which causes the platform to oscillate about East and affects the North accelerometer, northern velocity, and latitude determination. The error curves produced in the latitude channel by an initial azimuth misalignment are similar to those caused by levelling gyro drift. The unbounded nature of the resulting distance error makes it essential to keep the initial azimuth alignment error as small as possible, and preferably less than 0.1° . The effect of various misalignment angles is shown in Table 1.

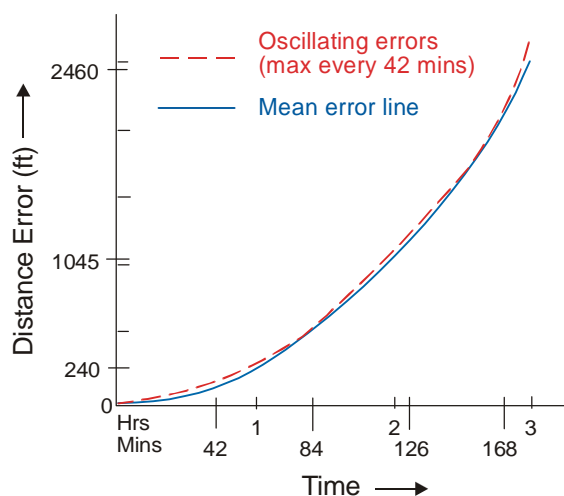
Table 1 Effect of Initial Azimuth Misalignment (North Channel) - LAT 55° N

| <i>Azimuth Error</i> | <i>Resultant East Gyro Drift</i> | <i>Distance Error after 1 hour</i> | <i>Max Velocity Error</i> |
|----------------------|----------------------------------|------------------------------------|---------------------------|
| 0.03° | 0.005°/hr | 2,220 ft | 1 ft/sec |
| 0.1° | 0.015°/hr | 6,600 ft | 3 ft/sec |
| 0.2° | 0.030°/hr | 13,310 ft | 6 ft/sec |
| 0.3° | 0.045°/hr | 19,970 ft | 9 ft/sec |

14. Azimuth misalignment also results in slightly incorrect accelerations being sensed by the misaligned accelerometers. The resultant errors may become significant under prolonged accelerations, e.g. during long accelerated climbs.

Azimuth Gyro Drift

15. Azimuth gyro drift ($\delta\psi$), like azimuth misalignment, registers as East levelling gyro drift and produces an increasing azimuth alignment error. The errors produced oscillate about means which increase with time. The increasing mean velocity error produces an unbounded distance error which follows a parabolic growth rate (illustrated in Fig 6).

7-11 Fig 6 Azimuth Gyro Drift - Distance Error

Errors in the Vertical Channel

16. Most of the errors in the horizontal channels have been shown to be bounded by the Schuler oscillations but this is not the case in the vertical channel. The vertical accelerometer must be corrected for the acceleration due to gravity (g_h) at the particular height before its output can be integrated to give rate of change and change of height. Gravity decreases with height according to the following relationship:

$$g_h = g_o \frac{R^2}{(R + h)^2} \text{ or } g_h = g_o - \frac{2g_o h}{R}$$

Where g_o = surface value of g

H = height

R = Earth radius

Any error in determining h will affect the calculation of g_h which in turn will increase the error. Therefore, errors in the height channel are not self-limiting and the channel is unstable.

17. The error growth is approximately exponential and, as an example, a step input error in vertical velocity of 0.1 ft/sec would generate a height error of 670 ft after half an hour and 15,000 ft after 1 hour. Thus, for most aircraft applications an INS vertical channel must be aided by another source - in the first instance by barometric altimetry.

Summary

18. A local vertical INS is inherently 'Schuler Tuned' and errors induced within the Schuler loop will cause the platform to oscillate about the horizontal. This oscillation results in some, but not all, errors being bounded as follows:

- a. Accelerometer errors, first stage integrator errors, and initial platform tilt errors yield a system velocity error which oscillates about a zero mean and so the distance error is bounded.
- b. Levelling gyro drift, azimuth misalignment and azimuth gyro drift cause a system velocity error which oscillates about a non-zero mean and thus the distance error is unbounded and oscillates about a ramp function of time, or a parabolic function in the case of azimuth gyro drift.

19. The vertical channel is not governed by Schuler oscillation and is inherently unstable due to the change of gravity with height.

20. Although the Schuler oscillations predominate in the short term (less than 4 hours), over the long term there are other periodic oscillations caused by interactions between the three axes.

MIXED INERTIAL SYSTEMS

Introduction

21. An INS is very accurate in the short term, but the introduction of errors is inevitable, and in a conventional local vertical system these cause the platform to oscillate about the horizontal. The shortcomings of such a system can be summarized as follows:

- a. The velocity error resulting from gyro drift oscillates about a non-zero mean and several applications, such as weapon aiming, require a very accurate velocity.
- b. The position error resulting from gyro drift is unbounded.
- c. High long-term accuracy requires very expensive components to minimize the errors.
- d. The system cannot be aligned in flight.
- e. The vertical channel is inherently unstable.

22. It is possible to overcome these disadvantages by 'mixing' the INS outputs with those of other navigation aids. These aids have errors which, although relatively large, do not increase with time and so a mixed system combines the short-term accuracy of a pure INS with the long term accuracy of another aid, thus enhancing the overall accuracy of both systems. Additionally, a mixed system can be aligned in flight.

23. The most common mixed systems are those which use Doppler as a reference velocity source which is used to damp the Schuler oscillations. Alternatively, an accurate fixing aid such as GPS could be used to bound the position error. In addition, it is possible to stabilize the vertical channel using barometric height. Sophisticated forms of mixing may involve several aids and use a software controlled statistical technique, such as Kalman Filtering, to continuously monitor and analyse the outputs to give the best results.

Doppler/IN Mixing

24. In a Doppler/IN system the Doppler and inertial velocities are compared to give an error signal which can be used in various configurations to modify the system performance and in particular to damp the Schuler oscillations. In a simple system the error signal is fed to the input of the first integrator, however in practice this leads to an unacceptable long time to reduce the error. In order to reduce the errors more quickly the error signal is in addition fed forward directly to modify the gyro torqueing signal; this arrangement is known as a 'Tuned Second Order' system.

25. The reductions in velocity error achieved with a tuned second order Doppler/IN system will have a significant effect on the accuracy of weapon delivery when compared with a pure INS. However, although the position error is slightly reduced, there is in general little to be gained in positional accuracy when the two systems are compared during the first 4 to 6 hours of flight.

Fix Monitored System

26. The problem of the unbounded position error in a pure INS or Doppler/INS can be reduced by coupling the system with an accurate fixing aid such as GPS. The latitude and longitude outputs from the fixing aid are compared with those from the INS and the resulting error signals are fed through suitable gains to update the inertial position. The difference signals are also used to provide a degree of damping to the platform. The fix-monitored arrangement has the disadvantage of relying to a certain extent on an external source of information whereas the pure INS and Doppler/INS are self-contained.

Airborne Alignment

27. The self alignment and reference alignment techniques are restricted to a fixed base as it is not possible for a pure INS to distinguish between the accelerations due to aircraft movement and those due to platform misalignment. An INS which is combined with an alternative velocity source or position information can however be aligned in flight. The platform is roughly aligned and then the error signals from either the external velocity or position information are used to level the platform and align it in azimuth. The accuracy of an airborne alignment is not as high as that obtained from a full self-alignment, but the technique does give the aircraft a rapid reaction capability and the ability to update the system during a long flight or after a transient equipment failure.

Vertical Channel Stabilization

28. The vertical channel does not display the same characteristics as the horizontal channels as it is inherently unstable due to the fact that the value of g varies with height. It is therefore necessary to supplement the vertical channel with another source of height reference in order to provide the accurate values of height and vertical velocity which are essential for weapon aiming calculations.

29. The barometric altimeter whilst inaccurate in the short term is very accurate in the long term and this characteristic can be used to stabilize the INS height and vertical velocity outputs. The inertial height output is compared with the barometric height to give an error signal which is fed back to the first integrator and this has the effect of stabilizing the accuracy in the long term whilst maintaining it in the short term.

KALMAN FILTERING

Introduction

30. The hardwired mixed systems described in the preceding paragraphs are inflexible because the feedback gains are fixed and have to be carefully chosen at the design stage, thus assigning a fixed level of relative performance to the sensors. In reality, the relative merits of each sensor will vary considerably and depend on such parameters as time of flight, range from a ground aid, flight conditions, and altitude. Consequently, the weighting factor applied to each sensor of a mixed system by a fixed gain loop is unlikely to be the true measure of the relative merits of the sensors and could possibly be significantly in error. By using a software controlled statistical technique, such as Kalman Filtering, these weighting factors can be optimized and continuously updated for any operating conditions. This method can use any number of sensors and can select the best information available at any particular time.

Kalman Filter Design

31. The Kalman Filtering process estimates each of the parameters which give rise to an error between the INS and one or more external sensors on the basis of maximum likelihood. By using a weighting factor which is continually revised the error between the external data and the INS is apportioned among all the possible error sources so that the probability of these errors occurring is greatest.

32. The computer holds an estimate of the system errors and uses known error propagation equations to forecast how these errors will behave with passing time. Thus, this error model will always maintain up to date values. When an external measurement is made the error held in the computer is compared with the measured error. All the quantities in the error model are then corrected in the light of the known variances of the external information and the variances of each quantity in the error model. The variances of the system errors are recalculated after each external measurement has been processed so that the errors of the next measurement can be apportioned in the optimum manner.

33. The design of a practical Kalman Filter for use in an aircraft system is complex. The first problem is to define a set of variables that specify the system. In practice, there is never enough information to enable the system to be perfectly modelled and there will frequently be limitations on computer time and storage. Extensive trials and simulation are necessary to enable the designer to define the error model and variables as accurately as possible within the computer limitations. Once designed however the filter performance is not affected by changes in aircraft role or tactics, and additional sensors can be incorporated into the system with relatively minor changes to the computer software.

Advantages of Kalman Filtering

34. By making better use of the information available, Kalman Filtering increases the flexibility and enhances the accuracy of a mixed system thus overcoming the disadvantages of a hardwired mixed system. Other important advantages are:

- a. Alignment and gyro drift trimming are improved.
- b. Weapon aiming accuracy is improved including the elimination of fixed bias errors.
- c. Post flight analysis of the navigation system and fault detection can be carried out.
- d. An estimate of system accuracy can be continuously displayed to the crew.

35. **Alignment and Gyro Drift Trimming.** A Kalman Filter can be used during alignment and for drift trimming the gyros. The times for full and rapid alignments can be reduced and the overall accuracy of the process improved. The filter can compensate for aircraft movement such as wind buffeting during ground alignment and also take account of the changing characteristics of components during the warm-up phase.

36. **Weapon Aiming Errors.** The Kalman Filter will directly affect weapon-aiming accuracy because of the improved navigation performance. It can also take account of fixed bias errors and in particular harmonization and windscreen distortion. By calculating the errors in the delivery of practice weapons in training sorties, the aircraft can be calibrated, and the filter programmed to eliminate these fixed bias errors. Kalman Filtering also improves the height and vertical velocity outputs which are essential for accurate weapon aiming.

37. **Post Flight Analysis and Fault Detection.** An important secondary application of Kalman Filtering is the post flight analysis of the navigation system. During flight, all reference data can be recorded and subsequently fed into a computer containing a much more comprehensive error model than it is possible to accommodate in an airborne computer. The method of using the data is the same as in flight but as the error model is more complete maximum use can be made of data which was previously unused. This post flight analysis highlights shortcomings in the airborne filter which may then be amended. By the use of post flight analysis data can be used to show when any sensor is not presenting navigation information within the expected variance due perhaps to progressive deterioration of components or incipient failure. Such a facility enables the thorough testing of the sensor to be carried out at an earlier stage than might otherwise have been possible.

38. **Estimate of System Accuracy.** When using statistical filtering an estimate of the navigation system accuracy is continuously available. This information may be displayed to the crew directly as a figure of merit reflecting the accuracy of the navigation outputs. Alternatively, the filter can automatically reject input data that is in error by more than 3 or 4 standard deviations and an indication given to the crew.

Summary

39. A navigation system using a Kalman Filtering technique is far more flexible and accurate than a more conventional system and has several secondary benefits. The advantages obtained from Kalman Filtering are limited only by the ability to accurately model the system parameters and the likely errors, within the computer time and space available.

CHAPTER 12 - GROUND CONTROLLED APPROACH (GCA)

Introduction

1. A Ground Controlled Approach (GCA) is a procedure by which air traffic controllers pass a series of verbal instructions, based on radar observations, to enable the pilot to make an approach to the runway in conditions of bad weather and poor visibility. This chapter describes a GCA in general terms, but full definitions and explanations of the procedure, responsibilities and associated R/T phraseology, are contained in the Civil Aviation Authority (CAA) Radiotelephony Manual (CAP 413). NATO standard, AATCP-2, Radiotelephony Phraseology also details R/T procedures. Provided that the pilot is proficient at instrument flying, the procedure is relatively straightforward, but variations can occur to take account of local conditions and the prevailing air traffic situation. Procedures at civilian airfields, and at military airfields of other nations, may vary in detail from those described, but the principles remain the same.
2. A full GCA consists of two elements:
 - a. A Surveillance Radar element, at which stage the aircraft will be identified and then vectored onto the final approach path.
 - b. A Precision Approach Radar (PAR) element, during which the pilot will be given instructions in height and azimuth, with respect to the runway touchdown point.
3. **Air Traffic Control (ATC).** A complete GCA procedure will normally involve three ATC controllers, in sequence:
 - a. **The Approach Controller.** The Radar Approach Controller will carry out the initial identification of aircraft wishing to enter the Military Air Traffic Control Zone (MATZ).
 - b. **The Radar Director.** The Radar Director will action the sequencing of aircraft for an ordered approach to the runway.
 - c. **The Talkdown Controller.** The Talkdown Controller will interpret the information presented by the PAR equipment and pass precise heading and height instructions to the pilot, to enable him to complete the final approach.

THE SURVEILLANCE RADAR ELEMENT

Approach Control

4. Aircraft wishing to penetrate a MATZ should make initial contact with the Radar Approach Controller (RAC) on the designated 'Approach' frequency, unless the airfield's promulgated Terminal Procedures dictate otherwise. The RAC will utilize the surveillance radar display to locate and identify the aircraft and vector it towards the final approach path to the runway in use. Aircraft are identified by IFF/SSR equipment, or by their response to manoeuvring instructions.
5. After identification, the RAC will pass any information necessary, and then, having positioned the aircraft at a suitable height, will hand over control to the Radar Director.

Radar Direction

6. The Radar Director will sequence aircraft for an ordered approach to the runway in use, by means of 'normal pattern' and 'short pattern' radar circuits:

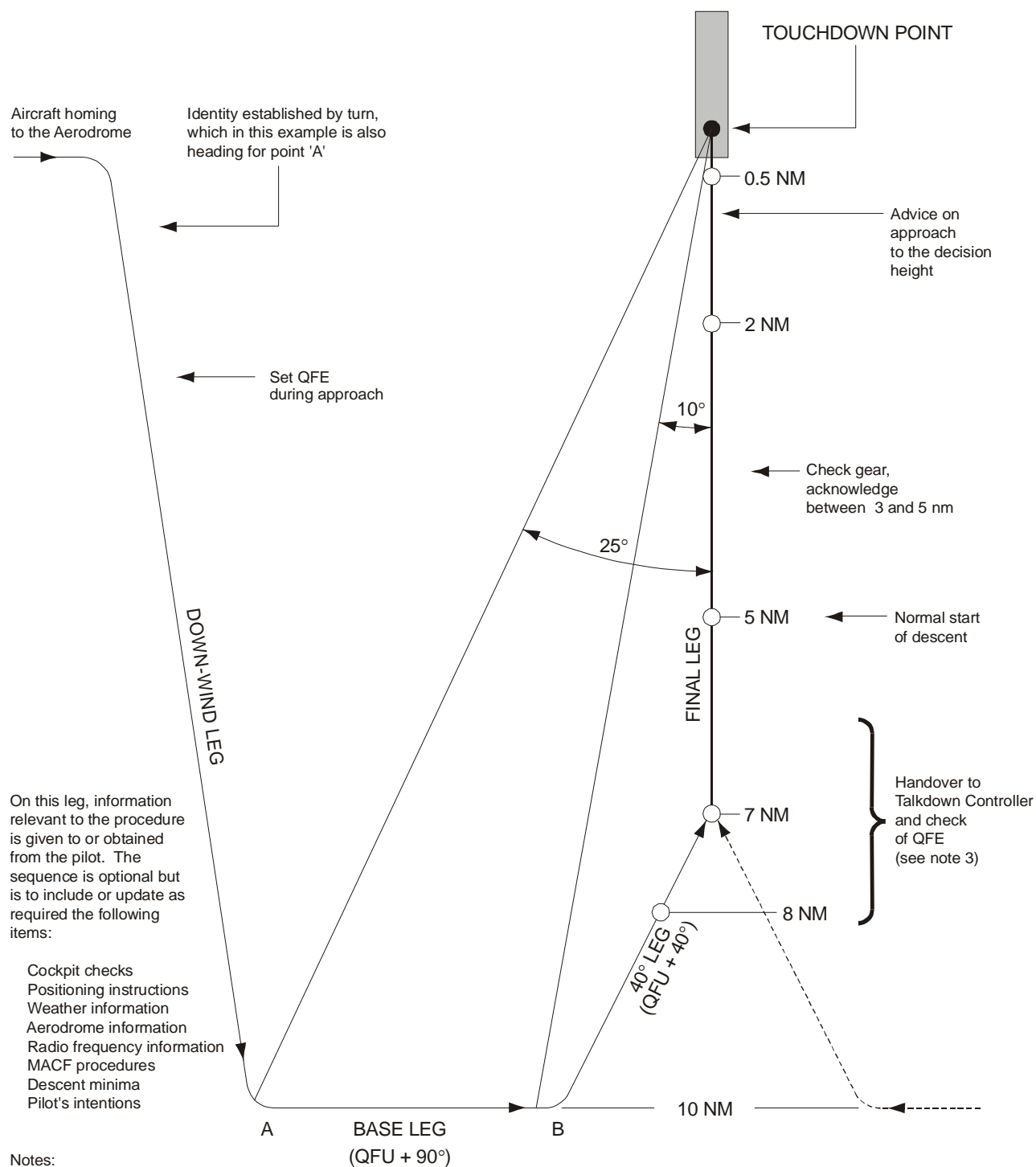
a. **The Normal Pattern Circuit.** The main elements of a normal pattern are a base leg, a converging heading and the final approach (see Fig 1). For multiple circuits, a downwind leg will also be included. The pattern is orientated upon the bearing of the runway centreline (the magnetic bearing of the runway centreline is referred to as 'QFU').

b. **The Short Pattern Circuit.** An aircraft overshooting from an instrument approach, and precluded by weather conditions, or other reasons, from carrying out a visual circuit or a normal pattern radar circuit, may be repositioned on the final approach using the short pattern circuit procedure. This will consist of a reciprocal track, parallel to the runway centreline, with an inbound turn to position the aircraft back on the centreline, at a point just before it intercepts the glidepath.

7. **Downwind Leg.** The Director will pass magnetic headings and heights to fly (based on QFE) to the pilot, in order to vector the aircraft to the downwind leg. The downwind leg extends from a point abeam the runway threshold, to a point situated at a range of 10 nm from touchdown and $\pm 25^\circ$ from the reciprocal of the runway centreline (Point A in Fig 1). The pilot should acknowledge all messages; executive instructions are to be read back verbatim. The Director will pass the Instrument Approach Minima (IAM) for the runway in use (missed approach and communications failure (MACF) instructions will also be passed to pilots not familiar with the airfield). The pilot will be asked for his decision height (DH) based on the IAM, the pilot's instrument rating and any other allowances necessary (see Flight Information Handbook). The pilot will also be asked for his intentions at the end of the approach (e.g. to land, join circuit, depart etc). During the downwind leg, the aircrew will carry out pre-landing checks.

8. **Base Leg.** At a range of approximately 10 nm downwind (depending on aircraft type, the needs of separation and controller availability) the aircraft is directed on to a base leg, the track of which is at an angle of 90° to the runway QFU. On Fig 1, the base leg is shown as from Point A to Point B (the point where a line bearing $\pm 10^\circ$ from the reciprocal of the runway QFU intercepts the base leg).

7-12 Fig 1 Standard GCA - Normal Pattern



Notes:

(1) The 10-mile distance can be varied to suit the aircraft type and adjacent traffic patterns

(2) Position of handover to Talkdown Controller may be varied to suit aircraft type

(3) QFE to be set prior to commencing final approach

9. Converging Heading. The use of a converging heading, $\pm 40^\circ$ from the QFU, alleviates the use of a 90° turn and provides a gradual, smooth closure with the final track, to the advantage of both the pilot and the controller. Depending upon the precise distance flown downwind, the aircraft will normally intercept the extended centreline of the runway at about 7 nm from touchdown.

10. **Height Control.** The final talkdown is normally started from 1,500 ft above touchdown zone elevation (TDZE). If, because of terrain clearance or other air traffic considerations, it is not possible to fly the whole circuit at 1,500 ft, the pilot is normally instructed to reduce height to 1,500 ft before the glidepath is reached.

11. **Handover for Final Approach.** From the converging 40° Leg, the aircraft is turned to close with the extended centreline of the runway. At this point, the Director will hand the aircraft over to the Talkdown Controller, for the final precision approach.

The Surveillance Radar Equipment

12. There are two types of Surveillance Radar Equipment (SRE) in current use; the 'Watchman' and the AR 15/2A radar equipment. Both systems are primary radar, and each provides 360° scan, plan position indicator (PPI) displays for use by controllers.

a. **Watchman Radar.** The Watchman radar operates in frequency diversity in the E and F bands. The azimuth beam width is nominally 1.5°. This radar can produce short pulse widths for improved target discrimination at short ranges, and a long 'chirp' pulse for detecting targets of small radar echoing areas at long range. It also has a Moving Target Indicator (MTI) capability, to assist with discriminating between moving targets and 'clutter'. The maximum range is 60, 80 or 120 nm, depending upon the particular installation.

b. **AR 15/2A Radar.** The AR 15/2A radar operates in the E/F band. It has duplicate transmitter/receiver circuits, which utilize the common aerial, but can use two discrete frequencies by transmitting fractionally apart. It has an MTI facility, and a maximum range of 60 nm.

13. If necessary, the surveillance radar PPI display can be used to provide a Surveillance Radar Approach (SRA) to the runway touchdown point (see para 25).

THE PRECISION APPROACH RADAR ELEMENT

The Talkdown

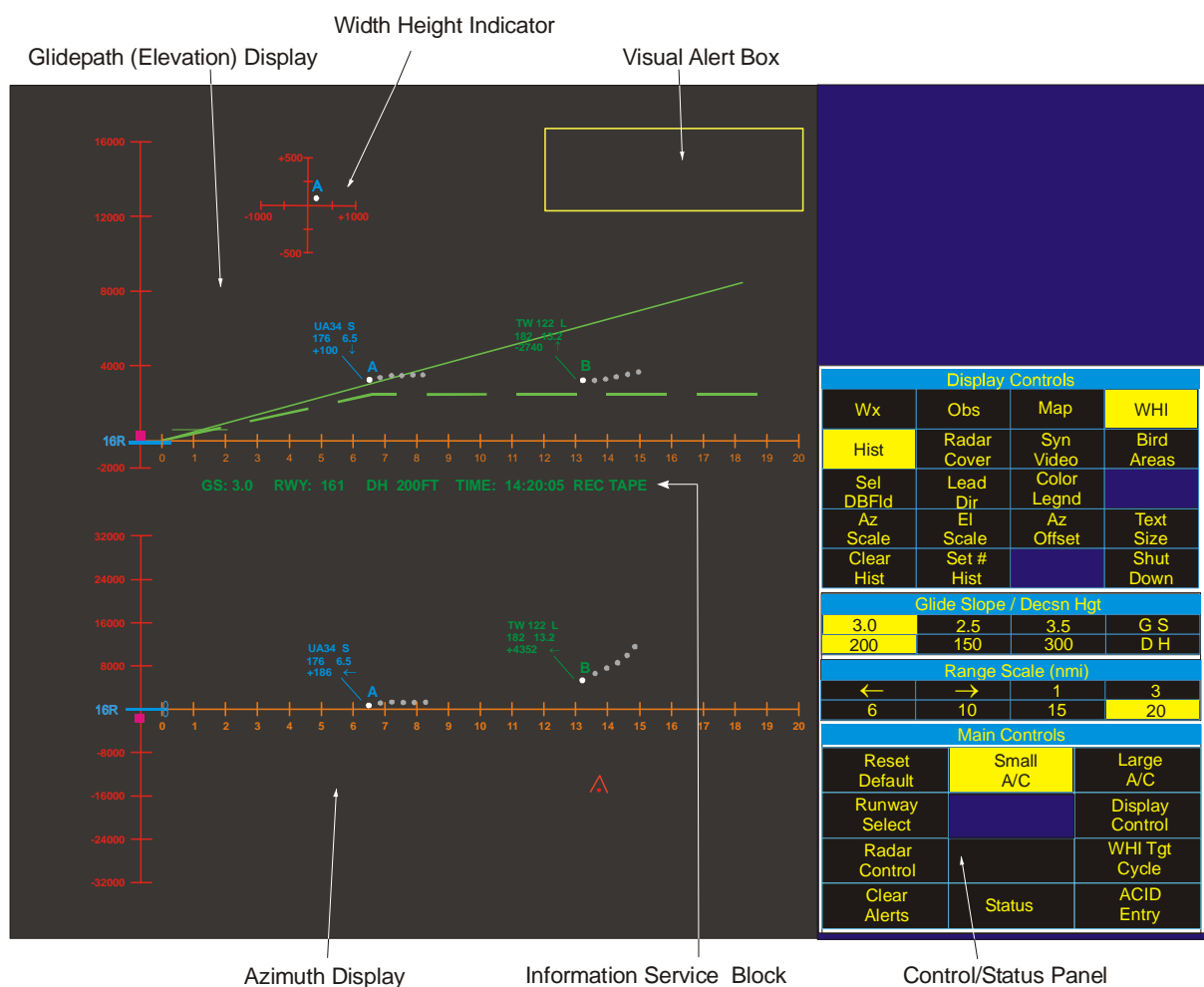
14. The Talkdown Controller will interpret the information presented by the PAR displays (Fig 2), and will pass precise heading and height instructions to the pilot, to enable him to follow the designated glidepath and thus complete the approach.

15. The aim of a PAR 'talkdown' is to enable the pilot to obtain at least one of the required visual references at, or before, decision height (DH), and to be in a position to continue the approach to land. The accuracy of a PAR procedure depends to a large extent on the proficiency of both the pilot and the Talkdown Controller.

16. The Talkdown Controller takes over at about 7 nm from touchdown, usually on a different frequency from that used by the Director. When positive contact has been established, the Talkdown Controller will ask the pilot to read back the altimeter pressure setting. The Controller will then instruct the pilot not to acknowledge further instructions. Immediately prior to the descent point, the Controller will warn the pilot of the impending descent. The slope of the PAR elevation centreline (the glideslope) is normally set for 3°, although 2.5° may sometimes be offered. For a 3° glideslope, the descent point will be at a range of 5 nm, provided the aircraft is at 1,500 ft on the QFE. At this point, the Controller will instruct the pilot to start a descent, and thereafter he will pass instructions to

enable the pilot to fly the aircraft down the glidepath whilst maintaining the azimuth centreline. The Talkdown Controller will interrupt his instructions with short, regular breaks in transmission. These breaks provide the pilot with an opportunity, where necessary, to pass a short message to the Controller without detracting from the continuity of the procedure. At a range between 5 nm and 3 nm from touchdown, the Controller instructs the pilot to confirm that the undercarriage is down (this check is not required for aircraft with fixed undercarriage). After obtaining suitable clearance from the Aerodrome Controller, the Talkdown Controller will pass authority for the pilot to land. Descent instructions are continued until the Talkdown Controller advises the pilot that he is approaching his DH. From this point onwards, only advisory information is passed, and the pilot must continue the descent visually or carry out the missed approach procedure if conditions do not permit a landing. The pilot will normally remain on the Talkdown frequency until safely landed.

7-12 Fig 2 The Precision Approach Radar Display



17. PAR is classed as a 'precision' approach aid. Equipment limitations, which take into account obstacles on the approach path, are published as Instrument Approach Minima (IAM) in the Terminal Approach Procedures for each airfield.

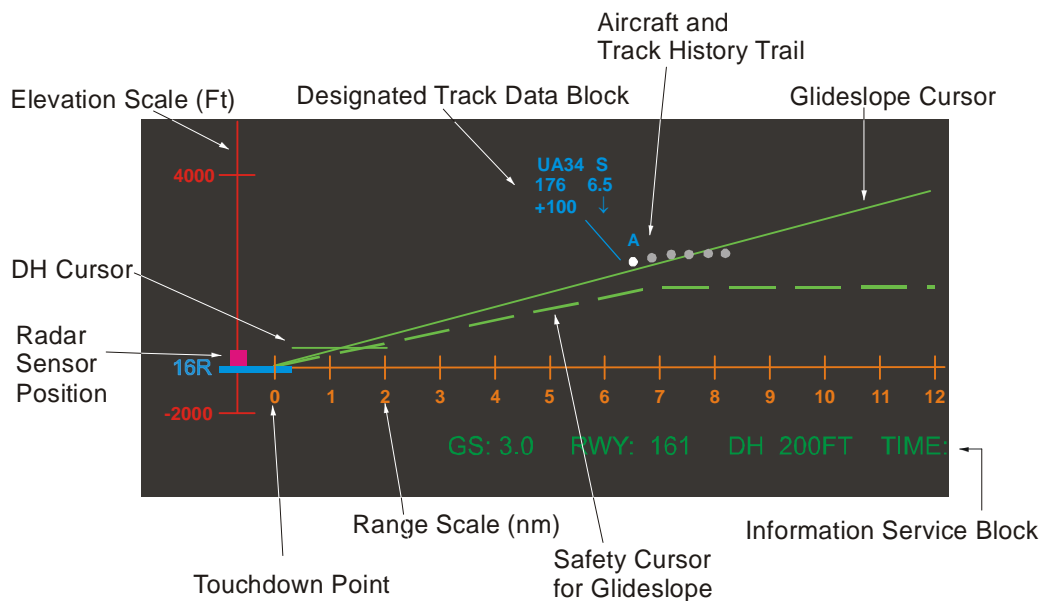
The Precision Approach Radar Equipment

18. The aircraft's final approach path to the runway is scanned accurately by the precision approach radar (PAR). The 'PAR-2000' equipment is a short wavelength (three centimetre) radar, with two linear phased array aerals, each transmitting a narrow beam. The vertical aerial defines the aircraft's

elevation, whilst the horizontal aerial defines the aircraft's azimuth. The PAR-2000 is less susceptible to weather interference than previous equipment.

19. **The PAR-2000 Display.** The PAR-2000 screen (Fig 2) will display multiple aircraft tracks to the controller. The position of each aircraft is shown simultaneously on two displays. The upper display shows range and height (see expanded illustration at Fig 3), whilst the lower display shows range and azimuth. An Information Service Block, located vertically between the two displays, provides the controller with approach data. The right-hand part of the screen incorporates a control/status panel which facilitates operator interface by means of a 'trackball' (combination roller ball/mouse).

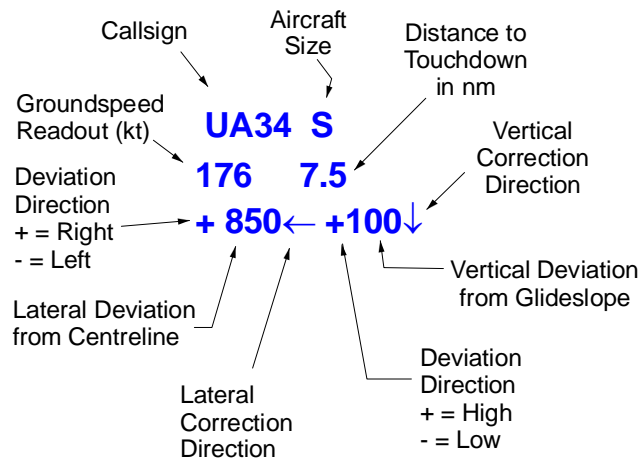
7-12 Fig 3 Detail from the Glidepath Display



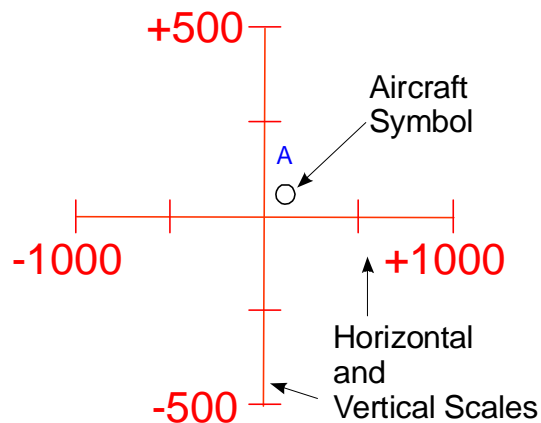
20. **Change of Parameters.** Being a computerized, digital format, the PAR-2000 display provides a large degree of inherent flexibility for the operator, including:

- Change of Scale.** The displays can be presented at different range scales (from 1 nm to 20 nm).
- Glideslope.** The glideslope can be adjusted from 2.0° to 6.0°, in steps of 0.1°.
- Decision Height.** The controller can set and display the pilot's DH in 10 ft increments (see Fig 3).

21. **Designated Track Data Block.** When an aircraft has been designated by a controller, a Designated Track Data Block appears on each display alongside its respective track symbol. This data block contains three rows of information - the top row contains aircraft details, the middle row gives groundspeed and distance from touchdown, whilst the bottom row shows data on vertical and lateral displacement from the PAR centreline. An example of a Designated Track Data Block is illustrated in Fig 4.

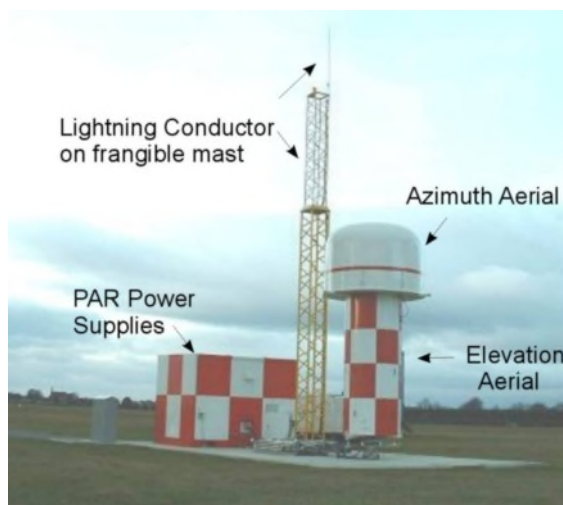
7-12 Fig 4 Example of a Designated Track Data Block

22. **The Width Height Indicator.** In addition to showing a representation of the aircraft's flight path, the display shows its precise displacement from the glideslope and azimuth centrelines in several formats. One such format is the Width Height Indicator (WHI), which illustrates azimuth and elevation displacement against a scale in feet, as seen from the pilot's viewpoint. In the example shown in Fig 5, the aircraft is slightly high, and slightly right of the PAR centreline.

7-12 Fig 5 The Width Height Indicator

23. **Runway Changes.** The PAR-2000 system can provide radar cover for approaches on up to 3 runways (i.e. 6 touchdown points), although, at some airfields this number may be reduced due to sighting restrictions. The touchdown point for each runway is pre-set into the PAR computer as an offset value relative to the location of the aerial. To re-align the PAR to suit a change of runway, the radar aerial assembly, shown in Fig 6, is rotated in azimuth, whereupon it automatically locates a surveyed radar reflector, to confirm the correct re-alignment. When this action is complete, the PAR screen will display an "Available for Operational Use" message. If the radar becomes misaligned or loses sight of the runway reflector for a prolonged period of time, the system will generate a 'critical' alert message on the display screen. The whole re-alignment process can be completed in approximately 1½ minutes.

7-12

Fig 6 The PAR Airfield Installation**OTHER APPROACHES****Straight-in Approach**

24. As a variation on a full GCA procedure an aircraft may be fed directly into the GCA pattern from an airfield approach aid such as an NDB, VOR, or TACAN. The pattern is arranged so that the aircraft is vectored onto the extended centreline at a range of 10 to 15 nm. The radar controller vectors the aircraft in to seven miles, where the Talkdown Controller takes over for the final approach. This procedure is often used when an aircraft is handed over from an Air Traffic Control Radar Unit to an airfield, under the Centralized Approach Control (CAC) procedure.

Surveillance Radar Approach

25. When precision approach radar is not available, the surveillance radar may be used to carry out a surveillance radar approach (SRA). Using the PPI display, the controller has no electronic glidepath information and there is less accuracy in azimuth when compared with a PAR. An SRA is therefore classed as a non-precision approach.

26. The controller passes headings to fly and other instructions to enable the pilot to fly a pattern similar to that of a standard GCA procedure.

27. The controller will pass advisory heights with range information, to assist the pilot to maintain the correct rate of descent for an equivalent glideslope angle. At 5 nm the pilot is requested to start a descent. If a 3° glideslope is used the approach height will be 1,500 ft and the rate of descent will be 300 ft per mile.

CHAPTER 13 - INSTRUMENT LANDING SYSTEM (ILS)

Introduction

1. The Instrument Landing System (ILS) is a pilot interpreted aid which provides azimuth, elevation, and limited range information to enable a runway approach to be made under instrument flight conditions. The ILS system comprises ground radar installations and an airborne receiver. The ground radar transmits an azimuth beam (localizer) defining the approximate centreline of the runway, and an elevation beam (glidepath) defining a slope for a safe descent. The ILS is classed as a precision approach aid.

Principle of Operation

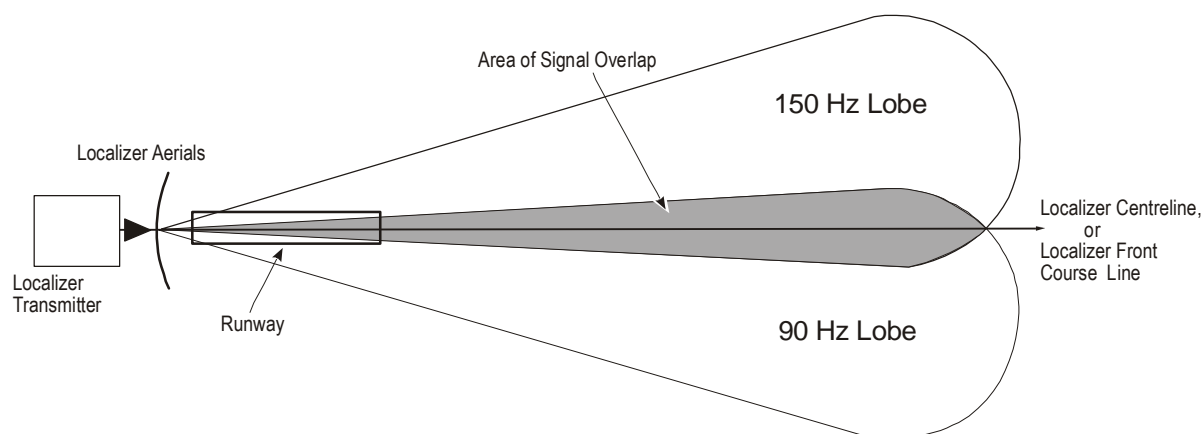
2. The ground transmitters emit localizer and glidepath signals simultaneously. These define the approach path to the runway using the principle of lobe comparison. The airborne receivers provide signals of aircraft displacement from the ILS centreline and glidepath which can be used to feed an ILS indicator on an HSI, a flight director, or a separate ILS indicator. They can also feed directly to an autopilot system. Ground marker beacons radiate vertically to provide range checks.

The Localizer

3. The ground localizer transmitter is usually installed approximately 1,000 ft beyond the upwind end of the instrument runway, with the aerial in line with the runway centreline. In some military installations the approach path is offset from the runway centreline because of siting difficulties. For this reason, the runway centreline and the localizer front course line are not necessarily coincident. Any angle of offset is generally less than 3° and the magnetic orientation of the runway and ILS centrelines are listed in the En-Route Supplement and TAPs.

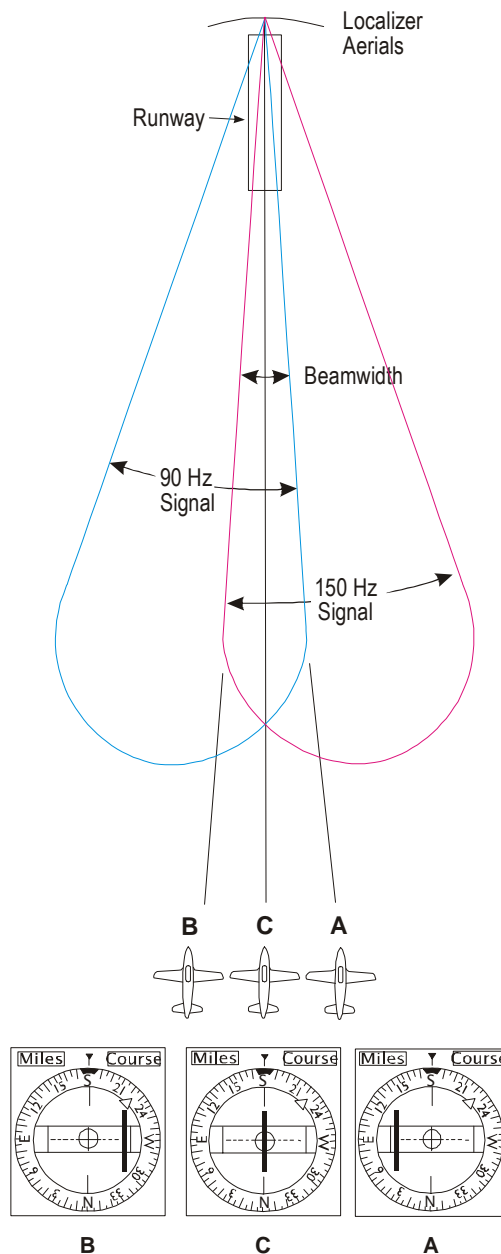
4. The localizer operates on VHF frequencies in the range 108.1 MHz to 111.95 MHz, utilizing those frequencies with an ODD first decimal place (e.g. 109.10 MHz, 111.35 MHz). The transmitter emits two lobes which overlap. The centre of the region of overlap provides an area of equal signal amplitude defining the approach path (see Fig 1). The two lobes are modulated at different frequencies, the beam to the right of the centreline (when approaching to land) at 150 Hz; that to the left at 90 Hz. The localizer transmitter also emits the identification code for the installation. This code consists of a two letter morse code signal transmitted every ten seconds, which can be received aurally by the pilot. In an emergency, ATC can interrupt the localizer channel and use that frequency for speech transmissions to the aircraft, but not vice versa. The range of the localizer signal is 25 nm at 2,000 ft within plus or minus 10° of the localizer front course line or 17 nm between 10° and 35° of the localizer front course line. Above 2,000 ft these ranges may be reduced.

7-13 Fig 1 ILS Localizer Lobes



5. The aircraft localizer receiver detects the two signals and produces two output voltages which are compared and the difference displayed as a deflection of the display deviation bar as shown in Fig 2. The aircraft at position A will receive a signal from the 150 Hz modulation which is stronger than that from the 90 Hz modulation. This will cause a large deflection of the deviation bar. Conversely an aircraft at B will receive a stronger signal from the 90 Hz lobe and the deflection will be equally large but in the opposite sense to that experienced at A. At position C, the two signals will be equal and there will be no deviation bar deflection. The display convention is such that the display centre represents the aircraft, and the deviation bar the ILS centreline, so, for example, at position A the deviation bar to the left indicates that the ILS centreline is to the left.

7-13 Fig 2 Behaviour of HSI Deviation Bar

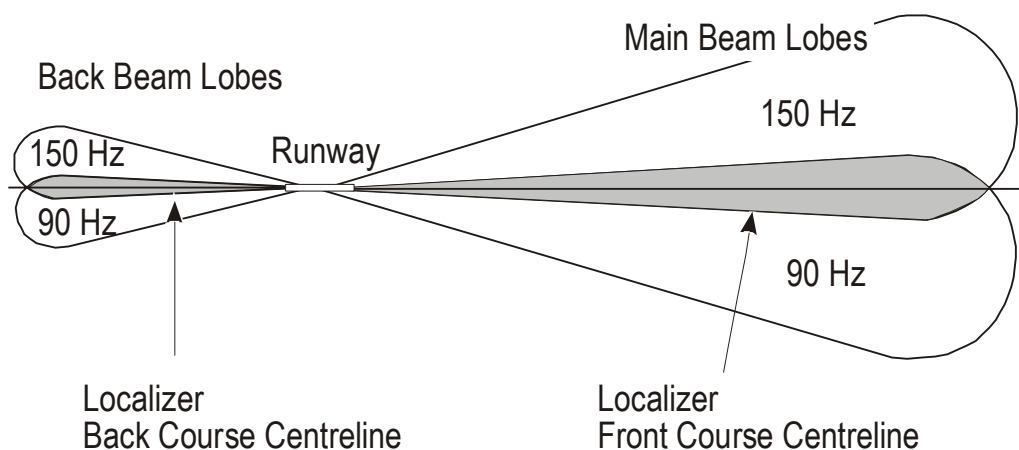


6. **Display Sensitivity.** Variations in azimuth between positions A and B in Fig 2 will cause the deviation bar to vary between the deflections shown. The display sensitivity will dictate at what angle, either side of the centreline, full-scale deflection occurs. This is normally adjusted so that full-scale deflection corresponds to $\pm 2.5^\circ$, which in turn corresponds to the beamwidth shown in Fig 2 where the two signals overlap. It should be noted that the displacement of the deviation bar reflects the angular

distance of the aircraft from the centreline and, thus, the closer the aircraft is to the transmitter the larger is the displacement of the deviation bar for any given linear distance from the centreline.

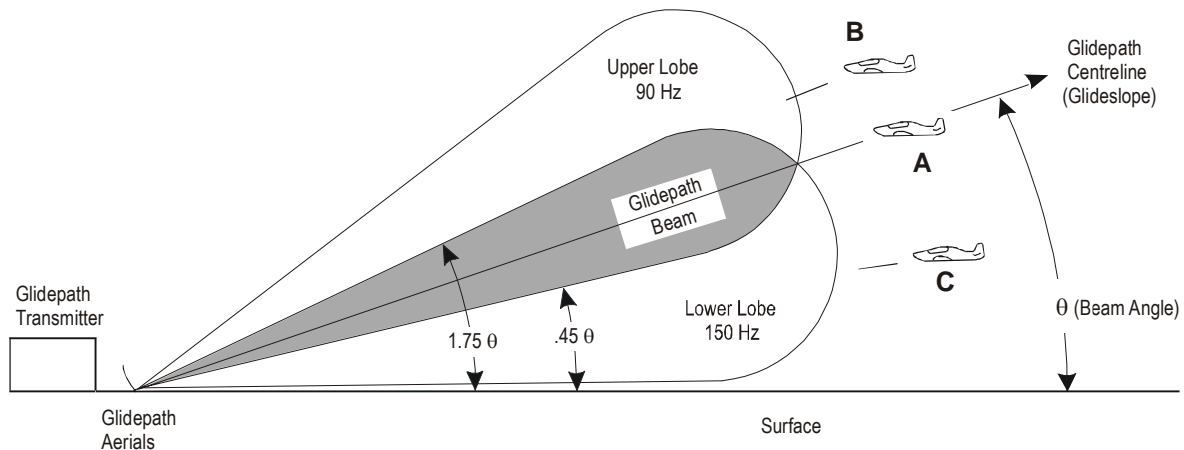
7. **Back Beams.** The majority of localizer transmitters have some signal 'spillage', and produce sidelobes in the direction opposite to the front course line. These signals may form lobes known as a 'back beam' (Fig 3). Back beam signals should not normally be used for navigation. However, a few ILS installations transmit a back beam intentionally. These facilities are promulgated in Flight Information Publications (FLIPs) and can be used by aircraft to fly a 'back course' towards the airfield overhead. Care is essential when flying a back course, as it is less accurate than the front course beam, and there is no associated glidepath information. In addition, because the back beam is narrower than the main beam, minor heading changes will cause large deflections of the Deviation Bar. When approaching the aerodrome along a back course, the localizer needle on a basic ILS indicator will operate in the reverse sense (i.e. if the aircraft is right of the centreline, the demand will show 'go right' instead of 'go left'). This problem can be overcome if using a Horizontal Situation Indicator (HSI), by setting the reciprocal of the back course centreline (i.e. set the front course centreline). The ILS demands will then be displayed in the correct sense. As an example, in Fig 3, if approaching on a back beam centreline of 090°, the pilot should set the HSI course to 270°.

7-13 Fig 3 ILS Back Beam

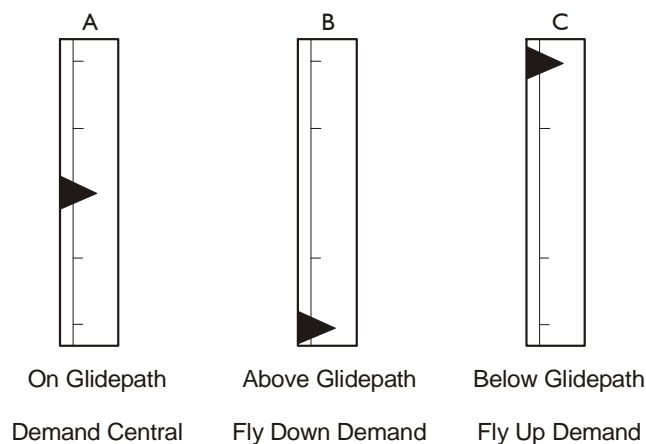


The Glidepath

8. The glidepath transmitter frequency is paired with the localizer frequency, but operates on UHF, in the range 328.6 MHz - 335.4 MHz. The glidepath produces two overlapping lobes, analogous to the localizer but in the vertical sense, as shown in Fig 4. The transmitter is offset some 300 to 400 ft to the side of the runway at a position approximately 750 to 1,000 ft from the threshold depending on the runway length and the touchdown point. The upper lobe is modulated at 90 Hz and the lower lobe at 150 Hz. The range of the glidepath signal is 10 nm at 2,000 ft up to 8° in azimuth either side of the localizer front course line.

7-13 Fig 4 ILS Glidepath Lobes

9. As with the localizer, the glidepath receiver compares the signal strength received from each lobe and deflects an indicator accordingly. The type of display found on an HSI is shown in Fig 5, which illustrates the three cases corresponding to positions of the aircraft in Fig 4. As with the localizer display, the centre of the indicator represents the aircraft, and the needle the glidepath.

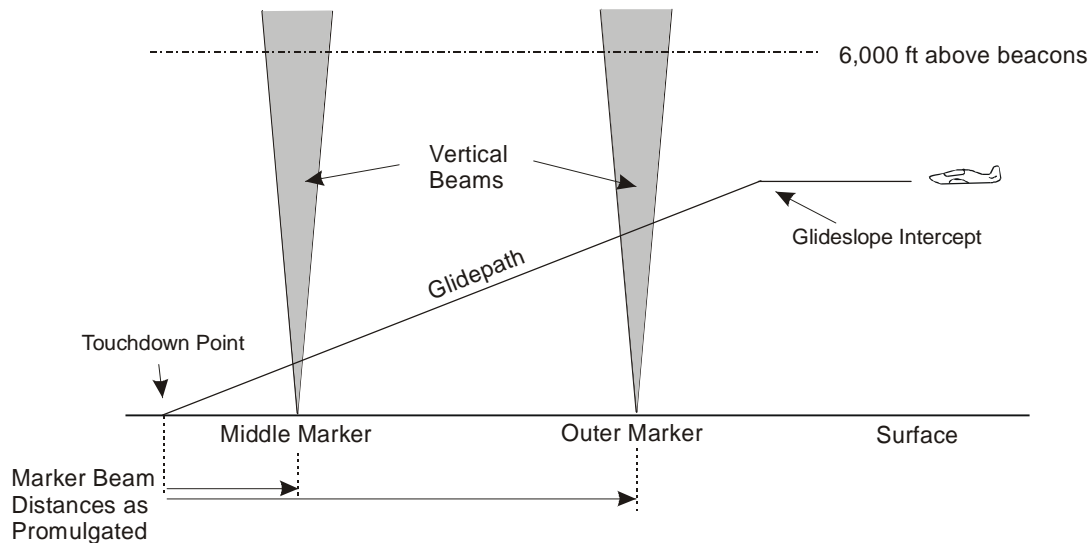
7-13 Fig 5 ILS Glidepath Indicator (HSI Display)

10. **Beamwidth.** The centreline of the glidepath beam is usually set at an angle of 3° but can be varied between 2° and 4° to suit local requirements. The beamwidth is a function of the beam angle (θ) and extends from 0.45θ to 1.75θ above the horizontal. For a 3° beam angle this equates to 1.35° to 5.25° above the horizontal. The edges of the beam correspond to full scale deflection of the indicator needle.

Marker Transmitters and Range Information

11. Limited information about the range from touchdown is provided by up to three marker beacon transmitters operating on a frequency of 75 Hz. The beacons are situated in line with the approach path and at fixed distances from the touchdown point (see Fig 6). These distances are promulgated to users in FLIPs. The marker beacons each radiate signals in a vertical fan-shaped beam, to a height of at least 6,000 ft above the beacons. As the aircraft crosses each fan beam, the marker identifies itself by audio signal through the pilot's headset and by visual lamp. The outer marker (usually situated approximately 4.5 nm from touchdown) transmits a series of dashes at 400 Hz. The middle marker is located approximately 3,000 ft from touchdown at the beginning of the runway bar lighting system. The transmission consists of alternate dots and dashes at 1,300 Hz. An inner marker, if employed, is situated about 1,000 ft from touchdown, and will transmit a series of dots at 3,000 Hz.

7-13 Fig 6 ILS Marker Beams



12. **Distance Measuring Equipment (DME).** A DME is sometimes frequency paired with the ILS system, such that DME selection is automatic when the localizer frequency is selected. The DME is automatically adjusted to provide a readout of range from touchdown. The DME is used to either supplement or replace the ILS marker beacons.

13. **Locator Beacons.** A locator beacon is a low-powered non-directional beacon (NDB), which can be collocated with an ILS marker beacon. At some airfields, the locator beacon may replace one of the marker beacons. The locator beacon can be used for tracking purposes when positioning the aircraft to commence an ILS approach. It can also be used for a holding pattern.

Procedure

14. Navigation to acquire the localizer may be accomplished using self-navigation techniques, radar vectors, or a terminal procedure. Typically, the localizer is intercepted at an angle of approximately 40°. When the localizer needle begins to centralize, the aircraft is turned on to the centreline and a heading is taken up to allow for drift. Fairly coarse corrections may be necessary to maintain the centre of the beam, which is quite wide in the early stages.

15. If approaching at 1,500 ft QFE, the aircraft will intercept the glidepath at 5 nm from touchdown, approximately at the position of the outer marker. The indicator needle will start to move down from its full deflection up position. The aircraft descent should be initiated so that the descent is established as the needle centres. Since the beam is only about 500 ft deep at this point the aircraft must be established on the correct descent path as quickly as possible.

16. As the runway is approached, the beams become more sensitive and accurate, steady flying is necessary to follow the indications; for example the glidepath beam is only 70 ft deep at a height of 200 ft. Particular care must be taken not to descend below the glidepath since this increases the risk of striking an obstruction, and the localizer signal is weaker in that region. Where the localizer is offset, the beam centreline is positioned so that it crosses the runway centreline at about 4,000 ft to 4,500 ft from the glidepath origin, coinciding approximately with a height of 200 ft. On transferring to a visual approach, the aircraft may have to be turned through a few degrees to achieve alignment with the runway centreline.

17. **Back-up Timing.** The middle marker often coincides with the Missed Approach Point (MAP). Where this is the case, the Terminal Approach Chart sometimes includes a table of time intervals, at

varying speeds, between the outer and middle markers. This will enable the use of a stopwatch, started overhead the outer marker, as a back-up for the MAP, in the event of a middle marker failure.

Interference

18. **Interference with ILS Reception.** The ILS system can be susceptible to unwanted reflected energy arriving at the airborne receiver, after re-radiation from terrain or airport buildings. Such interfering signals arriving at the ILS receiver may manifest themselves in one of two ways:

- a. **Scalloping.** Interference may be displayed as oscillating demands ('scalloping').
- b. **Beam Bend.** Interference may be present as a 'beam bend' in the course centreline.

To protect against interference from static objects, the accuracy of ILS centreline beams is checked by periodic airborne calibration. However, in certain mountainous areas, and locations with difficult topographical detail, successful ILS operation is impossible. Extreme interference can also be caused by re-radiation of the ILS signals from moving objects, particularly other aircraft taxiing close to the ILS installations, or aircraft taking-off.

19. **Interference from FM Transmissions.** More recently, high power FM broadcasts have been permitted in the band 88 - 108 MHz, adjacent to the localizer band. FM interference can manifest itself in the same forms as other interference. It can be overcome by including improved discrimination circuits in the aircraft ILS receiver.

20. **ILS Warning Flags.** At some stage, during the processing of received signals, the Localizer and Glidepath signals will each be sampled. Provided each signal exceeds an acceptable level, a separate signal voltage is generated to withdraw its associated warning flag (Glidepath or Localizer) on the ILS display. Therefore, the appearance of a warning flag will indicate only one of two possible conditions:

- a. The signal strength in the equipment is inadequate.
- b. The warning flag power supply has failed.

Warning flags are therefore not indicative of any interference. Whether the ILS equipment is protected against FM interference or not, aircrew should always carry out identification of the ILS transmission callsign, and monitor indications during the procedure, using all available cross-checks, including markers and DME information.

ILS Categories

21. **Facility Performance Categories.** ICAO grades the specification and performance of ILS ground facilities as Category I, II or III:

- a. **Facility Performance Category I.** A Category I ILS provides guidance down to 200 ft or less, above the ILS reference point.
- b. **Facility Performance Category II.** A Category II ILS provides guidance down to 50 ft or less, above the ILS reference point.
- c. **Facility Performance Category III.** A Category III ILS, with the aid of ancillary equipment where necessary, provides guidance down to, and along the surface of the runway.

22. **Operational Performance Categories.** The operational performance categories for aircraft approaches using ILS are shown in Table 1.

Table 1 Precision Approaches - Categories of Operation

| | |
|---|--|
| Category I | A Category I operation is a precision instrument approach and landing using ILS, Microwave Landing System (MLS) or Precision Approach Radar (PAR) with a decision height not lower than 200 ft and with a runway visual range (RVR) not less than 550 m. |
| Category II | A Category II operation is a precision instrument approach and landing using ILS or MLS with a decision height below 200 ft but not lower than 100 ft, and with a RVR of not less than 300 m. |
| Category IIIA | A Category IIIA operation is a precision instrument approach and landing using ILS or MLS, with a decision height lower than 100 ft, and a RVR of not less than 200 m. |
| Category IIIB | A Category IIIB operation is a precision instrument approach and landing using ILS or MLS, with a decision height lower than 50 ft, or no decision height and a RVR lower than 200 m but not less than 75 m. |
| Category IIIC | Category IIIC operation has no decision height limitation, and will provide guidance to, and along the runway and taxiways without reliance on external visual reference (see Note). |
| Note: Table 1 is a summary, based on the operational figures promulgated in JAR-OPS 1 (Subpart E) for Categories I to IIIB inclusive. Category IIIC is defined by ICAO International Standards and Recommended Practices (Annex 10, Vol 1). | |

23. **ILS Category II/III Operations.** To operate aircraft to the limits of Category II or III, the following requirements must be met:

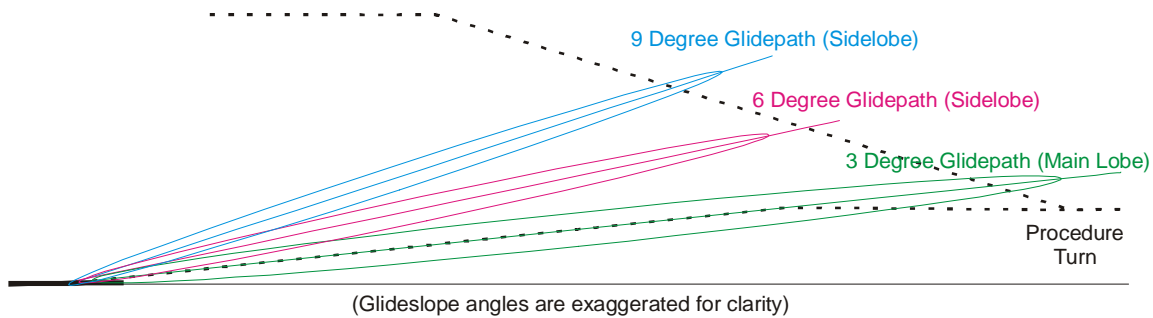
- a. The aircraft must be appropriately equipped and certificated.
- b. The aerodrome must be approved for such operations by the national authorities.
- c. The crew must be trained and qualified specific to the operation and the aircraft type.

24. **Low Visibility Procedures (LVP).** LVP is the term used to describe those procedures applied at an aerodrome for the purpose of ensuring safe operations during Category II and III approaches, and low visibility take-offs (where the RVR is less than 400 m). LVP may include special taxi routes, runway turn-off points and runway holding points; details will be promulgated in the AIP and FLIPs. To protect ILS signals, pre-take-off holding positions are more distant from the runway than the holding positions used in good weather. Taxiways lying within the ILS Sensitive Area are marked by a colour-coded taxiway centreline (alternate yellow/green lights). Pilots should avoid stopping their aircraft within the ILS Sensitive Area, and should make their "Runway Vacated" call only after the aircraft is clear of the Sensitive Area.

25. **Instrument Approach Criteria.** Instrument approach criteria for ILS equipped and cleared aircraft are promulgated in the Manual of Military Air Traffic Management (Chapter 1) and the Flight Information Handbook.

False ILS Glide Slope Capture

26. Different types of ILS glide slope systems are in use worldwide, with different signal characteristics in the area above the standard 3° glide slope. Sidelobes typically exist above the main 3° beam at 6° and 9°. If the ILS glide slope is intercepted from above the 3° beam, with the automatic flight control system engaged, the aircraft can capture a false glide slope within a sidelobe, resulting in an unexpected pitch-up command and causing the airspeed to drop towards a stall situation.

7-13 Fig 7 Typical ILS Glide Slope Lobe Pattern

27. Between the 3° and the 9° glide paths, the signal strength changes which can result in an observable movement of the ILS glide slope on the primary flight display. A signal reversal may be present in the 9° sidelobe and is sometimes present in the 6° sidelobe. This reversal activates the glide slope capture mode after which the autopilot follows the glide slope signal without restrictions. During flight tests the reversal resulted in the automatic flight control system commanding a severe pitch-up attitude.

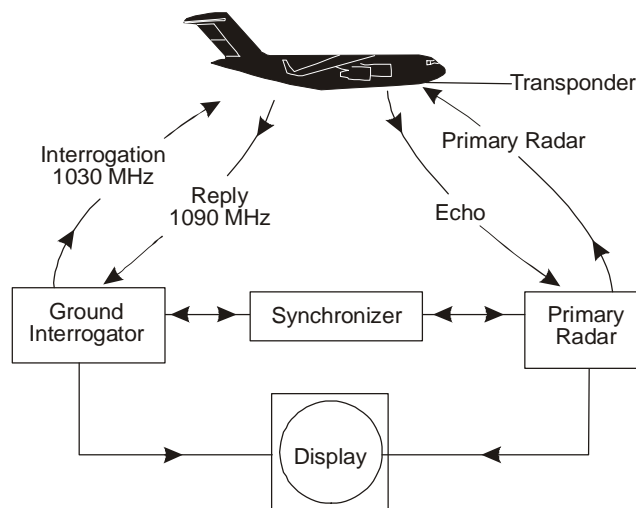
28. The ILS Glide Path is designed to be captured from below after aligning on the Localiser, either via ATC vectors or a published procedure. Ideally, if the aircraft is above the Glide Path, pilots should request an ATC reposition for a normal intercept and capture from below. If the operational circumstances prevent this, then a capture from above should only be done with extreme caution, using the expected aircraft height versus distance profile as a check, and only using basic flight director modes such as Vertical Speed until the aircraft is confirmed on the real Glide Path. Type specific procedures may limit rates of descent on an approach, and flying it as a non-precision approach, obeying any step-fixes and the MDH/A, should be considered.

CHAPTER 14 - IFF/SSR

Introduction

1. The military system, Identification, Friend or Foe (IFF), and the equivalent civilian system, Secondary Surveillance Radar (SSR), are examples of secondary radar systems. Whereas a primary radar system depends on receiving a radar echo reflected passively by a target, in a secondary radar system a transmitted radar signal is used to trigger a response from equipment in the target.
2. IFF/SSR is used by air defence (AD) and air traffic control ground installations, and increasingly by AD aircraft, to identify radar returns. The systems are normally operated in conjunction with a primary radar, the two aerials being either co-mounted, or arranged to scan synchronously. The arrangement of a typical system is shown schematically in Fig 1. The ground (or AD aircraft) transmitter, known as the interrogator, transmits a coded interrogation signal which is received and decoded by transponders in friendly aircraft. Depending on the mode to which the transponder is set, a coded reply signal is transmitted back to the interrogator. This reply signal is decoded and shown on the radar display along with the primary radar response.

7-14 Fig 1 Secondary Identification and Primary Radar Systems



3. Although IFF and SSR have a number of differences, the operating principle of each is basically the same and RAF aircraft are fitted with transponders which can operate with both systems.
4. **Operating Modes.** Military IFF has three operating modes, 1, 2, and 3. Civilian SSR has potentially five modes, A, B, C, D, and S, although Mode D is not currently used. IFF Mode 3 and SSR Mode A are identical and normally referred to as Mode 3/A.

Interrogation Signal

5. The basic interrogation signal is transmitted at a frequency of 1030 MHz and consists of a pair of pulses, each pulse having a width of 0.85 μ s. The separation of the pulses determines the mode of interrogation as shown in Table 1.

7-14 Table 1 Interrogation Signal Formats

| <i>Mode</i> | <i>Pulse Pair Spacing (μs)</i> |
|-------------|---|
| 1 | 3 |
| 2 | 5 |
| 3/A | 8 |
| B | 17 |
| C | 21 |
| D | 25 |
| S | Multiple Pulses |

Normal Replies

6. If the transponder has been set to LOW or NORM, it will reply to an interrogation signal provided that the appropriate mode has been selected.

7. A normal reply is transmitted on a frequency of 1090 MHz and consists of a pair of pulses (known as the framing pulses) separated by 20.3 μ s, with up to twelve information pulses in between. The framing pulses are referred to as F1 and F2, while the twelve information pulses are in four groups of three, designated A, B, C and D. Within a group, the pulses are annotated 1, 2, and 4, (A1, A2, A4; B1, B2, B4 etc), each pulse representing one digit of a three-digit binary number. In this way, the presence or absence of pulses allows each group to represent a decimal number from 0 to 7. For example, in group A:

| A4 | A2 | A1 | | |
|----|----|----|-----------------|-----|
| 0 | 0 | 0 | = No A pulse | = 0 |
| 0 | 0 | 1 | = A1 pulse only | = 1 |
| 0 | 1 | 0 | = A2 pulse only | = 2 |
| 0 | 1 | 1 | = A2 + A1 | = 3 |
| 1 | 0 | 0 | = A4 pulse only | = 4 |
| 1 | 0 | 1 | = A4 + A1 | = 5 |
| 1 | 1 | 0 | = A4 + A2 | = 6 |
| 1 | 1 | 1 | = A4 + A2 + A1 | = 7 |

It is thus possible to transmit 4096 codes, the code being set on the controller. For example, to transmit a code of 4167 the transmitted pulses would be:

| A4 | B1 | C2 + C4 | D1 + D2 + D4 |
|----|----|---------|--------------|
| 4 | 1 | 6 | 7 |

The transmitted code is displayed on a digital indicator alongside the PPI radar display.

8. There are a total of seven normal reply modes:

- a. **Military Mode 1.** Military Mode 1 replies comprise the framing pulses and the information pulses reflecting the Mode 1 code set on the cockpit control panel.
- b. **Military Mode 2.** Military Mode 2 has the same form as Mode 1. However, the code is not selectable in flight, but is preset on the transponder unit.

- c. **Common Mode 3/A.** Mode 3/A has the same form as Mode 1, but the controller has a separate set of code selection switches so that replies can be made to Mode 1 and Mode 3/A simultaneously. Mode 3/A is the mode normally used by ATC agencies to establish and maintain the identity of an aircraft, to assist in the transfer of control between agencies, and to supplement primary radar information.
- d. **Civil Mode B.** Mode B is the same as Mode 3/A but can only be used as an alternative to the latter.
- e. **Civil Mode C.** Mode C is used for the automatic reporting of altitude. The transponder, in association with an encoding altimeter, replies with a code train indicating the aircraft's height relative to a 1013.25 mb pressure datum. The code uses 11 of the 12 information pulses and a change occurs every 100 ft. Mode C is used by air traffic controllers to confirm that aircraft are maintaining, vacating, reaching, or passing assigned flight levels, and to monitor the vertical separation between transponding aircraft, without recourse to ground/air communication.
- f. **Civil Mode D.** Mode D is not currently used.
- g. **Civil Mode S.** Mode S was introduced to support the automation of some ATC functions. Its full title is 'Mode Select' and it provides a two-way data link facility. Mode S has all the functionality of Modes A and C, in addition to providing other information (Mode S transponders can be installed on their own to replace Modes A and C). One feature of the system is that each aircraft fitted with Mode S is assigned a unique address code. The address, together with the other information, is transmitted once per second in a signal known as a 'squitter'. This signal can be received by ATC units and other Mode S capable aircraft. Mode S is a required component of the Traffic Alert and Collision Avoidance System (see Volume 7, Chapter 16), now mandatory in many ATC regions.

Identification Replies

9. In addition to the normal replies discussed so far, identification replies may be initiated, when required, by a switch on the control panel. These identification replies consist of the normal code followed by a pulse, or pulses, which is transmitted for 20 seconds after operation of the switch. The switch is spring-loaded to the off position and is normally marked I/P or IDENT. Transmission of the signal enables the ground controller to carry out rapid identification of a particular aircraft among the many which may appear on the display, operating in the same mode.

10. **Military Ident Reply.** This reply is given only on military Mode 1 when the I/P switch is operated. It consists of the selected information pulse train followed by a second identical pulse train with the second F1 pulse 4.35 μ s after the first F2 pulse.

11. **Civil Ident Reply.** Civil identification replies are given on military Mode 2, common Mode 3/A, and civil Mode B when the identification facility is selected. It consists of a normal information code train followed by a further pulse, known as the SP1 pulse, 4.35 μ s after the F2 pulse.

Emergency and Special Purpose Replies

12. Operation of the emergency facility allows an emergency reply to be transmitted on military Modes 1 and 2, common Mode 3/A, and civil Mode B. This reply consists of a coded information pulse train followed by three repeats containing the framing pulses only, with no information pulses. On common

Mode 3/A or civil Mode B the information pulses contained in the first frame depend on the position of a civil/military emergency switch. With the switch set to CIV, the first frame consists of an automatic transmission of code 7700. With the switch to MIL, the first frame consists of the code which is already selected. In reply to military Modes 1 and 2, the first frame consists of the selected information pulse code. No emergency reply is transmitted on civil Mode C. Selection of the emergency facility overrides identification replies. Additionally, Code 3/A 7500 (unlawful interference) and 7600 (communications failure) are internationally recognized.

R/T Phraseology

13. A full list of standard phrases used in connection with IFF/SSR is published in Civil Aviation Publication Radiotelephony Manual (CAP 413). The more frequently used standard phrases are shown in Table 2.

Table 2 Standard ATC IFF/SSR Phraseology

| <i>Phrase</i> | <i>Meaning</i> |
|--|---|
| Squawk (code) | Set the code as instructed |
| Confirm squawk | Confirm the code set on the transponder |
| Reset squawk (code) | Reselect assigned code |
| Squawk Ident | Operate the special position identification feature |
| Squawk Mayday | Select Emergency |
| Squawk Standby | Select the standby feature |
| Squawk Charlie | Select altitude reporting feature |
| Check altimeter setting and confirm (level) | Check pressure setting and confirm your level |
| Stop squawk Charlie | Deselect altitude reporting |
| Stop squawk Charlie, wrong indication | Stop altitude report, incorrect level readout |
| * Confirm (level) | Check and confirm your level |
| ** Check selected level. Cleared level is (correct cleared level) | Check and confirm your cleared level |
| Confirm you are squawking assigned code (code assigned to the aircraft by ATC) | To verify that 7500 has been set intentionally |

Notes:

*Used to verify the accuracy of the Mode C derived level information displayed to the controller.

**Where selected flight level is seen to be at variance with an ATC clearance, controllers shall not state on the frequency the incorrect SFL as observed on the situation display. However, controllers may query the discrepancy using this phraseology. For ATC purposes, the generic phrase 'selected level' is used to encompass both altitude and flight level.

SYSTEM REFINEMENTS

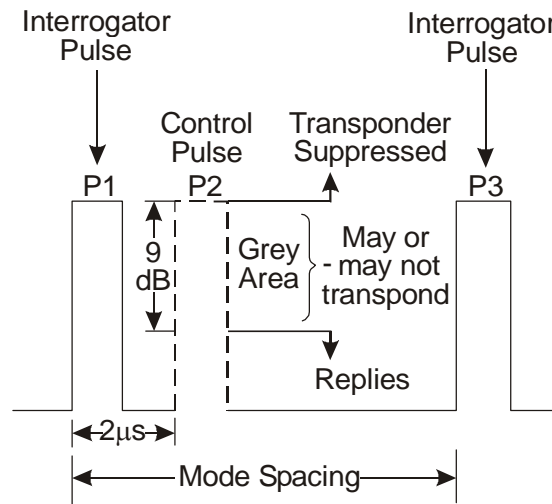
Three-pulse Sidelobe Suppression

14. In secondary radar, sidelobes are effective at greater ranges than in primary radar since transponder transmissions are detected rather than target echoes. It is thus necessary to suppress

any interrogator sidelobes which would be capable of triggering responses and the 3-pulse sidelobe suppression system has been adopted as the international standard technique.

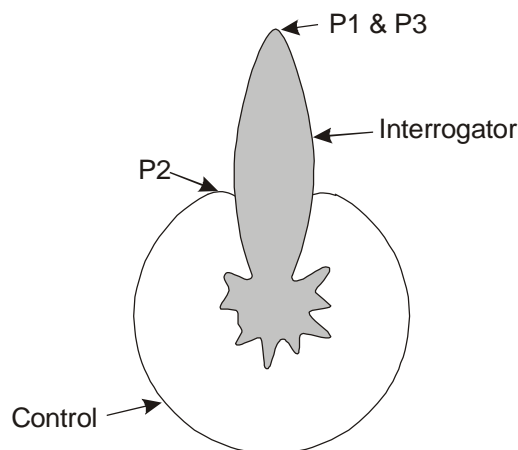
15. Each interrogation consists of a group of three pulses denoted P1, P2, and P3. The spacing of the pulses is shown in Fig 2; P2, known as the control pulse, is spaced at a constant $2\ \mu\text{s}$ from P1, and the spacing between P1 and P3, the interrogation pulses, is according to the mode as listed in Table 1.

7-14 Fig 2 3-Pulse Sidelobe Suppression - Pulse Spacing



16. The interrogation pulses P1 and P3 are transmitted from a rotating interrogator aerial, and the interrogation pulse P2 is transmitted from the control aerial, producing radiation patterns as shown in Fig 3. The interrogator aerial produces a high-power, narrow beam with lower-power sidelobes. The control aerial radiates an omnidirectional signal which is modified to produce a radiation pattern trough in the direction of interrogation.

7-14 Fig 3 3-Pulse Sidelobe Suppression - Aerial Polar Diagrams



17. By comparing the relative amplitudes of the control and interrogator pulses, the transponder determines whether the interrogation is a correct, main lobe one, or due to a sidelobe. P1 and P3 amplitudes are greater than a P2 amplitude only for a correct main lobe interrogation and the transponder will then reply; if P2 is greater than P1 and P3 the transponder is suppressed. The discrimination levels are indicated on Fig 2 which shows that there is a 'grey area' where the transponder may or may not reply.

Defruiting

18. Since all IFF/SSR equipments work on the same transmit and receive frequencies, any interrogator can trigger any transponder which is within range and selected to the appropriate mode. Thus any ground station can receive replies from transponders interrogated by other nearby ground stations. These unwanted replies appear as interference or 'fruit'. Defruiting is the process whereby this interference is removed. Adjacent interrogators are operated at different pulse recurrence frequencies (PRFs) and comparator circuits only pass replies at the correct home station PRF.

Garbling

19. Because the length of a transponder code train is about 20 μ s it is not always possible to decipher replies from aircraft closer than 2 or 3 miles to each other on a radial from the interrogator. The reply signals may garble and the decoder equipment can cause the generation of false targets between the aircraft or cause cancellation of all or part of either or both actual returns. False targets or cancellation may occur even though altitude separation between aircraft exists. Circuits in the decoder equipment are used to cancel garbled replies, and controllers will often ensure that only one aircraft within a formation has a transponder operating.

Mode Interlacing

20. In order to use the different modes for various functions, it is necessary for them to be transmitted separately from each other, and on a sequential basis. The mixing of the mode transmissions is known as mode interlacing. Each mode is selected at its PRF rate, and each mode sequence is selected at aerial rotation rate. Generally speaking, the use of more than three-mode interlace is not satisfactory operationally, since the number of hits per scan for each mode transmitted falls to a non-effective level. However, the interlacing of four or five modes can be achieved if necessary.

CHAPTER 15 - GROUND PROXIMITY WARNING SYSTEMS

Introduction

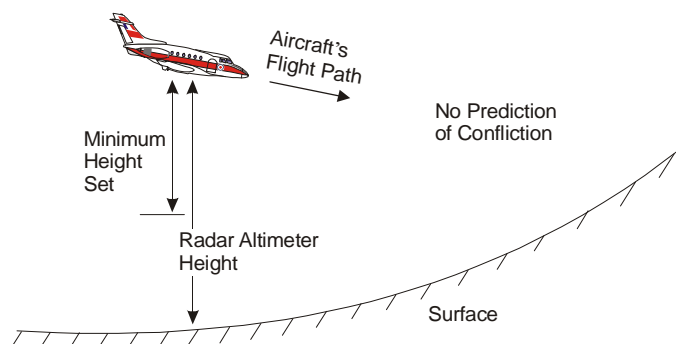
1. Ground Proximity Warning Systems (GPWS) are designed to aid aircrew with their situational awareness, specifically with respect to the terrain below and ahead of the aircraft. Some advanced GPWS also give advice of potential conflict with obstacles, and can be used for secondary functions such as navigation solutions and weapon aiming.

Non-predictive Ground Collision Avoidance Systems

2. The first generation of GPWS consist of little more than a radar altimeter (radalt), as described in Volume 5, Chapter 3. In such a system, the radalt obtains an accurate vertical measurement of the distance from the aircraft to the Earth's surface immediately below it. This height measurement is compared, continually, with a Minimum Height (MH), preset by the crew. When the radalt pointer indicates a height less than the MH, a set of electrical contacts will close, resulting in the illumination of a low-level warning lamp. These contacts may also operate an audio warning tone.

3. However, such a system works on an instantaneous height reference, based immediately below the aircraft. It is incapable of predicting the relationship of the aircraft's flight path and the terrain ahead of the aircraft (see Fig 1). In the case illustrated, there would be no warning until the radalt height is less than the MH set. Such a system is known as a Non-predictive Ground Collision Avoidance System.

7-15 Fig 1 A Simple GPWS with no Predictive Capability

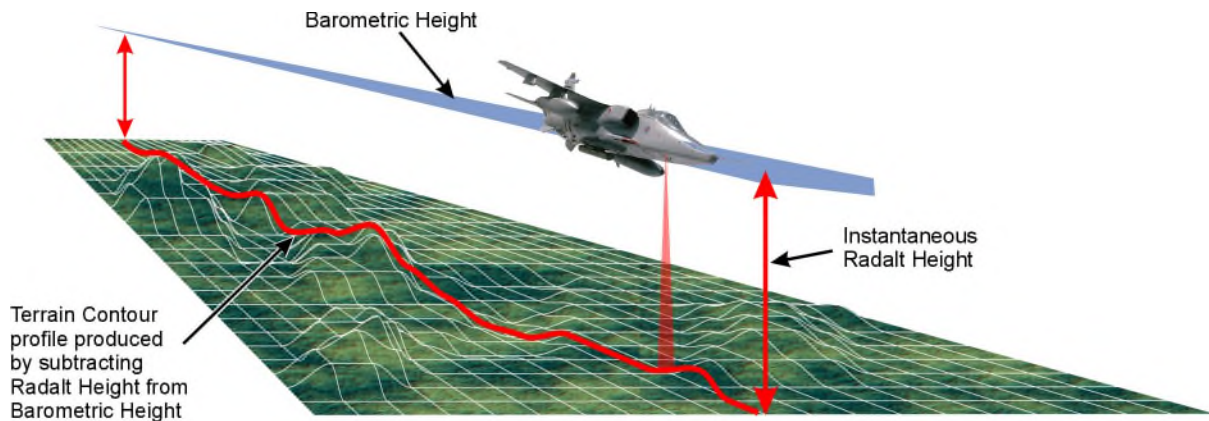


4. The development of terrain-referenced navigation systems, and digital terrain databases provided the impetus for more capable and advanced GPWS.

TERRAIN-REFERENCED NAVIGATION SYSTEMS

Introduction

5. Terrain-referenced navigation (TRN) was developed in the 1960s and 1970s, to provide US cruise missiles with a means of updating their INS guidance systems. A TRN system obtains a sequence of radalt readings, which, when subtracted from the barometric height, produces a sequence of ground heights (see Fig 2). The system then searches a digital terrain database for a match against the radalt-generated data. When a precise match is found, the missile's position can be established. This missile system was called TERrain-CONtour Matching (TERCOM).

7-15 Fig 2 The Principle of Operation of a simple Terrain-referenced Navigation System

6. Because the TRN utilises the changing height profile to determine horizontal position, TRN works best when the terrain is of reasonable undulation. Extremely flat terrain would lead to a larger circle of uncertainty for position. For this reason, TERCOM is integrated with an inertial navigation system (INS). The TRN will take planned 'waypoint fixes' in areas of distinctive topography, with the INS steering from one fix area to another. From the results of the planned 'waypoint fixes', corrections can be made to the missile's navigation system.

7. TRN systems are not reliant on other external aids, and are unaffected by jamming environments.

TERPROM

8. One of the more recent, commercially developed TRN systems is TERPROM (TERrain PROfile Matching). This uses a Kalman filter-based model of the navigation system to examine the terrain database and predict the next radalt reading. The difference between the predicted and actual readings is then fed back as an error signal to refine the system accuracy. TERPROM continues to be developed, and can now provide obstacle/terrain avoidance, navigation and weapon aiming solutions. TERPROM is primarily a software product, and can be hosted on an aircraft's mission computer hardware. It will produce a quoted navigational accuracy of better than 30 metres circular error probable (CEP) horizontally and 5 metres root-mean-square (RMS) vertically, when flying at low level (below 5,000 feet).

The Digital Terrain Elevation Database (DTED)

9. All TRN systems require a terrain database in digital form. The main agency developing and compiling such databases for military applications is the US National Imagery and Mapping Agency (NIMA). NIMA has lead the development of a series of standard digital datasets, termed Digital Terrain Elevation Data (DTED).

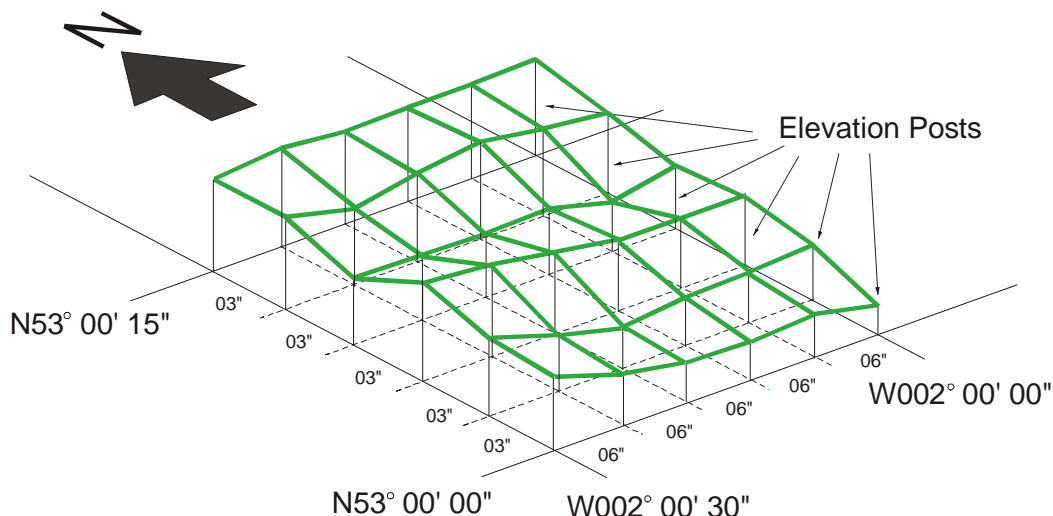
10. A DTED is a geographic matrix of terrain elevation data points at precise increments of latitude and longitude. These data points are converted into a numerical format for computer storage and analysis.

11. **DTED Data Files.** The DTED data files are structured by geographical areas, each 'cell' containing data within an area of one degree of latitude by one degree of longitude. To provide overlap between adjacent data files, the cell coverage includes the integer degree values on all sides of the cell.

12. **Elevation Posts.** Within each cell, the terrain elevation is expressed in metres, at precise points. Each point of elevation is known as an elevation 'post' (Fig 3). The locations of elevation posts are defined by the

intersection of rows and columns within the matrix. The required matrix intervals, defined in geographic arc seconds, may vary slightly within each data set, dependent upon the latitude.

7-15 Fig 3 Illustration showing a sample of Digital Terrain Elevation Database



13. **DTED Levels.** By varying the horizontal distance between elevation posts, different levels of DTED standard can be defined. The present DTED levels are:

- a. **DTED Level 0.** DTED Level 0 has an elevation post spacing of 30 arc seconds. It was generated with the support of select international mapping organizations, and is of value to scientific, technical and other communities. It allows a gross representation of the Earth's surface for general modelling and assessment activities. It is not intended for any precision activity involving the safety of the public.
- b. **DTED Level 1.** DTED Level 1 is a medium resolution elevation data source for military and other activities, and is a uniform matrix with post spacing every 3 arc seconds. The information content is approximately equivalent to the contour information represented on a map of 1:250,000 scale.
- c. **DTED Level 2.** DTED Level 2 is a high resolution elevation data source for military and other activities, and is a uniform matrix with post spacing of one arc second. The information content is approximately equivalent to the contour information represented on a map of 1:50,000 scale.
- d. **DTED Level 3, 4 and 5.** At present, DTED levels 3, 4 and 5 are only proposals.

Table 1 summarizes post spacings for the various levels of DTED.

Table 1 DTED Levels

| DTED Level | Post Spacing | Approx Ground Distance |
|------------|--------------------|------------------------|
| 0 | 30 seconds arc | ~ 1,000 m |
| 1 | 3.0 seconds arc | ~ 100 m |
| 2 | 1.0 second arc | ~ 30 m |
| 3 | 0.333 seconds arc | ~ 10 m |
| 4 | 0.111 seconds arc | ~ 3 m |
| 5 | 0.0370 seconds arc | ~ 1 m |

14. *Thinning of Databases.* To save software storage, a terrain database can be 'thinned' in size. In this process, selected columns or rows of elevation posts, deemed not to be significant, can be deleted.

PREDICTIVE GROUND COLLISION AVOIDANCE SYSTEMS

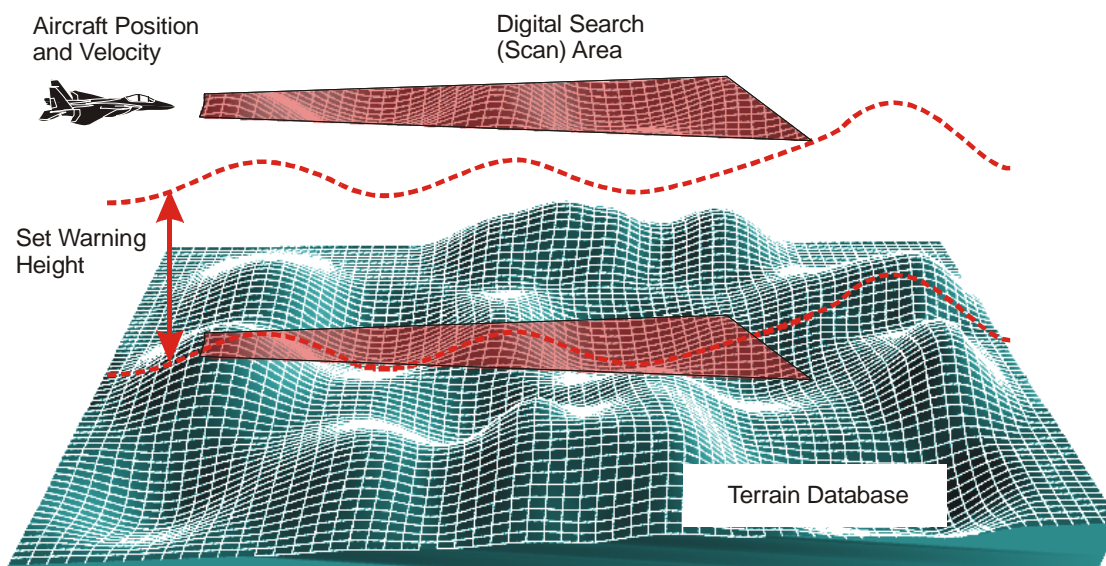
Requirements of a Predictive System

15. The latest GPWS are capable of analysing:

- a. The aircraft's position in space.
- b. The aircraft's instantaneous velocities.
- c. The terrain within a specified search area ('scan' area) ahead of the aircraft.

By carrying out comparisons with this data, it is possible to predict imminent conflict with the terrain. This function is called a Predictive Ground Collision Avoidance System (PGCAS). The basic requirements of a PGCAS are illustrated in Fig 4 and described in the following paragraphs.

4-29 Fig 4 Requirements of a Typical PGCAS



16. **Aircraft Position and Velocity.** The aircraft's position and kinematic state will be provided by the aircraft's navigation system, which may be based upon:

- a. Inertial Navigation System (INS).
- b. Global Positioning System (GPS).

Most systems will use a mixture of sensor sources to allow a measure of redundancy.

17. **Database Storage.** The terrain database is held in digital format in the aircraft. The aircraft's concept of operations will dictate the requirements for the level of terrain data held. A typical fast-jet will require the most precise and concentrated database available, but only over a limited theatre of operations. If deployed to another geographical area, then a 'theatre upload' of terrain database can take place. On the other hand, a strategic transport aircraft might require a world-wide database on-board at all times. Even in this latter case, memory storage space might be limited, therefore the data density level will have to be varied, to be less concentrated for 'en route' areas, and perhaps highly concentrated in the surrounds of terminal airfields.

18. **Scan Area.** The PGCAS must mathematically search an area of the digital terrain database, forward from the aircraft's present position. The size and shape of this 'scan' area is defined by the PGCAS. In some systems, the size and shape of the scan area will be responsive to:

- a. **Groundspeed Changes.** The faster the aircraft's speed, the longer the length of the scan area.
- b. **Navigational Accuracy.** Any variations in the estimated accuracy of the positional output from the navigational system may be used to adjust the scan pattern or size.
- c. **Turns.** A scan area based purely ahead of the aircraft is insufficient if predictive warning is required during turning manoeuvres. A sophisticated system will expand the width of the scan area to look into the turn, in addition to monitoring directly ahead.

19. **Set Warning Height.** A PGCAS will have the capability to store a Set Warning Height (SWH). This SWH might be set by:

- a. **The pilot.** The pilot will be able to set the required warning height to suit the required ground clearance. A cockpit indication will show the height set.
- b. **The Aircraft.** The PGCAS might store and utilize a series of built-in minimum heights, dependent upon the stage of flight.

In some PGCAS, the SWH is automatically factored for safety; it will be increased if the navigation system detects inaccuracy in either position or velocity data.

20. **Warnings.** If the predicted aircraft flight path penetrates the SWH, the PGCAS will give warning to the crew by Audio and/or Visual means.

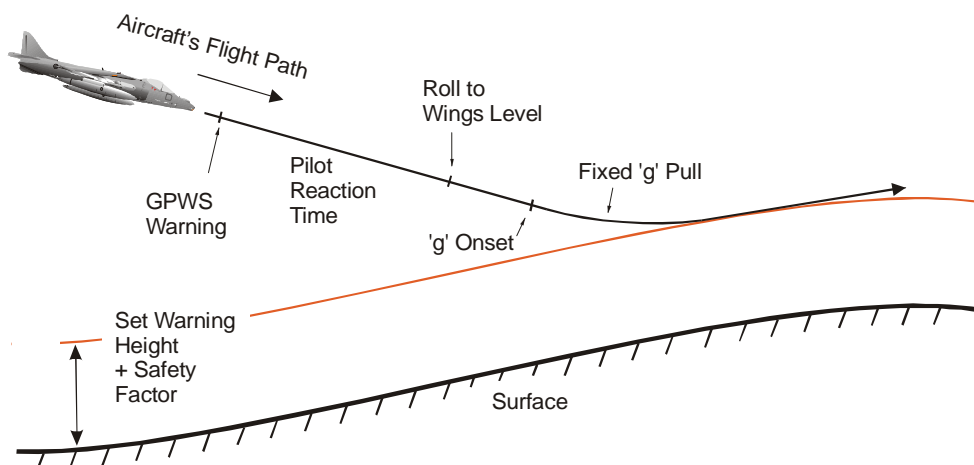
Operation of the PGCAS

21. The PGCAS uses the known aircraft position, the aircraft's velocities, the scan area and the stored terrain data to provide a warning whenever the aircraft is going to penetrate the SWH.

22. The PGCAS carries out continuous calculations based upon the following (see Fig 5):

- a. The profile of the terrain ahead of the aircraft, increased according to the SWH and known obstacles.
- b. A timed recovery manoeuvre based upon:
 - (1) A set period to allow for the pilot's reactions.
 - (2) A time for the aircraft to roll to wings level, using a set roll rate.
 - (3) A recovery 'pull' at a pre-set 'g' level, with an allowance for a specific 'g' onset rate.
- c. The time remaining on present flight parameters before the recovery manoeuvre intersects the SWH/terrain profile.

As the time remaining, at c above, approaches zero, a cautionary warning can be generated. When the time variable reaches zero, a survival warning is generated, and the pilot must initiate a pull-up.

7-15 Fig 5 Operation of the Predictive Ground Collision Avoidance System

23. The PGCAS does not make direct use of the radalt data, therefore warnings can be provided regardless of aircraft attitude, turn rate or roll rate.

24. **Worst-case Profile.** As previously mentioned, the SWH, illustrated in Fig 5, may be increased by a safety factor if errors are detected in position or velocity information. In addition, any known obstructions can be added to the terrain (see para 30). In some PCGAS, this amended SWH profile is referred to as the 'worst-case profile'.

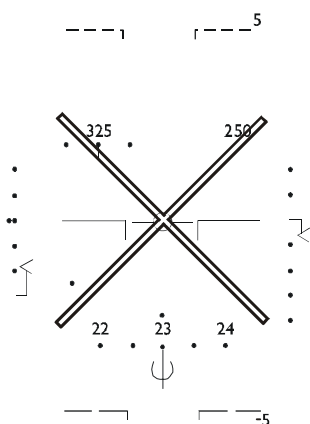
Crew Warnings

25. The GPWS can give warnings to the crew in either audio or visual formats.

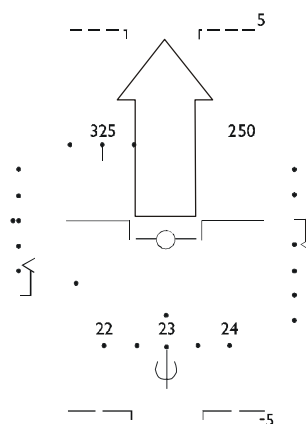
- a. **Audio Warnings.** Audio warnings might be in the form of a warning tone or as a pre-recorded voice message. The voice message is generally chosen as the prime method in most GPWS. A typical voice warning might say "PULL-UP, PULL-UP".
- b. **Visual Cues.** Visual warnings can be presented as cues in the Head-up Display (HUD). The cue might take the form of a breakaway cross (Fig 6a), or an arrow indicating direction to fly in order to achieve safe terrain separation (Fig 6b).

7-15 Fig 6 GPWS HUD Warnings

a Breakaway Cross Cue



b Arrow Cue



26. **Audio Levels.** An audio warning must be capable of being heard, even over aircraft and cockpit noise levels in a busy environment. During evaluation trials, assessment can be made of the

background noise for a specific aircraft type. A volume setting can then be established, giving 100% chance of a warning being heard.

27. **Nuisance Warnings.** Due to the design constraints within a GPWS, it is possible that the crew will continue to receive warnings, even though they have taken action to remedy the situation. A continual flow of audio warnings can be a nuisance, and provides potential for distraction, even to experienced aircrew. For this reason, the GPWS can usually be manually deselected if required.

System Failures

28. The predictive system does not rely on a continuous input from the radalt. Therefore, the PGCAS ground proximity warnings will still be provided, irrespective of radalt status.

29. The system will revert to a non-predictive system if:

- a. The GPWS leaves the geographical coverage of DTED.
- b. Position or velocity sensor failures are detected (this reversionary mode does require a serviceable radalt).

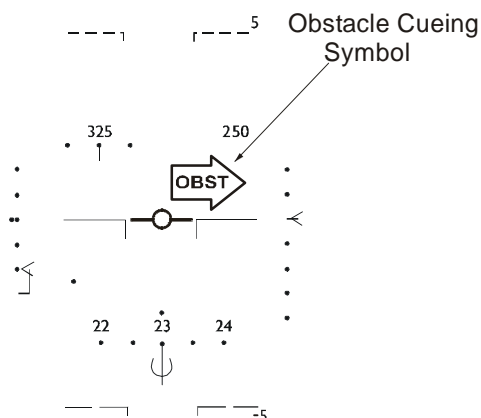
SECONDARY FUNCTIONS OF THE GPWS

Obstacle Warning and Cueing

30. Some GPWS are capable of providing warnings to the pilot, to advise of proximity to vertical obstructions. Details of the location and elevation of known obstructions (man-made features such as radio masts, chimneys, etc) will be held in digital format in files that are separate to the main DTED. This obstruction data can be included as part of the PGCAS to add obstacle avoidance into the predictive avoidance. In addition, obstruction warnings can be used to assist low level operations, in order to increase aircrew situational awareness.

31. **Digital Vertical Obstruction Files.** The obstruction data within the TRN is provided by the Defence Mapping Agency (DMA) by means of the Digital Vertical Obstruction File (DVOF). The DVOF will be updated regularly, and will contain all known obstacles above a stated height above ground level, plus a selection of smaller obstacles. The DVOF can be amended or changed quickly to meet theatre requirements.

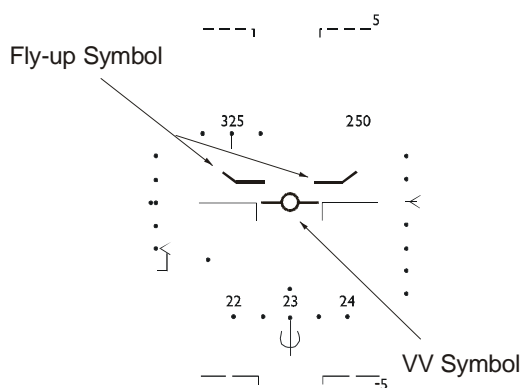
32. **Obstacle Warning and Cueing.** The obstacle warning and cueing (OWC) facility searches the database within its own specified scan area. When an obstacle is detected ahead of, or to either side of the aircraft, the system will generate audio and/or visual cues. An audio cue might be a voice warning saying "OBSTRUCTION". A visual cue can be presented in the HUD. A simple HUD format is to use just a text warning ("OBST") for obstacles directly ahead, or combined with an arrow showing that the obstacle is left or right of the aircraft's fore-aft axis. The arrow cue will not point directly at the obstruction; it merely advises the pilot which side of the nose he must look to search visually for the obstruction. Fig 7 shows a HUD indicating an obstruction to the right. More complex HUD displays may overlay a warning marker on an image of the obstacle.

7-15 Fig 7 HUD Obstacle Cueing

33. **Obstacle Priority System.** A priority system will be built into the obstacle cueing, such that the HUD displays the most significant obstacle at that moment (this might be the largest obstruction, or the one closest to track). Because multiple obstacles may lie within the scan area, but only the priority one is cued, it is important for aircrew to scan visually for other obstructions, as normal. Furthermore, although the OWC will indicate obstructions such as radio masts, it may not account for any associated guide wires.

Terrain Avoidance Cueing

34. Some GPWS systems have a terrain avoidance cueing (TAC) mode. The TAC mode provides pitch commands to give an indication as to how close the aircraft is to the minimum height setting. The pitch commands can be sent to the autopilot for automatic terrain following, or to the HUD for the pilot to follow manually. The HUD illustration in Fig 8 shows a 'fly-up' symbol above the aircraft's Velocity Vector (VV) symbol. The TAC facility is a passive software solution. As such, it not to be confused with active warning systems like Terrain Following Radar.

7-15 Fig 8 HUD Terrain Avoidance Cueing

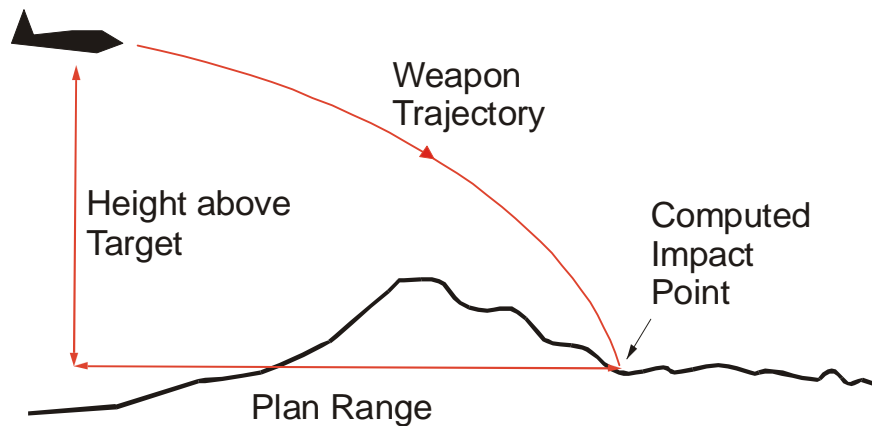
Navigation Solution

35. As described in paras 5 to 7, the TRN component of the GPWS is capable of determining the aircraft position in 3D. The TRN can therefore be used as a stand-alone sensor, or be integrated with the other navigation sensors, to produce a navigation solution giving a high level of accuracy in horizontal and vertical planes.

Weapon Aiming

36. The TRN can provide a passive target ranging facility, based on its accurate knowledge of aircraft height and target elevation. The system can calculate an instantaneous range from the aircraft to any point in the terrain database (Fig 9). This range can be used for weapon aiming calculations. Additionally, it can be integrated with some forward-looking sensor or HUD to provide the range, bearing and elevation of a target in the pilot's field of view. Such a system has the advantages of being totally self-contained and resistant to countermeasures.

7-15 Fig 9 Weapon Aiming Function



SUMMARY

General

37. The combination of trade names and abbreviations concerned with GPWS operation can appear confusing (e.g. GPWS, E(nhanced)GPWS, TAWS (Terrain Awareness and Warning System)). However, a predictive GPWS must employ the principles described within this chapter, ie to compare the aircraft's position and velocity with an area of digital terrain database ahead of the aircraft, to determine whether the aircraft is about to intercept a pre-set minimum height.

38. The ability to choose and mix component parts gives an endless permutation for design. The precise components of a GPWS will depend upon the manufacturer's choice for the TRN system, navigational sensors and radar altimeter. A typical fit for a high-performance military front-line aircraft might be:

- a. INS/GPS mix, for position and velocity sources.
- b. Radar Altimeter.
- c. TERPROM, hosted on a Terrain Processing Module in the mission computer.
- d. DTED Level 1.
- e. A DVOF.

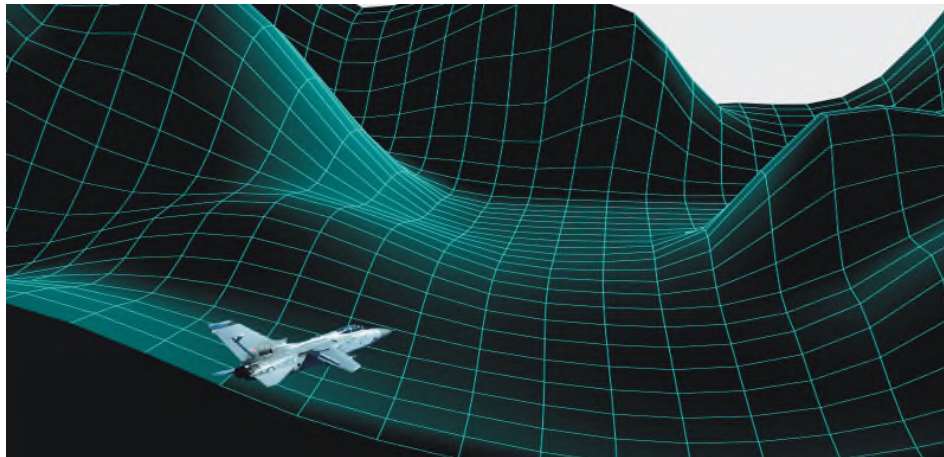
39. A GPWS system in a civilian transport aircraft will not require the full TRN elements and functions needed for military aircraft.

Future Developments

40. GPWS and TRN systems are developing continuously. Terrain database information can be used for many purposes, and might, for example, be displayed to aircrew in plan format, on colour screens.

41. It is possible to produce a three-dimensional lattice by joining the tops of the DTED elevation posts. In this manner, a synthetic view of the terrain ahead of the aircraft can be constructed (see Fig 10). It may be possible to display this terrain as ridge lines, superimposed on the HUD, or other optical displays.

4-29 Fig 10 Illustration of a Synthetic Terrain Matrix



CHAPTER 16 - AIRBORNE COLLISION AVOIDANCE SYSTEMS

Introduction

1. The increasing intensity of air traffic has led to an increased risk of an airborne collision. In an effort to improve airborne safety, most civil aircraft and increasingly more military aircraft are fitted with equipment designed to provide collision warning. Airborne Collision Avoidance Systems (ACAS) are designed to provide collision avoidance protection and airspace situational awareness between aircraft independent of air traffic control. The principal benefit of ACAS as a 'last resort' collision avoidance system is enhanced safety rather than operational benefits. A Traffic Alert and Collision Avoidance System (TCAS) is a type of an Airborne Collision Avoidance System. A detailed description of ACAS can be found in the International Civil Aviation Organization (ICAO) Document 9863, Airborne Collision Avoidance System (ACAS) Manual.

TRAFFIC ALERT AND COLLISION AVOIDANCE SYSTEM (TCAS)

General Description

2. TCAS utilizes Secondary Surveillance Radar (SSR) technology and is an example of a secondary radar system. The system operates without any reference to ground-based systems and utilizes the transmissions made by aircraft on IFF frequencies for ATC purposes. TCAS monitors the airspace around an aircraft for other aircraft equipped with a corresponding active transponder. It warns pilots of the presence of other transponder-equipped aircraft, which may present a threat of mid-air collision. The equipment interrogates the transponders of other aircraft and using the replies, tracks the slant range, altitude (Mode C) and relative bearing of each contact. Using successive replies, TCAS calculates the time to reach the Closest Point of Approach (CPA) with the intruder aircraft by dividing the range by the closure rate. The time to CPA is the parameter used for issuing alerts. With Mode C equipped aircraft, TCAS also calculates the time to co-altitude.

Capability Levels

3. There are three levels of capability within TCAS technology:
 - a. **TCAS I.** TCAS I provides the aircraft crew with bearing and distance information of a possible hazard, which helps them to visually identify the traffic by issuing a Traffic Advisory (TA) alert (see para 6a). The TA display will indicate the traffic's range and bearing but will not recommend an escape manoeuvre.
 - b. **TCAS II.** TCAS II provides the same information as TCAS I, but, in addition, gives a recommended vertical escape manoeuvre by issuing a Resolution Advisory (RA) alert (see para 6b). Although bearing information is provided on the traffic, it is only to aid visual identification.
 - c. **TCAS III.** TCAS III is an improvement on TCAS II, since it has all the features of TCAS II but can provide RAs in the horizontal as well as vertical plane. This system is not yet fully developed.
4. ACAS carriage requirements are detailed in Article 39(2) and Schedule 5 of the Air Navigation Order 2009. Military Requirements are contained within:

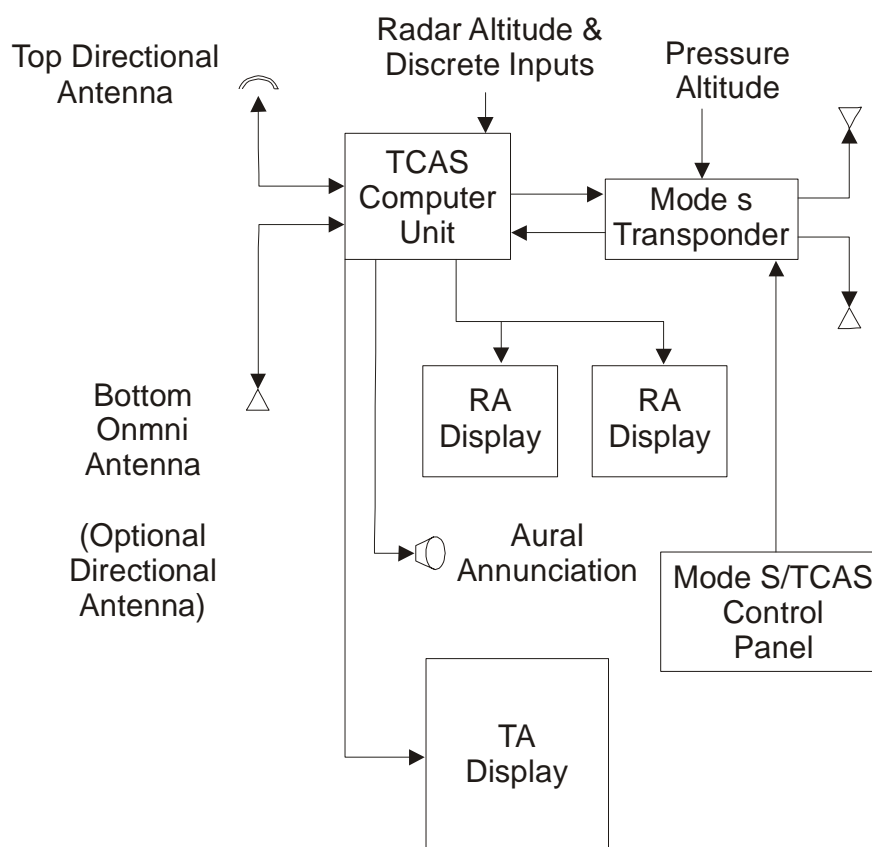
- a. MAA Regulatory Publications: Air Traffic Management (ATM) 3000 Series Regulatory Article RA 3013 Airborne Collision Avoidance Systems.
- b. MAA Manual of Military Air Traffic Management Chapter 13: Airborne Collision Avoidance System: Traffic Alert and Collision Avoidance System Regulatory Cross-Reference.

Some aircraft in UK military service employ a Traffic Advisory System (TAS). TAS is a TCAS I derivative that interrogates aircraft out to 20 nautical miles (nm), but only displays aircraft out to 7nm. Because of this similar capability, yet reduced functionality of a TCAS I, it is designated as a TAS.

Components

5. The equipment required to provide a TCAS II system consists of (see Fig 1):
 - a. A TCAS Computer with Transmit/Receive Unit and two aeralis. The upper aeral is directional; the lower aeral may be either directional or omni-directional.
 - b. A Mode S transponder with top and bottom omni-directional aerial.
 - c. A combined Mode S/TCAS control panel.
 - d. A cockpit display and aural warning generator.

7-16 Fig 1 TCAS II Component Diagram



Alerts

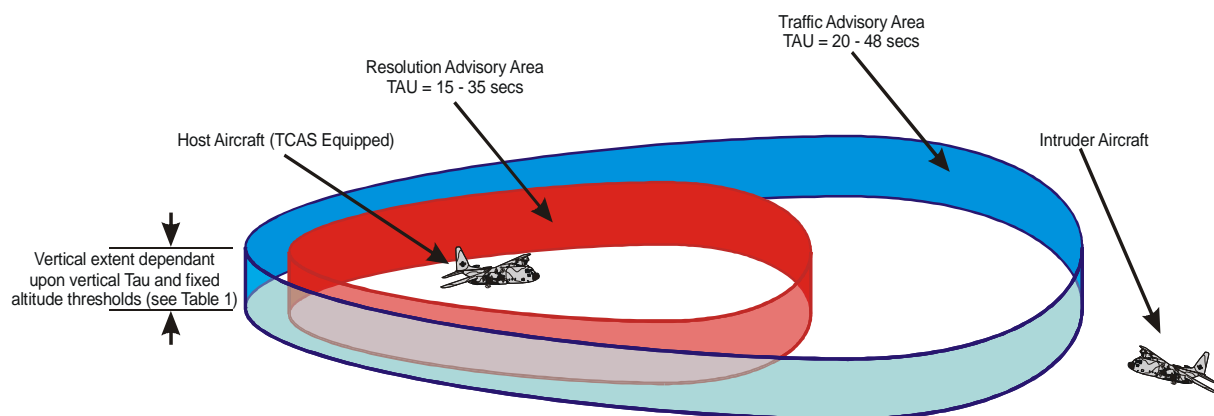
6. Airborne Collision Avoidance Systems can issue two types of alert, a Traffic Advisory (TA) and a Resolution Advisory (RA).

- a. Traffic Advisory. TA alerts warn the pilot of possible conflicting aircraft but do not recommend avoiding maneuvers. They help the pilot in the visual search for the intruder aircraft and prepare for a possible RA.
- b. Resolution Advisory. RA alerts recommend vertical manoeuvres that will either increase or maintain the existing vertical separation between aircraft. Where both aircraft are fitted with TCAS II, the ACASs co-ordinate their RAs through the Mode S data link in order to select complementary resolution senses (one may recommend a climb, the other may recommend a descent or to maintain level in order to provide separation between aircraft).

Protected Airspace

7. The idea of protected airspace is based on studies carried out by Dr John Morrell into the fundamental physics of the collision avoidance problem. Morrell published his findings in 1956 and introduced the concept of Tau (τ - the Greek letter 't'), which he defined as the range of a conflicting aircraft divided by its closure rate, thus the defining parameter for TCAS is time. All TCAS technology is based on this concept. TCAS II Protected Airspace is illustrated in Fig 2. The use of the term 'protected airspace' around a TCAS equipped aircraft does not prevent an intruder aircraft from penetrating that airspace. The protected airspace merely defines the volume of air in which a TA or RA is triggered.

7-16 Fig 2 TCAS II Protected Airspace



Sensitivity Levels

8. **Sensitivity levels (SL).** The values of Tau which are used to trigger either TA or RA warnings depend on the sensitivity level (SL) of the equipment and are between 20 and 48 seconds and 15 to 35 seconds respectively. Increasing the SL increases the trigger values of Tau. The highest values (48 secs TA and 35 secs RA) apply to level 7, which is used above 20,000 ft; the lowest values (20 secs TA and 15 secs RA) are used at lower levels (see Table 1). The longer warning times are necessary at high levels where manoeuvres take longer to take effect and speeds are generally higher. The SL mode is selected by the pilot via the TCAS Control Panel.

- a. **Standby.** When selected to Standby, TCAS operates at SL 1 and TCAS does not transmit any interrogations. Standby is normally only selected on the ground or if TCAS has failed and is the only way that SL 1 is selected.
- b. **TA-ONLY.** When selected to TA-ONLY, TCAS is placed into SL 2. At this setting, TCAS performs all surveillance functions and will issue TAs as required but will inhibit RAs.
- c. **TA-RA.** When selected to TA-RA TCAS automatically selects the appropriate SL based on the altitude of the host aircraft. The SL settings are given in Table 1.
- d. **Ground Based Control.** TCAS II has the capability for ground-based control of the SL built into its design, allowing the SL to be reduced from the ground using a Mode S uplink message. This facility is already being used at major international airfields with complex taxi patterns as an aid to the ATC ground controller.

Fig 3 shows the relationship between Range and Closure Speed for sensitivity level 5 for both TA and RA. Note that the boundary lines are modified at close range to provide added protection against slow closure encounters (see Note 2, Table 1).

Table 1 Sensitivity Level Definition and Alarm Thresholds

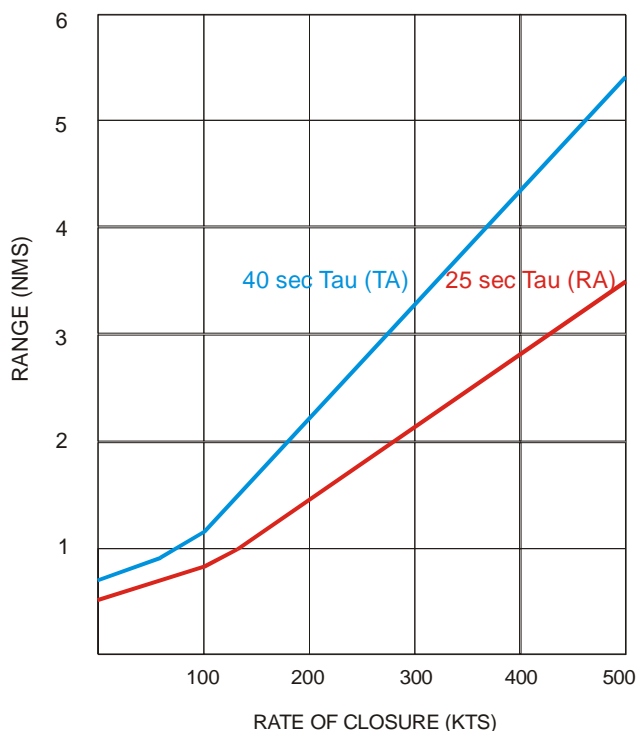
| Own Altitude (feet) | Sensitivity Level | (1) Tau (seconds) | | (2) DMOD (nm) | | (3) ZTHR (ft) Alt Threshold | | (4) ALIM (ft) |
|------------------------|----------------------|-------------------------|-----|---------------------|------|-----------------------------------|-----|---------------------|
| | | TA | RA | TA | RA | TA | RA | RA |
| < 1000 (AGL) | 2 | 20 | N/A | 0.30 | N/A | 850 | N/A | N/A |
| 1000 – 2350 (AGL) | 3 | 25 | 15 | 0.33 | 0.20 | 850 | 600 | 300 |
| 2350 – 5000 | 4 | 30 | 20 | 0.48 | 0.35 | 850 | 600 | 300 |
| 5000 – 10000 | 5 | 40 | 25 | 0.75 | 0.55 | 850 | 600 | 350 |
| 10000 – 20000 | 6 | 45 | 30 | 1.00 | 0.80 | 850 | 600 | 400 |
| 20000 – 42000 | 7 | 48 | 35 | 1.30 | 1.10 | 850 | 700 | 600 |
| > 42000 | 7 | 48 | 35 | 1.30 | 1.10 | 1200 | 800 | 700 |

Notes:

- Tau** – Time to go to CPA. Time to CPA is the range Tau, and time to co-altitude is the vertical Tau.

$$\text{Range Tau} = \text{slant range (nm)} / \text{closing speed (kt)} \times 3600$$

$$\text{Vertical Tau} = \text{altitude separation (ft)} / \text{vertical closing speed (ft/min)} \times 60$$
- DMOD** – With very low closure rates, intruder aircraft can get very close in range without crossing the Tau boundaries and thus triggering a TA or RA. In these circumstances, a modified definition of range Tau is used where the system uses a set range to trigger alerts. This range is called Distance MODification (DMOD).
- ZTHR** – Where the vertical closure rate between the TCAS and intruder aircraft is low, or when they are close but diverging in altitude, TCAS uses a fixed altitude threshold, in conjunction with the vertical Tau, to determine whether a TA or RA should be issued. This value is referred to as ZTHR.
- ALIM** – The latest software issue for TCAS introduced a horizontal miss distance (HMD) filter to reduce the number of RAs against intruder aircraft having a large horizontal separation at CPA. As part of the range test, the HMD filter can also terminate an RA prior to the Altitude LIMit (ALIM) being obtained to minimize altitude displacement when the filter predicts that the horizontal separation at CPA will be large.

7-16 Fig 3 TA/RA Tau Values for Sensitivity Level 5

9. **Non-threat Traffic.** Apart from traffic which generates RAs and TAs, as described earlier, TCAS will also show all other transponding traffic within range, giving the pilot a good situational awareness. All intruder aircraft more than 1,200 feet above or below the host aircraft, or more than 6 nm away are considered to be non-threat traffic. When the relative height is less than 1,200 ft or the range is less than 6 nm, the traffic is considered as 'Proximity' traffic, although it may still not be treated as a threat.

Operation

10. TCAS uses the same transponder principle, and the same frequencies, as the ground-based IFF/SSR system, which has been in use by Air Traffic Control for many years. For successful operation, TCAS depends on all aircraft carrying a serviceable IFF/SSR transponder; any aircraft not so equipped will be invisible to the system.

11. TCAS equipped aircraft will elicit information from all aircraft within range by sending out an interrogation signal on 1030 MHz. Suitably equipped aircraft will respond to this interrogation on 1090 MHz. The directional antenna enables TCAS to calculate the responding aircrafts' bearings while the times to reply permit range calculations to be carried out. Aircraft with IFF/SSR Mode C also transmit their altitudes. In addition, TCAS equipped aircraft transmit an omni-directional signal from Mode S transponders, known as 'Squitter', once per second, on 1090 MHz. This will alert any other TCAS equipped aircraft of their location.

12. TCAS aircraft use the data they have collected to compile a list of aircraft in their vicinity. The system can hold information on up to 45 contacts, of which it can display 30 and calculate RAs on 3 simultaneously. Once the TCAS equipped aircraft has compiled its list, it updates it once per second, thereby ensuring that it is operating on current data.

13. Using this compiled list, TCAS determines which, if any, of the contacts represents a potential collision threat. It does this by predicting the future positions of all contacts by applying the appropriate rates of change of range and altitude to their current positions. Once the threats have been determined, TCAS will provide the appropriate warnings to ensure that vertical separation is at least 300 ft at low altitudes or 700ft at high altitudes (see Table 1, ALIM). When the potential conflict is another TCAS equipped aircraft, the two TCAS computers will 'agree' a course of action and co-ordinate their respective RAs.

TCAS Aural / Visual Warnings

14. **Aural Warnings.** If the criteria for issuing a TA or RA are met, the crew will be alerted with an aural warning. The warning will depend on the version of the equipment software and operators should be familiar with the current state of their equipment. Warnings in use at the time of publication are given in Table 2.

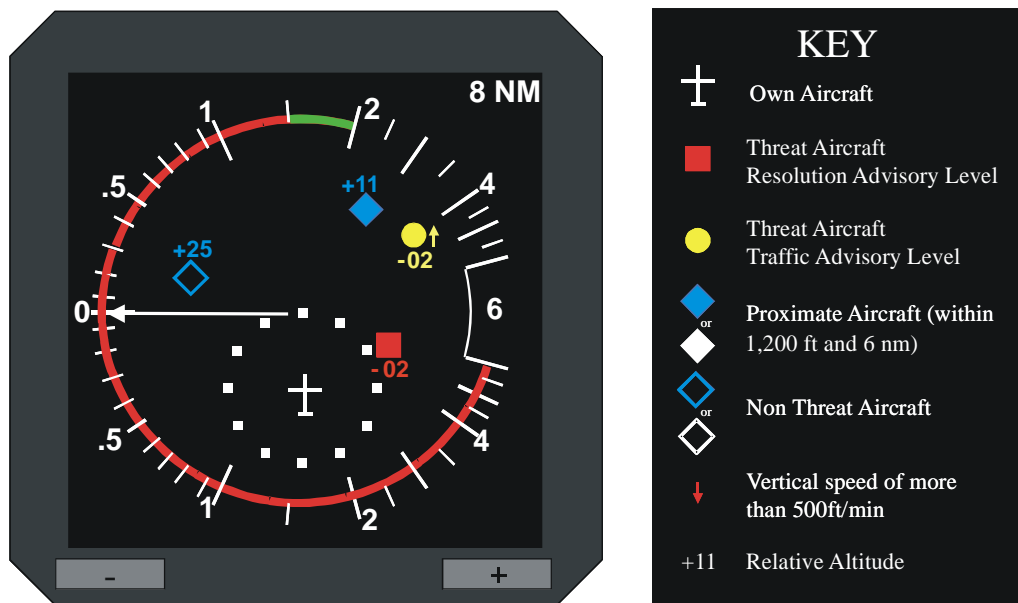
Table 2 TCAS Aural Annunciations

| TCAS Advisory | Version 7.1 | Version 7.0 | Version 6.04a |
|--|---|--------------------------------------|---|
| Traffic Advisory (TA) | Traffic, Traffic | | |
| Climb RA | <i>Climb, Climb</i> | | <i>Climb, Climb, Climb</i> |
| Descend RA | <i>Descend, Descend</i> | | <i>Descend, Descend, Descend</i> |
| Altitude Crossing Climb RA | <i>Climb, Crossing Climb; Climb, Crossing Climb</i> | | |
| Altitude Crossing Descend RA | <i>Descend, Crossing Descend; Descend, Crossing Descend</i> | | |
| Reduce Climb RA | <i>Level Off, Level Off</i> | <i>Adjust Vertical Speed, Adjust</i> | <i>Reduce Climb, Reduce Climb</i> |
| Reduce Descent RA | <i>Level Off, Level Off</i> | <i>Adjust Vertical Speed, Adjust</i> | <i>Reduce Descent, Reduce Descent</i> |
| RA Reversal to Climb RA | <i>Climb, Climb NOW; Climb, Climb NOW</i> | | |
| RA Reversal to Descend RA | <i>Descend, Descend NOW; Descend, Descend NOW</i> | | |
| Increase Climb RA | <i>Increase Climb, Increase Climb</i> | | |
| Increase Descent RA | <i>Increase Descent, Increase Descent</i> | | |
| Maintain Rate RA | <i>Maintain Vertical Speed, Maintain</i> | | <i>Monitor Vertical Speed</i> |
| Altitude Crossing Maintain Rate RA (Climb and Descend) | <i>Maintain Vertical Speed, Crossing Maintain</i> | | <i>Monitor Vertical Speed</i> |
| Weakening of RA | <i>Level Off, Level Off</i> | <i>Adjust Vertical Speed, Adjust</i> | <i>Monitor Vertical Speed</i> |
| Preventive RA (no change in vertical speed required) | <i>Monitor Vertical Speed</i> | | <i>Monitor Vertical Speed, Monitor Vertical Speed</i> |
| RA Removed | <i>Clear of Conflict</i> | | |

15. **Visual Warnings.** In 'non-glass' cockpits, TCAS information is normally presented on a combined TCAS/VSI display. The precise implementation in glass cockpit aircraft (Electronic Flight Information Systems (EFIS)) will vary, but RA information is normally shown as a commanded pitch manoeuvre via the flight director system on the Primary Flight Display, or Head-up Display, whereas TA information will appear separately on the Navigation Display.

16. **The TCAS/VSI display.** Fig 4 shows a combined TCAS and VSI display. The pointer and the scale round the perimeter represent an electronic version of a VSI. RA information is shown by the red and green arcs; the aircraft should ideally fly at a rate of climb or descent shown by the green arc and should avoid the red arc. In order to de-clutter the number of traffic symbols, displays are usually configurable in terms of range and ABOVE/BELOW. When in normal mode the display will show transponding aircraft 2700' above and below the host aircraft to a range selected. When ABOVE is selected the upper height, range displayed is extended and when BELOW is selected the lower height range displayed is extended. Depending on software this allows the display to show traffic up to 9900' vertically displaced from the host aircraft for climb/descent phases of flight. The ABOVE/BELOW and range selections have no control over which IFF equipped aircraft are interrogated by the host aircraft. If an intruder which generates a TA is outside the parameters selected for the display it will automatically generate a TA warning symbol irrespective of the display settings.

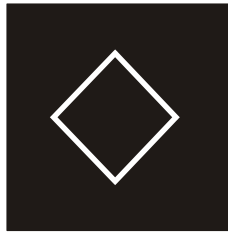
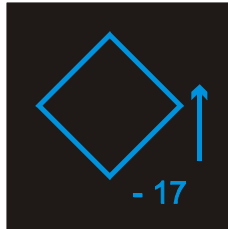
7-16 Fig 4 TCAS/VSI Display



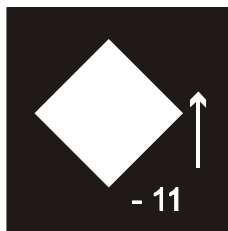
Display Symbology

17. **Display Symbology.** In Fig 4, the aircraft symbol represents the user's position and is surrounded by a circle of dots at a range of 2 nm. The display scale, 8 nm in the diagram, usually has ranges of 4, 8 and 16 nm which are user-selectable. The coloured shapes represent the various threat and non-threat aircraft. The symbols may have an altitude tag associated with them. This tag shows the relative altitude in hundreds of feet; a + sign indicates that the intruder is above the host aircraft, while a – sign shows that it is below. If the intruder is climbing or descending at a rate of 500 ft per minute or greater, a trend arrow, in the appropriate sense, will be shown next to the symbol. If the intruder is not reporting altitude, then no numbers or trend arrows will appear. A decode of the various symbols is given in the KEY to Fig 4; a more detailed explanation, with examples, is given in the following sub-paragraphs.

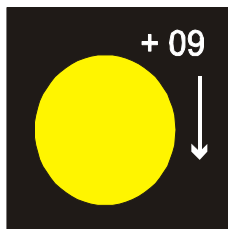
- a. **Non-threat Traffic.** The basic symbol for non-threat traffic is an open white diamond, as shown in Fig 5, with the alternative being an open blue (Cyan) diamond (Fig 6). In the example at Fig 5, there is no altitude information associated with the symbol, indicating that it is from an aircraft which is not reporting altitude, but since it is not a threat then it must be more than 6 nm away. In the example at Fig 6, the contact is 1,700 ft below the host aircraft and is climbing at a rate of at least 500 ft per minute.

7-16 Fig 5 Non-threat traffic, not reporting altitude**7-16 Fig 6 Alternative display for non-threat contact – Cyan instead of White**

b. **Proximate (or Proximity) Traffic.** A filled white diamond (Fig 7) represents traffic which is within 1,200 ft vertically or 6 nm laterally. This is known as 'Proximate Traffic' but is not considered to be a threat. As with non-threat traffic, the symbol may also be Cyan in colour.

7-16 Fig 7 Proximity Traffic 1,100 ft below, climbing at 500 ft/min or more

c. **Traffic Advisory.** Traffic considered to be a potential threat is displayed as a yellow circle (Fig 8). In the example shown, the traffic is 900 feet above the host aircraft and is descending at a rate of at least 500 ft per minute.

7-16 Fig 8 Traffic Advisory traffic 900 ft above, descending at 500 ft/min or more

d. **Resolution Advisory.** A solid red square (Fig 9) indicates that the intruder aircraft is considered to be a collision hazard. This symbol appears together with an aural warning and a vertical manoeuvre indication on the VSI. An RA can also be upgraded or altered once issued – for example an aircraft that has responded to a 'CLIMB CLIMB' RA may then be issued an enhanced RA ('ADJUST VERTICAL SPEED, ADJUST') or an RA reversal ('DESCEND, DESCEND NOW'). In Fig 8, the contact is now 500 feet below the host aircraft at a steady altitude.

7-16 Fig 9 Resolution Advisory traffic 500 ft below**Reaction to TACS II Warnings**

18. **Reaction to TCAS II Warnings.** The appropriate reactions to TCAS warnings are summarized below:

- a. TCAS II equipment reacts to the transponders of other aircraft.
- b. Warnings are based on the time to the predicted Closest Point of Approach (CPA) of vertical and horizontal distances preset in the software.
- c. When given a Traffic Advisory (TA) warning, pilots are advised not to take avoiding action but to look for the conflicting traffic.
- d. When given a Resolution Advisory (RA) warning, pilots are expected to react immediately (within 3 secs) and advise the ATC unit as soon as is practicable.
- e. When given an enhanced RA or an RA reversal, pilots are expected to react within 1 sec. This reaction time is possible due to the heightened awareness of the pilot, having already followed an initial RA warning.

Any intruder traffic with basic IFF (no mode C or S) will generate a non-threat or proximate symbol as the TCAS cannot judge a relative altitude. A TA warning will also be generated if the horizontal separation is degraded because the TCAS assumes that there is no vertical separation. In this case an RA warning is not possible as the TCAS has no altitude information to base its vertical avoidance on. Some detail is given in the CAA Radiotelephony Manual (CAP 413) which includes the R/T phraseology to be used when reacting to a TCAS RA. TCAS RAs should be reported using the Air Traffic Control Occurrence Reporting (ATCOR) scheme.

Operating Restrictions

19. TCAS may need to be restricted to some degree to avoid safety conflicts in certain flight scenarios. Typical restrictions will include:

- a. If a stall or ground proximity warning takes priority.
- b. RAs are inhibited below 500 ft AGL.
- c. Descend RAs are inhibited below 1,000 feet AGL.
- d. RAs which call for an increased rate of descent are suppressed below 1,800 ft AGL.
- e. There may be situations where RAs may be inhibited by the user by selecting TA-ONLY (see para 8b)

(1) Aircrew manuals or regulations may direct operators to inhibit RAs in situations where the aircraft does not have normal levels of performance. For example, when a multi-engined

aircraft has an engine shut down, the crew may not be able to comply with an RA warning and so operating in TA only mode may be more appropriate.

(2) To prevent RAs between aircraft flying in formation.

(3) In the visual circuit. For example, the TCAS would issue a RA between an aircraft turning finals against initials join traffic. Other procedures and airmanship negate the collision risk. By operating TCAS in TA-RA in the circuit, there is also a danger that pilots would get used to ignoring RAs.

Limitations

20. TCAS is designed to complement, not replace, air traffic control systems and good airmanship. It has limitations in that:

- a. The system only works with transponding aircraft. Thus, a TCAS fitted aircraft will not detect a non-transponding aircraft.
- b. TAs are intended to enhance situational awareness and assist in visual acquisition of conflicting traffic, however, visually acquired traffic may not be the same traffic causing the TA.
- c. TAs can be issued against any transponder-equipped aircraft that respond to Mode C interrogations, even if the aircraft does not have altitude-reporting capability.
- d. Avoiding action will only be given against transponding aircraft if that aircraft is operating Mode C IFF. TCAS cannot take RA action on an aircraft which is only using Mode A since it has no height information on the other aircraft.
- e. The system relies on the pilot taking immediate avoiding action based on the TCAS alert. Unless there is a clear safety reason not to act on an RA, then the pilot should react. Choosing to ignore the advisory under the belief that the threat has been identified may be worse than having no TCAS at all, since the pilot is not necessarily aware of the coordination that has taken place between the TCAS units. There is also the possibility that TCAS has reacted to a pop-up contact that the pilot was not aware of but needs to avoid.
- f. TCAS does not provide the precise location of an intruder aircraft, but rather its FL, range and closure. The information allows the pilot to look for the intruder under TA conditions or react to the intruder under RA conditions.
- g. The depicted position of an aircraft may be in error due to limitations in the azimuth performance of the TCAS directional antenna and computer. The displayed position may lag the actual location of the aircraft due to processing delays or the host aircraft manoeuvring.

Legislation

21. ACAS carriage requirements are detailed in Article 39(2) and Schedule 5 of the Air Navigation Order (CAP 393), issued by the UK Civil Aviation Authority (CAA). Similar regulations apply in many other countries around the world and military aircraft will become increasingly restricted in their operations if they cannot comply with these directives. Guidance for operators of UK military aircraft fitted with TCAS is given in the Manual of Military Air Traffic Management (Chapter 13).

The Future

22. Although working TCAS III systems have been trialled, problems still exist with achieving sufficiently accurate azimuth information to enable horizontal RAs to be given. TCAS IV is a separate development of TCAS II, which aims to achieve horizontal RAs using the advanced datalink capabilities of Mode S which will enable TCAS equipped aircraft to exchange more accurate navigational information.

23. There are other areas of research currently being pursued, which could lead to radical changes within the Air Traffic Control regime. One such concept is the Automatic Dependant Surveillance – Broadcast (ADS-B) system in which an aircraft's position, altitude, vector and other information are broadcast via datalink rather than being supplied on demand by the conventional transponder system.

FLARM

Introduction

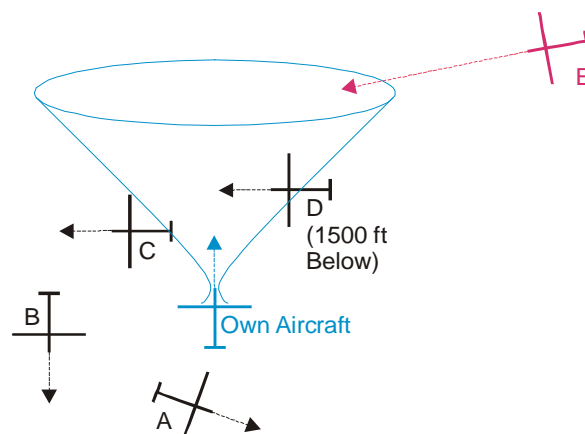
24. FLARM (Flight Alarm) is a collision avoidance system designed for use in gliders. The equipment gives warning of the risk of collision with other FLARM equipped aircraft and also the location of non-threatening FLARM equipped aircraft.

Principle of Operation

25. FLARM uses a barometric sensor and in-built GPS receiver to determine the aircraft position in space. The equipment uses this data every second to determine the predicted flight path of the aircraft and transmits the information via radio to other FLARM equipped aircraft within range. The operating range is dependent upon the antenna installation and is typically 2 to 5 km. Signals received from other aircraft are processed and compared to the predicted flight path of the receiving aircraft. If the unit determines that there is a risk of collision, audio and visual warnings of the greatest danger is given. Some FLARM systems can be loaded with a database to provide collision warnings against ground obstacles such as radio masts.

26. In each one second cycle, the equipment determines the aircrafts' absolute position in space, in terms of latitude, longitude and altitude, and also the track of the aircraft. From this information, the projected flight path of the aircraft over the next 18 seconds is determined. This digital data is transmitted over a common FLARM radio frequency and data from other FLARM equipped aircraft within range is received. Own aircraft data is compared to other aircraft data to determine whether the projected flight paths coincide and hence present a risk of collision. Fig 10, illustrates a possible scenario, in plan view, where several gliders are in close proximity utilizing the same thermal to gain height.

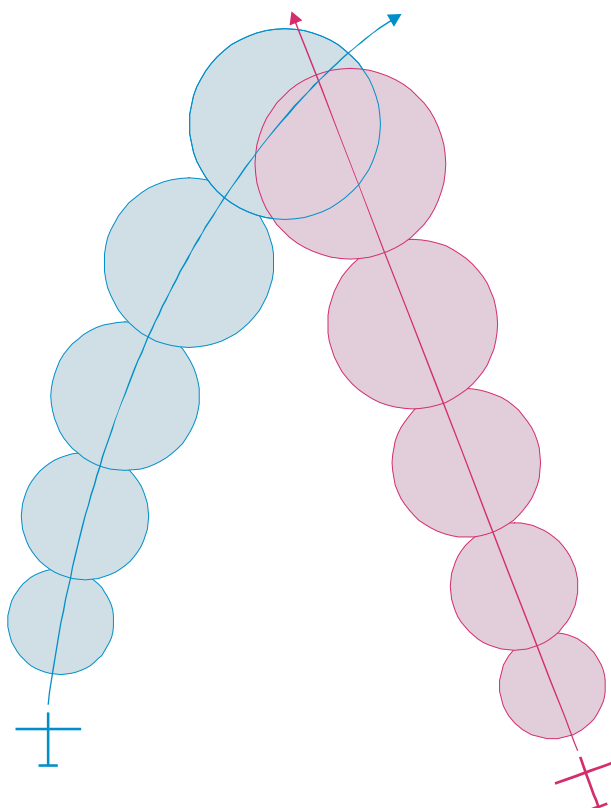
7-16 Fig 10 Illustration of FLARM Confliction Scenario



27. In Fig 10, the blue cone represents the volume of space that an aircraft could occupy over the following 18 seconds, based on the predicted flight path of the on board FLARM. Aircraft A, B, C and D are in close proximity using the same thermal. Aircraft E is joining the thermal. Aircraft A is very close but moving away and so does not pose a collision risk. Aircraft B and C are also very close and although at a similar altitude do not pose a threat within the next 18 seconds. Aircraft D does not pose a threat due to the altitude difference. Aircraft E, while more distant than the other aircraft is predicted to be within the volume of space that one's own aircraft may occupy within the next 18 seconds, and thus poses a risk of collision. The potential for a dangerous risk developing is increased by the fact that the aircraft within the thermal are turning to the left, and thus aircraft E may not be visible from one's own aircraft as it will be hidden below the fuselage. It must be remembered that the FLARM data is updated each second, and as such the scenario represented in Fig 10 will be continually changing. Thus, aircraft E may quickly cease to be a collision risk while the other aircraft may become so.

28. FLARM determines future time windows of up to 18 seconds. For each of these time windows the equipment calculates the volume of space that the aircraft could occupy. At the same time, it checks, using the received data from other aircraft, whether the flight paths of those aircraft could intersect that volume of space. Due to the way the data is processed, FLARM can not only predict potential collision risks, but also the time that any collision might take place. Thus, it can predict how imminent the threat is and indicate this through its warning display. Fig 11 illustrates the development of a potential collision. It can be seen that the calculated volume of space close to the aircraft is smaller than that furthest from the aircraft, due to the time differences. Thus, with multiple aircraft in close proximity, FLARM can determine which aircraft poses the greater threat and warn the pilot accordingly.

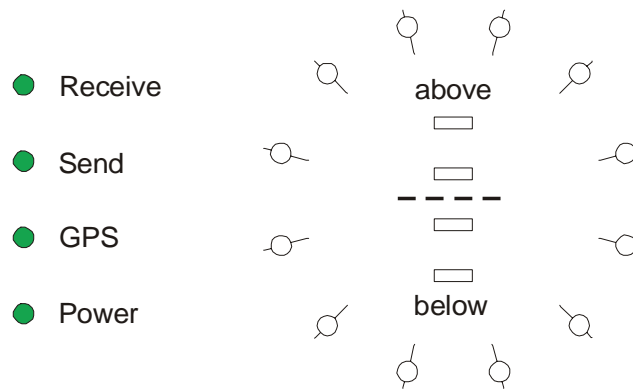
7-16 Fig 11 Illustration of Predicted Flight Paths



Warnings, Displays and Modes of Operation

29. Different equipments display the alert information in different ways, but all systems should give both aural and visual warnings. Fig 12 shows a typical FLARM display using bi-colour LEDs, but conflicting aircraft can also be shown on moving map displays within the cockpit.

7-16 Fig 12 Typical LED FLARM Display



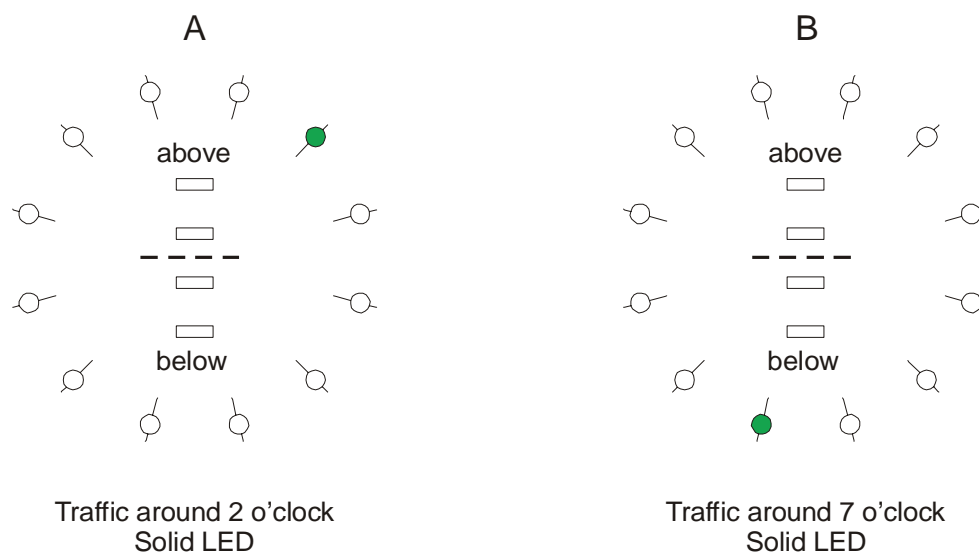
Modes of Operation. FLARM operates in two modes, Nearest and Collision Modes.

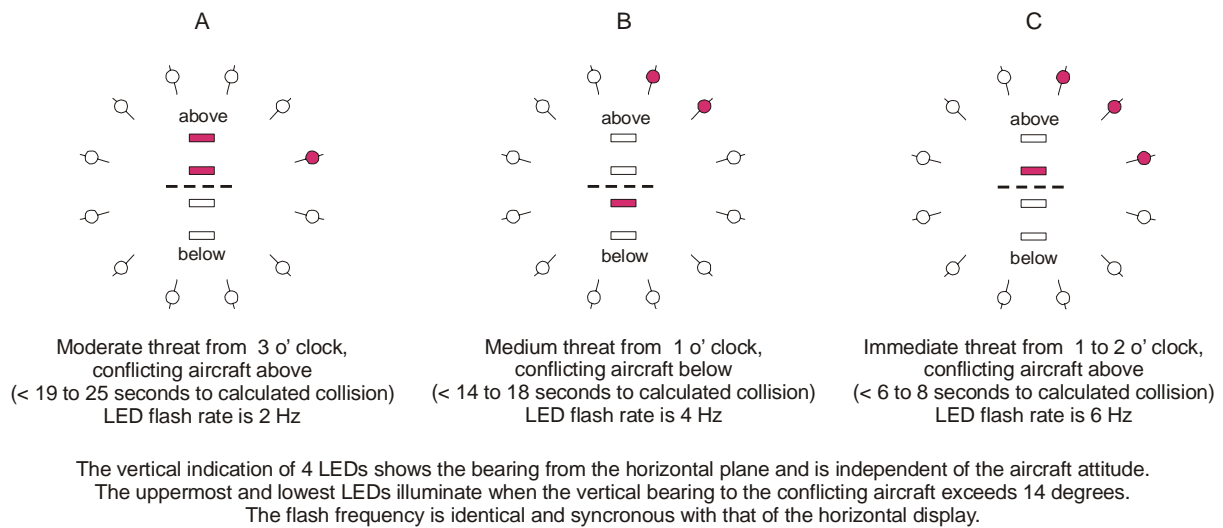
a. **Nearest Mode.** When operating in Nearest Mode, FLARM indicates the presence of other aircraft within range even though they do not pose a threat of collision. The information is limited to a configurable radius and a vertical separation of 500 m. Only one aircraft is displayed, by means of a green LED (see Fig 13 a and b) and the threat intensity is not indicated. When the equipment detects the risk of collision, it switches automatically to Collision Mode.

b. **Collision Mode.** When operating in Collision Mode, FLARM indicates the presence of the aircraft giving the most immediate risk of collision using aural and visual warnings (see Fig XX a, b and c). In this example, the number of illuminated LEDs and their flash rate indicates the threat level. When the risk of collision is over, the equipment will revert to Nearest Mode. Collision Mode can be selected as the default setting.

30. FLARM uses aircraft track to predict the flight path and so, in strong wind conditions, with high values of drift, distorted threat bearings can be indicated.

7-16 Fig 13 Examples of FLARM Alerts in Nearest Mode



7-16 Fig 14 Examples of FLARM Alerts in Collision Mode

PowerFlarm

31. A development of the basic FLARM is PowerFlarm. Whereas FLARM has the ability to detect other FLARM equipped aircraft, PowerFlarm also has the ability to detect, and give warnings with regard to Automatic Dependent Surveillance-Broadcast (ADS-B) equipped aircraft and those fitted with Mode C or S transponders. ADS-B is a surveillance tracking system for aircraft that transmits aircraft position and velocity data each second. ADS-B Out transmits aircraft identification, position, altitude and velocity, providing Air Traffic Controllers with real time information on aircraft in their area. ADS-B will be mandated on most aircraft in the USA by 2020 and on aircraft weighing over 5,700 kg or having a maximum cruising speed of over 250 kt, in Europe from 2017.

Limitations of FLARM

32. FLARM is designed for situational awareness only. It has several limitations and must not be expected to give totally reliable warnings.

- a. Basic FLARM will only indicate the presence of other FLARM equipped aircraft, however PowerFlarm has increased functionality (Para 31).
- b. FLARM operates at short range which can be compromised by equipment fit and poor antenna positioning.
- c. Basic FLARM does not 'communicate' with conventional transponders or Airborne Collision Avoidance Systems (ACAS).
- d. FLARM relies on good quality GPS reception to determine the current aircraft position accurately.
- e. With high values of drift, inaccurate threat bearings may be displayed.
- f. FLARM warns of only the most immediate threat and the warning time is very short. Other aircraft may pose a threat but not be displayed. As a result, FLARM does not replace good airmanship practices and lookout. The short range of the system makes it unsuitable for fast moving aircraft.
- g. FLARM only warns of the risk of collision, it does not offer avoiding action.
- h. FLARM radio communication frequencies are unprotected and so the possibility of interference is always present.

CHAPTER 17 - TELEVISION, LOW-LIGHT TELEVISION AND NIGHT VISION GOGGLES

TELEVISION

Introduction

1. Television has a number of applications in avionic systems, probably the most familiar being reconnaissance and missile guidance. Although airborne television is different from its commercial and domestic counterparts in terms of size and ruggedness, the principles of operation are identical.

Television Principles

2. A television camera receives light energy from a scene and converts it into electrical energy, using either or both of the following techniques:

a. **Photoemission.** The photoemission technique relies on the fact that light energies are sufficient to cause the ejection of electrons from the surface of materials such as sodium, caesium, and potassium. The number of electrons emitted is directly proportional to the incident light levels.

b. **Photoconduction.** Photoconduction is the process by which the conductivity of materials such as selenium, arsenic trisulphide, and lead monoxide is increased by exposure to higher light levels. Electrons are not ejected, but are moved to a higher energy level.

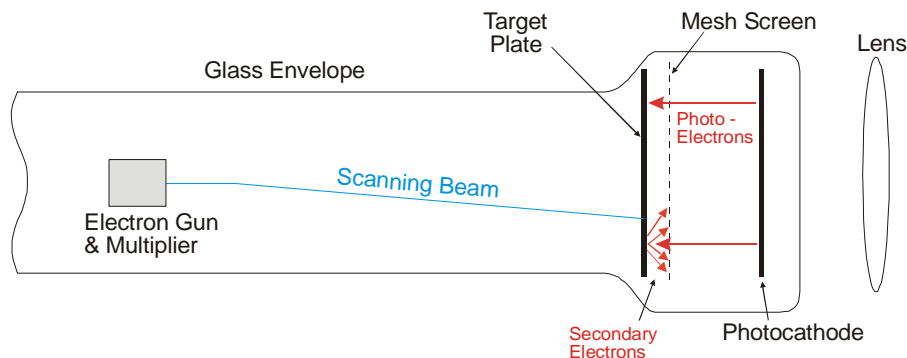
3. Cameras can be either scanning or non-scanning. The former scan the scene with a tube which converts incident light into voltages proportional to the changing light intensity. Staring arrays are used in non-scanning cameras; these employ a detector for each point of the picture (pixel). The varying voltage from scanning, or the output from the staring array, can be amplified and transmitted by wire or fibre optic cable, or it can be modulated onto a radio frequency carrier wave and radiated to another area.

4. The photoemission principle is used in the imaging orthicon type of camera (para 6); the photoconduction principle is used in the vidicon camera (para 9). Non-scanning, solid state, charge-coupled device cameras use both principles (para 11).

5. The display, or TV receiver, synthesizes the original scene by deflecting an electron beam spot across the fluorescent screen of a CRT and varying its brightness in accordance with the received signals. The scanning process is carried out sufficiently rapidly that an illusion of continuous, non-flickering, motion is achieved.

The Imaging Orthicon

6. The elements of an imaging orthicon are contained in a cylindrical glass tube envelope which has an enlarged section at one end (Fig 1). The device is closed at the enlarged end by an optically flat glass plate which has a continuous photosensitive coating (photocathode) deposited on the inside.

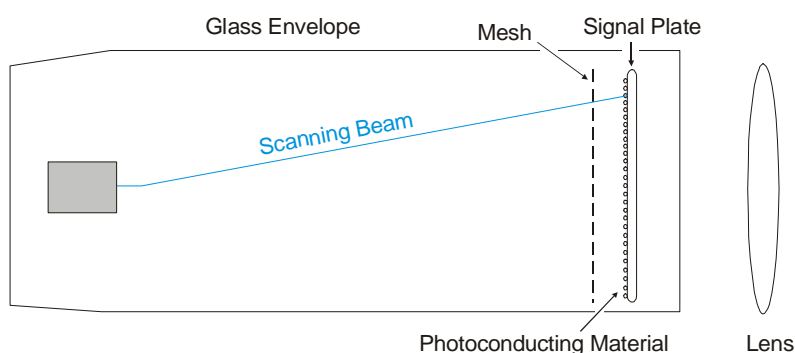
7-17 Fig 1 Simplified Construction of an Image Orthicon Tube

7. The optical image is focused on the lens side of the glass plate, and photoelectrons are liberated from the photocathode, in proportion to the light intensity at any point. These electrons stream through the enlarged cylindrical portion of the tube, and a fine mesh screen, until they encounter the target plate. At the target, they induce a secondary emission of electrons from the surface. These electrons are collected by the fine mesh screen that lies parallel and close to the target. The departure of the secondary electrons leaves positive charges on the target, with an intensity proportional to the light distribution of the original optical image. The changes in electrical potential, induced by the secondary emissions, are transferred to the opposite face of the target and, as the streams of electrons from the photocathode continue to fall on the target, the intensity of electrical image continually increases. As the target is very thin, there is negligible leakage of the charge parallel to the target surface, which would otherwise degrade the image detail.

8. The stored charge image on the reverse side of the target is scanned by an electron beam; the current in the returning beam varies in amplitude according to the variations in the intensity of the successive portions of the image being scanned. An electron multiplier produces an amplified reproduction of the current in the returning beam; the output is an analogue video signal which can be displayed or stored.

The Vidicon

9. The elements of the vidicon camera are contained in a cylindrical glass envelope, as shown in Fig 2. The signal plate, onto which the optical image is focused, has a very thin layer of photoconductive material deposited on it. This material has high electrical resistance in the dark, but becomes progressively less resistant as the amount of light increases. Thus, the optical image induces a pattern of varying conductivity, which matches the distribution of brightness in the image. The target is scanned by an electron beam and the resulting video signal may be displayed or stored.

7-17 Fig 2 Simplified Construction of a Vidicon Tube

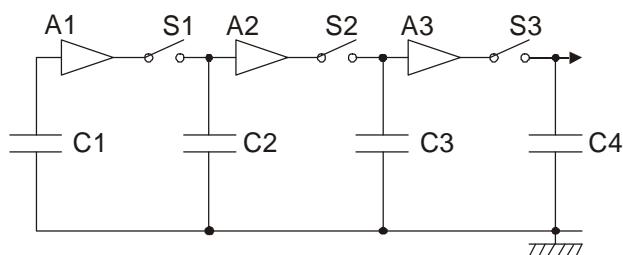
10. The scanning electron beam spot size imposes the only limitation on resolution. It is, therefore, possible to derive a high quality image from a photosensitive area no larger than about 1.3 cm^2 . This permits the use of small and comparatively inexpensive lenses, with a correspondingly large depth of field for any given aperture. In addition, the structure of the tube is simple and small, so cameras based on the vidicon principle are adaptable to a wide range of broadcasting, industrial, and military applications (e.g. missile guidance).

Charge Coupled Device (CCD) Systems

11. Television tubes are susceptible to shock, overload, and stray magnetic fields. They impose size and weight penalties in some applications, and they have a relatively short life expectancy. Solid state charge coupled devices, which are more rugged and reliable, have lower power requirements, and can be used to produce small, lightweight cameras with a long life.

12. A typical CCD sensor chip, about $8 \text{ mm} \times 10 \text{ mm}$ in size, contains a matrix of many thousands of silicon photodiodes. Each of these photodiodes effectively forms a capacitor which accumulates a charge proportional to the brightness of the light incident upon it. Fig 3 illustrates, schematically, a simplified arrangement in which C1, C2, C3, and C4 represent four such capacitors, which would form the first four pixels in one television line. The charge accumulated by each capacitor (eg C1) appears as a voltage at the output of the associated amplifier (eg A1). If all the switches, S1 to S3, were closed momentarily, the charge on C1 would be passed into C2, C2's charge would be transferred into C3, and so on all the way along the line. Thus, by momentarily closing all the switches synchronously, the charge representing the brightness level for each pixel can be made to proceed along the line of capacitors, at a rate dependent on the switching frequency, and can be detected at the end of the line and read serially as the pattern of the light level at each pixel. In practice, the stored charge in each capacitor is relayed to a second set of capacitors before being shifted, thus allowing the photodiodes to register the next pattern of illumination with no appreciable blank period.

7-17 Fig 3 Simplified Schematic CCD Arrangement



LOW-LIGHT TELEVISION (LLTV)

Introduction

13. There is a requirement for military aircraft to operate both by day and by night, and a television system could provide a solution to the visual acquisition problem of night operations. However, the systems described so far are, in general, only capable of producing a usable image down to the light levels associated with twilight. Variations of the vidicon and orthicon cameras can operate at lower light levels, but it is normally necessary to attach an image intensifier system to the camera for it to be usable down to the illumination level of starlight. Before describing the operation of these image intensifiers, the night environment will be reviewed.

The Night Environment

14. Fig 4 shows the range of natural illumination (lux is the SI unit of luminance; foot candles (fc) is the imperial unit). The human eye is at its maximum efficiency in daylight but can adjust to operate in lower light levels, albeit with a reduction in capability. Colour perception is lost once the illumination level is down to less than 1 lux; below this, the eye's resolving power is degraded rapidly such that at 10^{-4} lux the eye is only capable of discerning large, high-contrast objects, and then only after a prolonged period of dark adaptation. Clearly, unaided, the eye is not suitable for night operations; hence the need for some form of imaging system.

7-17 Fig 4 The Range of Natural Illumination

| Human Vision | Lux | fc | |
|--|-----------|-----------|--------------------|
| ↑ Photopic (good acuity, colour) | 10^5 | 10^4 | Sun on Snow |
| | 10^4 | 10^3 | Full Daylight |
| | 10^3 | 10^2 | Overcast Daylight |
| | 10^2 | 10 | Very Poor Daylight |
| | 10 | 1 | Twilight |
| | 1 | 10^{-1} | Deep Twilight |
| ----- | | | |
| ↓ Scotopic (poor acuity, no colour) | 10^{-1} | 10^{-2} | Full Moon |
| | 10^{-2} | 10^{-3} | Quarter Moon |
| | 10^{-3} | 10^{-4} | Starlight |
| | 10^{-4} | 10^{-5} | Overcast Starlight |
| ↓ | | | |

15. A television camera forms an image of the scene being viewed by utilizing a small proportion of the radiant energy which is incident upon the scene. The amount of energy which reaches the camera sensor depends mainly on the incident radiation level, the reflectance of the subject, and the light-collecting power of the optical system. Furthermore, the video signal current obtained from the sensor will be dependent upon the spectral energy distribution of the illuminant, and the spectral response and sensitivity of the sensor.

16. **Incident Radiation.** The level of incident radiation depends on many geographical and meteorological factors, such as the time of night, the latitude, the declination of the sun (season), the phase of the moon, and the degree of cloud cover. Worldwide surveys indicate that minimum illumination levels are practically constant everywhere at a value of about 1×10^{-4} lux; values less than this are rare, and in temperate regions the light level is above 10^{-3} lux for 82% of the time. Under clear, moonless conditions, the incident radiation has the following components:

- 30% direct or scattered starlight.
- 15% of zodiacal origin (caused by small particles reflecting sunlight).
- 5% of galactic origin.
- 40% from airglow (permanent luminescence of the night sky).
- 10% scattered light from these various sources.

The airglow phenomenon contributes nearly half of the incident radiation and originates from the ionization of rare gases in the upper atmosphere. Natural night sky spectral irradiance contains nearly 10 times as many incident photons per unit wavelength at $0.8 \mu\text{m}$ as at $0.4 \mu\text{m}$. In order to make use

of this high energy content, it is desirable that a sensor has the highest sensitivity in the red and near infra-red parts of the spectrum.

17. **Subject Reflectance.** The apparent brightness of an object depends not only on the level of incident radiation, but also on the manner in which the object reflects that illumination. If all objects reflected an equal amount of light, there would be no contrast between a target and its background, and it would probably not be seen. The reflectance of an object (the ratio of reflected to incident light) depends principally on the nature of the surface (eg its colour and texture), the wavelength and angle of the illumination, and the viewing angle. Table 1 shows some typical values of reflectance.

7-17 Table 1 Typical Reflectance Values

| | |
|----------------|--------------|
| Snow | 0.7 to 0.86 |
| Clouds | 0.5 to 0.75 |
| Limestone | 0.63 |
| Dry Sand | 0.24 |
| Wet Sand | 0.18 |
| Bare Ground | 0.03 to 0.2 |
| Water | 0.03 to 0.1 |
| Forest | 0.03 to 0.15 |
| Grass | 0.10 to 0.25 |
| Rock | 0.12 to 0.30 |
| Concrete | 0.15 to 0.35 |
| Blacktop Roads | 0.08 to 0.09 |

18. **Contrast.** Contrast is defined as:

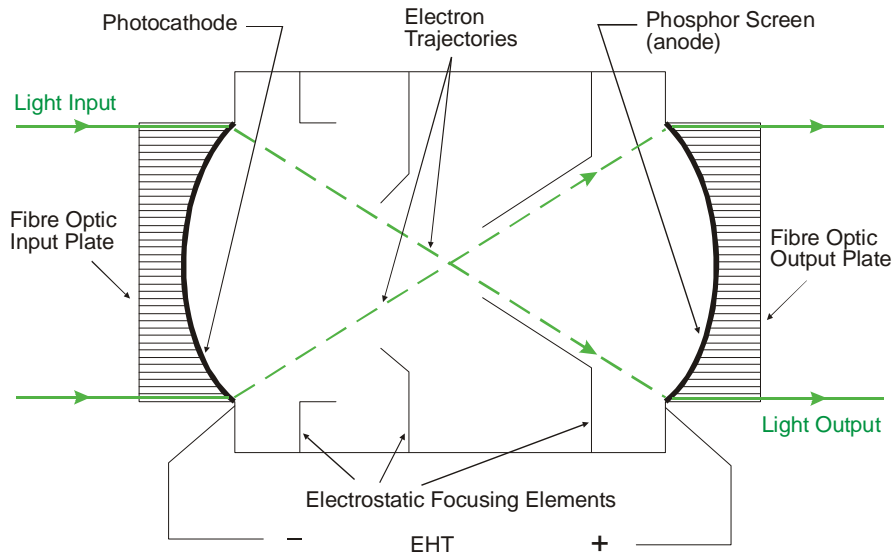
$$\frac{B_{\max} - B_{\min}}{B_{\max}}$$

where B_{\max} is the maximum luminance and B_{\min} is the minimum luminance in the scene; it is usually expressed as a percentage. At the camera, the contrast will often be reduced by the effect of the intervening medium, and in particular by haze, fog, and rain, resulting in a reduction in the effective range of the system.

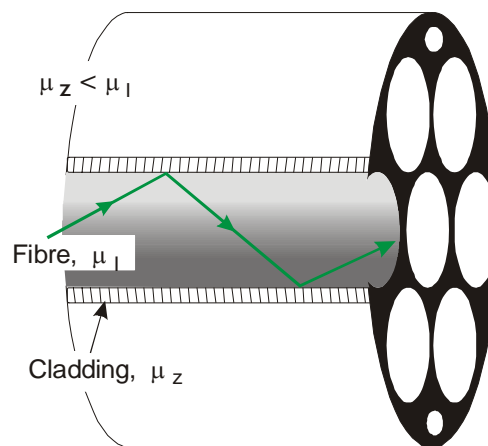
19. **Optical System.** The larger the diameter of the lens, the more capable it is of operating at low light levels. However, the maximum size of the lens will often be constrained by the physical limits of the airframe, by considerations of the increase in drag due to the flat plate effect, and by the need for the lens to resist the aerodynamic forces imposed by high-speed flight.

Image Intensification

20. A basic image intensifier is an electronic device which reproduces an image on a fluorescent screen. Fig 5 shows a schematic construction of such a device.

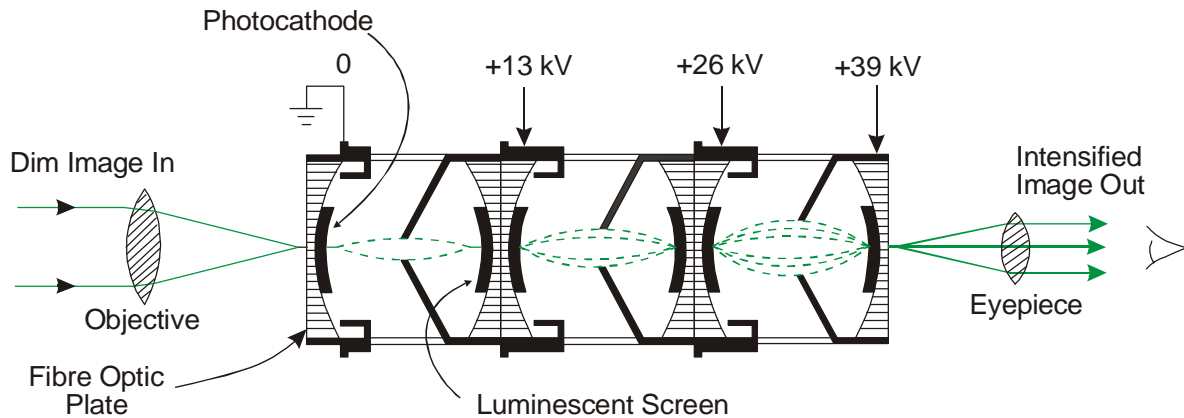
7-17 Fig 5 Construction of a Single Stage Image Intensifier

The input and output plates are fibre optic plates which are made up of a complex of minute glass tubes clad with another type of glass of lower refractive index, thus preventing cross-talk between adjacent fibres (Fig 6). These tubes form light guides, in which light entering at one end is trapped until it emerges at the other end. The shaped fibre optic plates also transform the flat optical image into a curved image, necessary for the electrostatic lens in the intensifier. Electrostatic focusing elements ensure that electrons released from a particular spot on the photocathode are focused onto a corresponding spot on the phosphor screen.

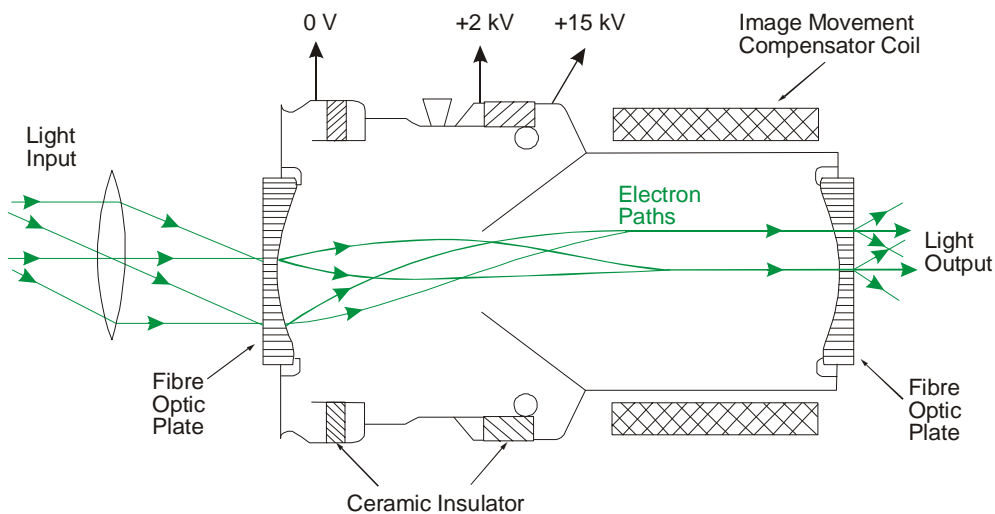
7-17 Fig 6 Construction of Fibre Optic Plat

21. The input plate is coated with a photocathode, and the output plate with a phosphor anode. When light falls on the photocathode, electrons are released which are accelerated towards the phosphor by the 15 kV field across the device. The increased energy acquired by the electrons is expended in exciting the phosphor, such that the image formed can be 40 to 50 times brighter than the original.

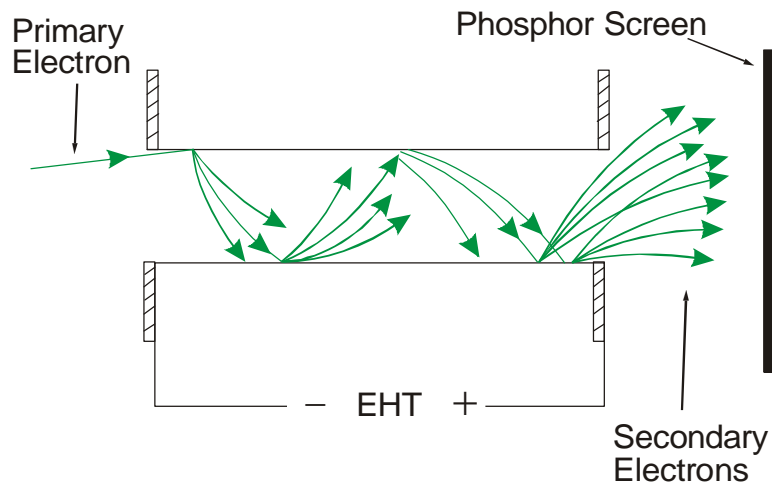
22. An extension of the basic intensifier is to arrange two or three tubes together in series to form a 'cascade' intensifier (Fig 7) which can achieve gains of up to 50,000 times at $0.4 \mu\text{m}$. A 3-stage cascade intensifier would be sufficient to make a simple vidicon into a LLTV tube, and similar devices have been used as simple, hand-held, direct view image intensifiers. The multi-stage systems have the disadvantages of lower picture quality, and increased size, weight, and cost.

7-17 Fig 7 Schematic Diagram of Modular-type Cascade Image Intensifier

23. If a LLTV system is used in a moving vehicle, there is likely to be some angular vibration which will degrade its resolution. To accommodate this problem, an image intensifier has been designed in which the output image can be magnetically deflected to counter the movement. The construction is shown in Fig 8. The front end is the same as a conventional image intensifier, but the output end has a relatively long, field-free, section where the electron beams are moving essentially parallel to each other. A transverse magnetic field in this area, controlled by position gyros, deflects the image as a whole to compensate precisely for the angular movement of the system.

7-17 Fig 8 Motion Compensated Intensifier

24. **Microchannel Plate (MCP) Intensifiers.** Microchannel plate intensifiers are capable of gains of up to 10^4 . A MCP is a thin plate of special glass with a matrix of fine holes (channels) through it. The holes are from $10\text{ }\mu\text{m}$ to $12\text{ }\mu\text{m}$ in diameter, and about 1 mm in length. The inside of each channel is coated with an electron rich material and, when a primary electron from the photocathode strikes the channel wall, secondary electrons are released. These, in turn, release more electrons as they move along the channel (Fig 9) and the channel thus acts as a miniature photomultiplier tube. A potential gradient is provided along the wall of the channel, accelerating the electrons before they strike the phosphor. MCP image intensifiers have the advantage of greater sensitivity, smaller size, and less weight than cascade devices.

7-17 Fig 9 Microchannel Plate - Principle of Operation

LLTV Limitations

25. The main limitations of LLTV are the limited field of view and blooming. The field of view (FOV) is typically $30^\circ \times 40^\circ$, which is barely adequate for the purpose of low-level visual navigation, for which a look-into-turn capability is needed. A narrower FOV may be useful for reconnaissance purposes, as this implies greater magnification and greater range. Blooming results from the amplification of a bright light such as a flare, beacon, or searchlight, in an otherwise low-light scene. This problem can be reduced by using filters which make light levels above a set threshold appear black.

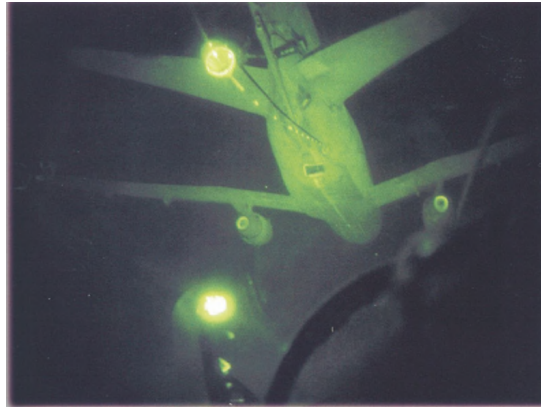
NIGHT VISION GOGGLES (NVGS)

General

26. Night vision goggles (NVGs) are one means of solving the FOV problem associated with a fixed LLTV system. The image intensifiers are mounted onto a helmet (see Fig 10) so that, although the instantaneous FOV is smaller than that of a LLTV, the crew member can look into the turn. NVGs are lightweight binocular devices and are sufficient for the rapid recognition of terrain obstacles at light levels down to overcast starlight.

7-17 Fig 10 Helmet-mounted Night Vision Goggles

27. Fig 11 shows a tanker aircraft viewed through NVGs from the cockpit of a fighter aircraft (note the clear example of blooming).

7-17 Fig 11 Tanker Aircraft viewed through NVGs

28. The prime disadvantages of NVGs are that their weight causes fatigue, they are a potential hazard in the event of ejection, and they amplify all light entering them, including that from internal cockpit displays. Operational aspects of using NVGs are discussed in more detail in Volume 12, Chapter 12.

29. Study of the spectral distribution of the night sky shows that there are ten times as many photons in the red and near infra-red region than there are in the blue region. Image intensifiers make use of this fact by being more sensitive to the red and near infra-red region of the spectrum. Unfortunately, cockpit displays contain light sources which are also rich in the red and near infra-red region, but at levels of radiation significantly higher than those emanating from the night scene. Consequently, NVGs can be blinded by cockpit illumination.

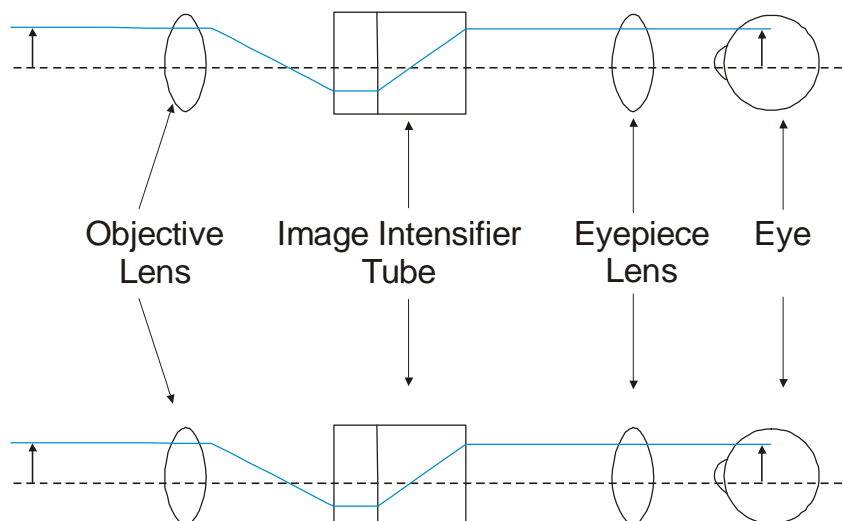
30. To overcome this problem, the spectral response of intensifiers used in NVGs must be modified by filtration to allow the use of the blue-green region of the visible spectrum for cockpit lights and displays. The cockpit lights are also filtered to reduce transmissions in the red and near infra-red regions of the spectrum. NVGs provide the user with improved night vision by the amplification of available natural light from the sky achieved using image intensification. Cockpit instruments lit by suitable NVG compatible lighting are viewed peripherally by the unaided eye.

Principles of Operation

31. The visible image produced by the NVG is derived solely from amplification of available light; NVGs cannot work in total darkness. However, very small amounts of light from cultural lighting, the moon and the stars provide sufficient illumination for flight on NVGs, even if the night is overcast. Dual optical channels (monocular) give stereoscopic vision which allows depth and speed perception.

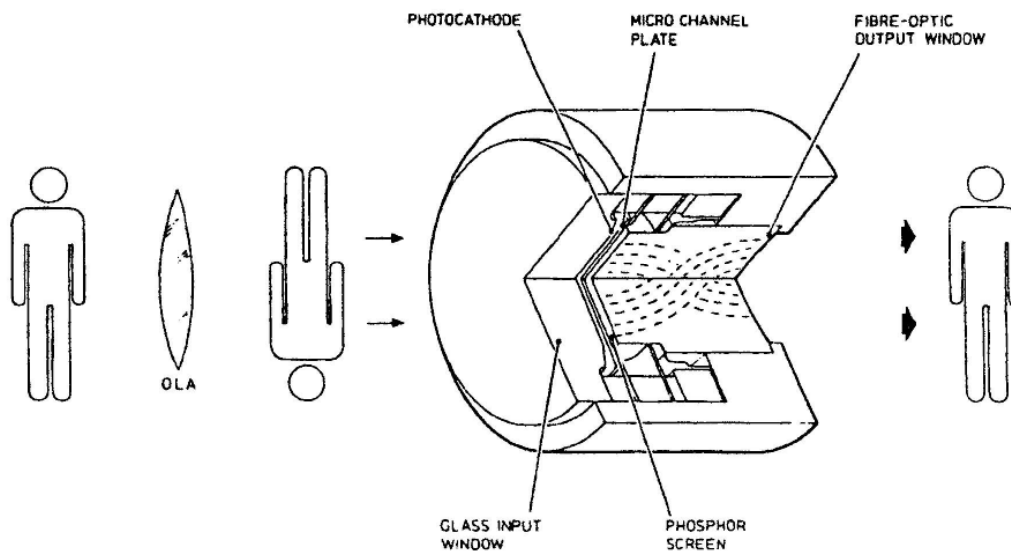
32. Reflected light from the viewed scene is collected by an objective lens in each monocular assembly and focused onto the image intensifier tube.

33. The intensifier tube consists of an input window, photocathode, microchannel plate, phosphor screen, fibre optic image inverter and annular power supply all in a sealed assembly about the size of a cotton reel. This is the core of any NVG system. (see Fig 12)

7-17 Fig 12 Optical System Schematic Diagram

34. The annular power supply uses switch-mode techniques to generate three EHT voltages, V1 (approx 800V), V2 (approx 800V) and V3 (approx 5.6 kV), from the 3.6V dc supplied by the battery.

35. Light entering each NVG monocular is focused onto the photocathode of the intensifier tube by the objective lens, where the emission of electrons from the photocathode converts the light image to an electron image. V1 then attracts electrons from the photocathode to the microchannel plate (see Fig 13).

7-17 Fig 13 Generation III Image Intensifier Tube (IIT)

36. The microchannel plate is the main gain producing component of the intensifier tube. It is the modern solid state equivalent of early vacuum tube devices (1st generation) and is a disc less than 0.4 mm in thickness, made from glass, which incorporates a series of minute holes (microchannels) (see Para 24). The walls of the microchannels are processed to enable electrons to be released. As electrons from the photocathode are accelerated down the microchannel by V2 they undergo collisions with the walls releasing additional electrons, avalanching at a rate of multiplication dependant upon V2 which is applied between the faces of the microchannel plate. The microchannel plate is in effect a disc of

miniature photo-multiplier tubes, where the output from each microchannel forms a pixel within the intensified image.

37. When these electrons emerge from the microchannels they are accelerated by V3 until they hit the phosphor screen where the electron image is converted back into light. The image from the phosphor screen is then passed through an inverting fibre optic converter to correct the orientation of the image. A fibre optic element is used in lieu of a lens to invert the image with minimum weight, assembly size and distortion.

38. The microchannel bore and individual fibre optic strand size, together with the size of the gaps between the photocathode, microchannel plate and phosphor screen (smallest is best), determine the limiting resolution of the tube.

39. Beyond a certain threshold the tube EHT power supply is current-regulated to maintain the same tube output brightness regardless of light input level; effectively a type of automatic gain control. However, at very low light levels with virtually no incident photons at the photocathode, electrons released by random ionisation events account for most light output from the phosphor screen. This is seen as scintillation (sparkling) and is normal when the tube is working hard. The output image is of the order of 20,000 times brighter than the input when the tube is working efficiently.

40. The final image from the intensifier tube is viewed through the monocular assembly eyepiece lens which can be adjusted (see para 51) to accommodate variations in the users eyes.

41. The optical magnification factor of the NVG is unity; the image presented to the user is actual size.

Caution

42. Where there is sufficient light to see with the naked eye, damage may occur to the unprotected image intensifier tubes even without power applied. NVGs should never be exposed to direct sunlight or pointed at the sun, even when switched off. Lens caps or daylight training filters should be fitted to the NVG when not in use to prevent inadvertent damage. They should also be protected from extremes of heat.

NVG Cockpit Compatibility

43. By careful filtering of cockpit lighting and the NVGs, compatibility is achieved which allows the cockpit instruments and warning captions to be readable with the unassisted eye whilst at the same time not affecting the performance of the NVGs. This is achieved by filtering the NVGs to amplify only light with a wavelength greater than around 665 nm, from red through to near-infra-red. This makes optimum use of the light energy available from the night sky. Conversely, the cockpit lighting is filtered to prevent emissions in this part of the spectrum, making use of shorter wavelengths and avoiding most red light. Thus NVG compatible lighting is predominantly green in colour, although special NVG compatible red warning captions have now been developed which can be seen even in bright daylight conditions.

Associated Aircraft Modifications

44. There are several considerations to the integration of NVGs to a given aircraft type.
 - a. Cockpit lighting must be compatible with NVG operation (see para 30).
 - b. External lighting needs to be NVG compatible, both visible aircraft lights and Infra-red (IR) light which is invisible to the naked eye. All external lighting must be considered including formation, anti-collision, taxi and landing lights.
 - c. Cockpit stowage of NVGs needs to be considered when the equipment is not in use.
 - d. Helmets are required to provide a mounting platform for the NVGs.

Technical Description

45. The following paragraphs describe NVGs in service at the time of writing. Users should refer to DAP 112G-1623-123 ; Night Vision Goggles, for the latest information.

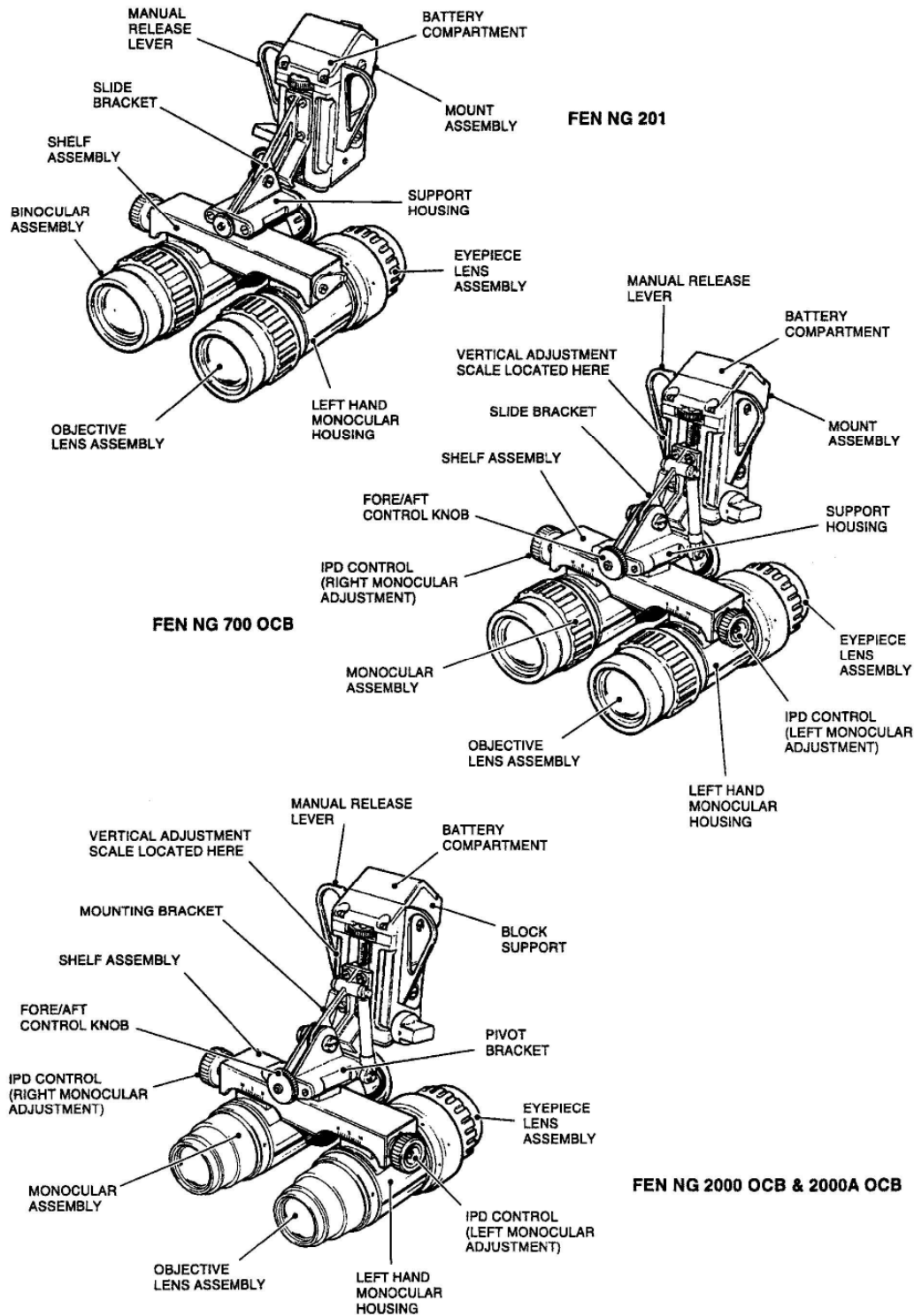
46. FEN NG 201 goggles (see fig 14) for fast-jet use consists of a binocular assembly, and utilise an auto detach bracket which embodies a gas motor that causes separation from the helmet during unpremeditated ejection. The objective lens can be focused anywhere from around 1 m distance to infinity and the eyepiece diopetre setting can be adjusted to accommodate the normal range of eyesight variations between individual users.

47. FEN NG 700 OCB is very similar but with a manual detach mount and the ON/OFF switch mounted on the left side of the bracket to facilitate its use in rotary wing aircraft. Focusing range is the same as FEN NG 201 and eyepiece diopetre range is similar.

48. FEN NG 2000 OCB is a lightened version of FEN NG 700 OCB, being fitted with the same mounting bracket but smaller objective lenses with focus fixed at infinity and is intended for use by the Royal Navy. These may be fitted with a neutral density filter which allows the goggles to be switched on in daylight conditions, however on bright sunny days, tube life may be reduced by burning if the sun or its reflection is imaged directly.

49. FEN NG 2000A OCB is further reduced in weight by utilising a lighter mounting bracket with the same smaller fixed focus objective lenses and is primarily for use in Hercules J type aircraft.

7-17 Fig 14 NVG Types



Monocular housing

50. The monocular housing (Fig 14) holds the intensifier tube and supports both the eyepiece and objective lens. Each monocular housing is fitted with a bleed nipple to enable purging with dry nitrogen. A gas-tight seal is provided by integrated O-ring seals fitted between the monocular housing and the objective and eyepiece lens cells. Purging with dry nitrogen is carried out to exclude air and water vapour. Water vapour can render the NVG unserviceable by encouraging mould growth on the optical coatings of the lens, or causing breakdown of the insulation within the EHT power supply leading to total failure of the intensifier tube. Air contains small amounts of helium, which has very

small molecules that are very difficult to exclude from the vacuum of the intensifier tube as they can diffuse through the seals. Once gas enters the vacuum it inhibits the acceleration of electrons, reducing tube gain, ultimately leading to shaded areas or complete tube failure which is why nitrogen purging is important.

NVG Adjustment

51. NVGs can be adjusted in the vertical and fore and aft axis. The fore and aft axis is particularly important as it allows the NVG to be adjusted for adequate clearance from Face Protection Visors (FPV) or aircrew respirators and for optimum field-of-view. FEN NGs 700 OCB, 2000 OCB and 2000A OCB are designed to be used at 30 mm eye relief, allowing face protection visors or aircrew NBC respirators to be fitted behind the NVGs. FEN NG 201 is designed to be used at 25 mm eye relief but can be used at 30 mm eye relief, allowing these items to be fitted but with a consequential reduction in the field of view, e.g. at 30 mm eye relief the FOV. is $39.5^{\circ} \pm 2.8^{\circ}$. As a general rule, eye relief should be adjusted to provide a 3 to 5 mm clearance between the FPV/respirator transparency and the eyepiece lens. This will give a maximum intensified FOV with minimum vignetting (obscuration of the intensified image) while allowing peripheral viewing of the cockpit instruments. The distance between the monocular sub-assemblies can be adjusted, with the monoculars being simultaneously moved inwards or outwards to suit the user's eye spacing. This is known as the Interpupillary Distance (IPD) Adjustment. Tilt adjustment is provided to enable the user to alter the sightline of the NVGs in elevation.

Face Protection Visor

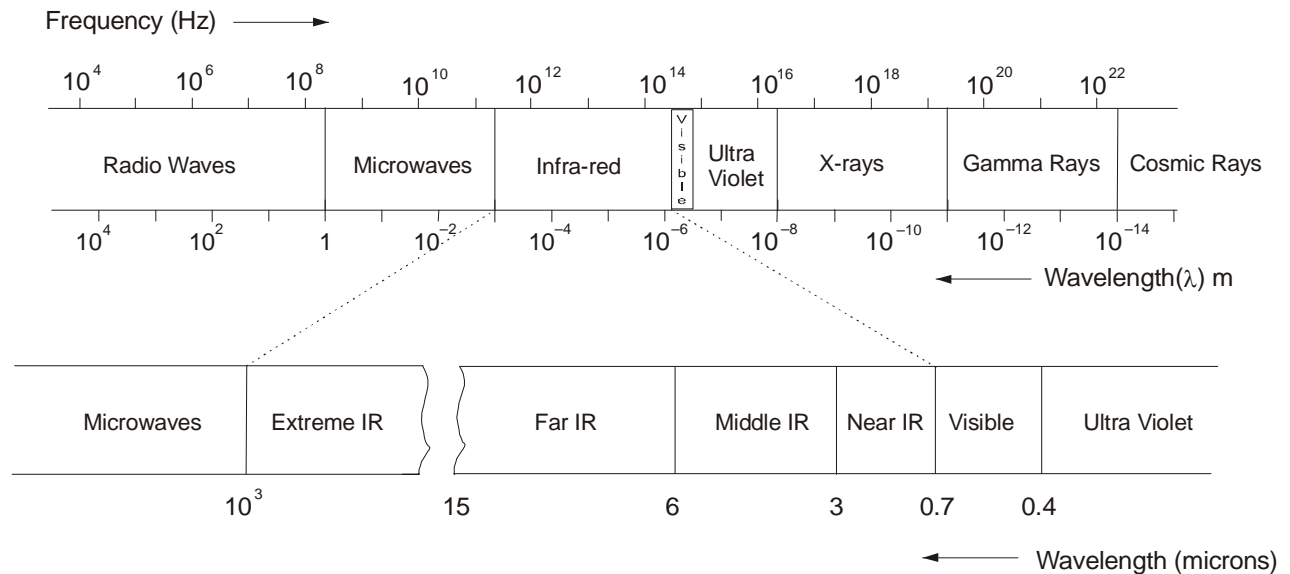
52. The NVG mount assembly prevents the deployment of the standard helmet visor when attached to the helmet and it is recommended that a FPV is worn instead. The FPV is made of polycarbonate, and when used in conjunction with a Mk 4 series aircrew protective helmet, provides adequate eye and face protection from wind blast at speeds of up to 400 kts and from the NVGs in the event of a survivable crash. The FPV is compatible with aircrew NBC spectacles but not with conventional aircrew spectacles.

CHAPTER 18 - INFRA-RED RADIATION

Characteristics of Infra-red Radiation

1. Infra-red (IR) radiation is electro-magnetic radiation and occupies that part of the electro-magnetic spectrum between visible light and microwaves. The IR part of the spectrum is sub-divided into Near IR, Middle IR, Far IR and Extreme IR. The position and division of the IR band, together with the appropriate wavelengths and frequencies, is shown in Fig 1.

7-18 Fig 1 Infra-red in the Electromagnetic Spectrum



Note: One micron (μ) = 10^{-6} metres and is now known as one micro-metre in the SI system.

2. All bodies with a temperature greater than absolute zero (0 K, -273°C) emit IR radiation and it may be propagated both in a vacuum and in a physical medium. As a part of the electro-magnetic spectrum it shares many of the attributes of, for example, light and radio waves; thus it can be reflected, refracted, diffracted and polarized, and it can be transmitted through many materials which are opaque to visible light.

Absorption and Emission

3. **Black body.** The radiation incident upon a body can be absorbed, reflected or transmitted by that body. If a body absorbs all of the incident radiation then it is termed a 'black body'. A black body is also an ideal emitter in that the radiation from a black body is greater than that from any other similar body at the same temperature.

4. **Emissivity (ϵ).** In IR, the black body is used as a standard and its absorbing and emitting efficiency is said to be unity; i.e. $\epsilon = 1$. Objects which are less efficient radiators, ($\epsilon < 1$), are termed 'grey bodies'. Emissivity is a function of the type of material and its surface finish, and it can vary with wavelength and temperature. When ϵ varies with wavelength the body is termed a selective radiator. The ϵ for metals is low, typically 0.1, and increases with increasing temperature; the ϵ for non-metals is high, typically 0.9, and decreases with increasing temperature.

Spectral Emittance

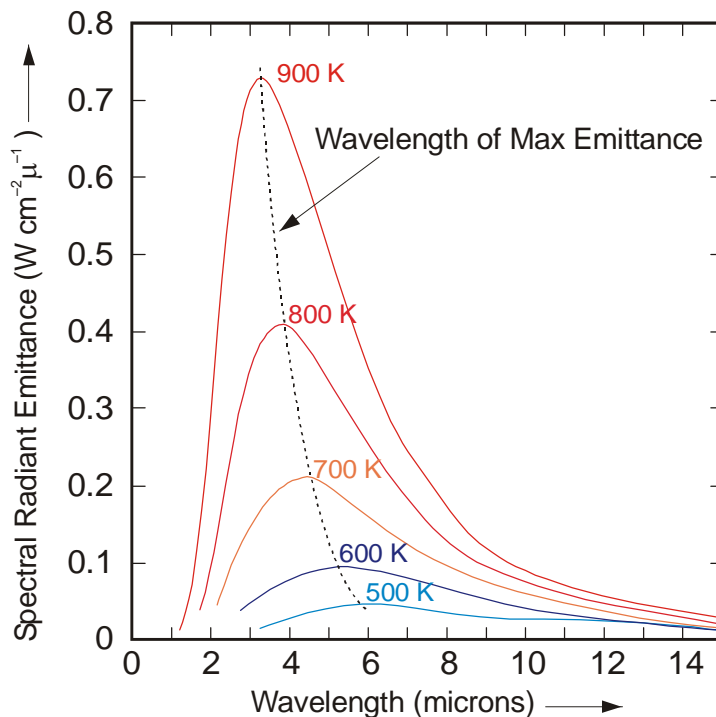
5. **Planck's Law.** A black body whose temperature is above absolute zero emits IR radiation over a range of wavelengths with different amounts of energy radiated at each wavelength. A description of this energy distribution is provided by the spectral emittance, W_λ , which is the power emitted by unit area of the radiating surface, per unit interval of wavelength. Max Planck determined that the distribution of energy is governed by the equation:

$$W_\lambda = \frac{2\pi c^2 h}{\lambda^5} \left(e^{\frac{hc}{kT\lambda}} - 1 \right)^{-1}$$

where λ = Wavelength
 h = Planck's constant
 T = Absolute temperature
 c = Velocity of light
 k = Boltzmann's constant

6. **Temperature/Emittance Relationship.** This rather complex relationship is best shown graphically, as in Fig 2 in which the spectral emittance is plotted against wavelength for a variety of temperatures. It will be seen that the total emittance, which is given by the area under the curve, increases rapidly with increasing temperature and that the wavelength of maximum emittance shifts towards the shorter wavelengths as the temperature is increased.

7-18 Fig 2 Distribution of IR Energy with Temperature



7. **Stefan-Boltzmann Law.** The total emittance of a black body is obtained by integrating the Max Planck equation which gives the result:

$$W = \sigma T^4$$

where W = Total emittance
 σ = Stefan Boltzmann constant
 T = Absolute temperature

For a grey body, the total radiant emittance is modified by the emissivity, thus:

$$W = \epsilon \sigma T^4$$

8. **Wien's Displacement Law.** The wavelength corresponding to the peak of radiation is governed by Wien's displacement law which states that the wavelength of peak radiation (λ_m), multiplied by the absolute temperature is a constant. Thus:

$$\lambda_m T = 2900 \mu K$$

By substituting $\lambda_m = 2900/T$ into Planck's expression it is found that:

$$W\lambda_m = 1.3 \times 10^{-15} T^5 \text{ expressed in Watts cm}^{-2}\mu^{-1}$$

ie the maximum spectral radiant emittance depends upon the fifth power of the temperature.

Geometric Spreading

9. The laws so far discussed relate to the radiation intensity at the surface of the radiating object. In general, radiation is detected at some distance from the object and the radiation intensity decreases with distance from the source as it spreads into an ever-increasing volume of space. Two types of source are of interest; the point source and the plane extended source.

10. **Point Source.** A point source radiates uniformly into a spherical volume. In this case the intensity of radiation varies as the inverse square of the distance between source and detector.

11. **Plane Extended Source.** When the radiating surface is a plane of finite dimensions radiating uniformly from all parts of the surface then the radiant intensity received by a detector varies with the angle between the line of sight and the normal to the surface. For a source of area A the total radiant emittance is WA . The radiant emittance received at a distance d and at an angle θ from the normal is given by:

$$\frac{WA}{2\pi d^2} \cos\theta$$

IR Sources

12. It is convenient to classify IR sources by the part they play in IR systems; ie as targets, as background, or as controlled sources. A target is an object which is to be detected, located or identified by means of IR techniques, while a background is any distribution or pattern of radiation, external to the observing equipment, which is capable of interfering with the desired observations. Clearly what might be considered a target in one situation could be regarded as background in another. As an example terrain features would be regarded as targets in a reconnaissance application but would be background in a low-level air intercept situation. Controlled sources are those which supply the power required for active IR systems (e.g. communications), or provide the standard for calibrating IR devices.

Targets

13. **Aircraft Target.** A supersonic aircraft generates three principle sources of detectable and usable IR energy. The typical jet pipe temperature of 773 K produces a peak of radiation, (from Wien's law), at 3.75μ . The exhaust plume produces two peaks generated by the gas constituents; one at 2.5 to 3.2μ due to carbon dioxide, the other at 4.2 to 4.5μ due to water vapour. The third source is due to leading edge kinetic heating giving a typical temperature of 338 K with a corresponding radiation peak at about 7μ .

14. **Reconnaissance.** Terrestrial IR reconnaissance and imaging relies on the IR radiation from the Earth which has a typical temperature of 300 K. The peak of radiation corresponding to this temperature is about 10μ and so systems must be designed to work at this wavelength.

Background Sources

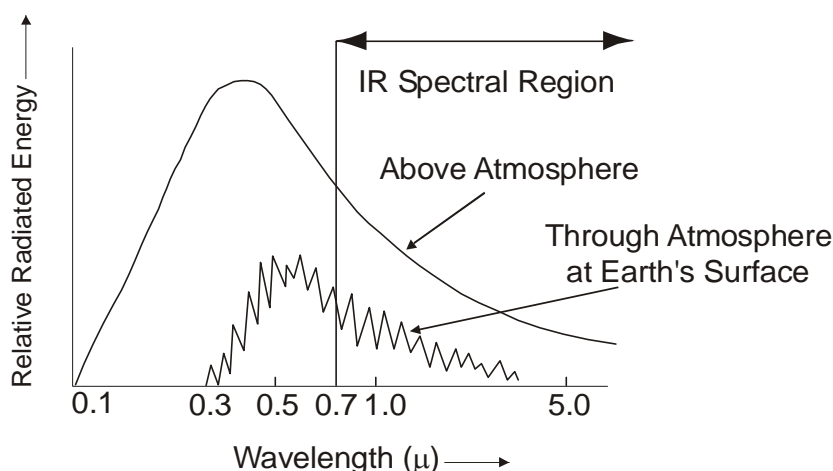
15. Regardless of the nature of the target source, a certain amount of background or interfering radiation will be present, appearing in the detection system as noise. The natural sources which produce this background radiation may be broadly classified as terrestrial or atmospheric and celestial.

16. **Terrestrial Sources.** Whenever an IR system is looking below the horizon it encounters the terrestrial background radiation. As all terrestrial constituents are above absolute zero they will radiate in the infra-red, and in addition IR radiation from the sun will be reflected. Green vegetation is a particularly strong reflector which accounts for its bright image in IR photographs or imaging systems. Conversely, water, which is a good reflector in the visible part of the spectrum, is a good absorber of IR, and therefore appears dark in IR images.

17. **Atmospheric and Celestial Sources.** Whenever an IR device looks above the horizon the sky provides the background radiation. The radiation characteristics of celestial sources depend on the source temperature together with modifications by the atmosphere.

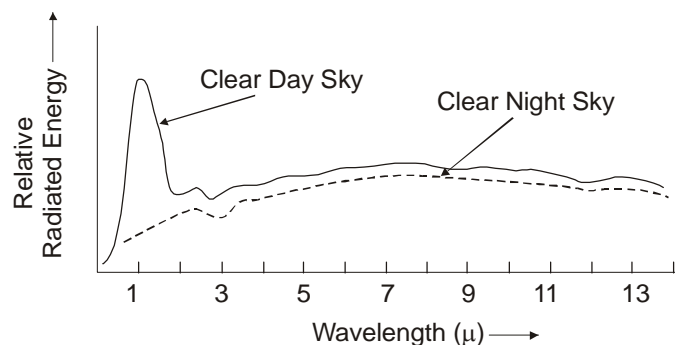
- a. **The Sun.** The sun approximates to a black body radiator at a temperature of 6,000 K and thus has a peak of radiation at 0.5μ , which corresponds to yellow-green light. The distribution of energy is shown in Fig 3 from which it will be seen that half of the radiant power occurs in the infra-red. The Earth's atmosphere changes the spectrum by absorption, scattering and some re-radiation such that although the distribution curve has essentially the same shape, the intensity is decreased and the shorter, ultraviolet, wavelengths are filtered out. The proportion of IR energy remains the same or perhaps may be slightly higher. Sunlight reflected from clouds, terrain and sea shows a similar energy distribution.

7-18 Fig 3 Spectral Distribution of Solar Radiation



- b. **The Moon.** The bulk of the energy received from the moon is re-radiated solar radiation, modified by reflection from the lunar surface, slight absorption by any lunar atmosphere and by the Earth's atmosphere. The moon is also a natural radiating source with a lunar daytime surface temperature up to 373 K and lunar night time temperature of about 120 K. The near sub-surface temperature remains constant at 230 K, corresponding to peak radiation at 12.6μ .
- b. **Sky.** Fig 4 shows a comparison of the spectral distribution due to a clear day and a clear night sky. At night, the short wavelength background radiation caused by the scattering of sunlight by air molecules, dust and other particles, disappears. At night there is a tendency for the Earth's surface and the atmosphere to blend with a loss of horizon since both are at the same temperature and have similar emissivities.
- c.

7-18 Fig 4 Spectral Energy Distribution of Background Radiation from the Sky



- d. **Clouds.** Clouds produce considerable variation in sky background, both by day and by night, with the greatest effect occurring at wavelengths shorter than 3μ due to solar radiation reflected from cloud surfaces. At wavelengths longer than 3μ , the background radiation intensity caused by clouds is higher than that of the clear sky. Low bright clouds produce a larger increase in background radiation intensity at this wavelength than do darker or higher clouds. As the cloud formation changes the sky background changes and the IR observer is presented with a varying background both in time and space. The most serious cloud effect on IR detection systems is that of the bright cloud edge. A small local area of IR radiation is produced which may be comparable in area to that of the target, and also brighter. Early IR homing missiles showed a greater affinity for cumulus cloud types than the target aircraft. Discrimination from this background effect requires the use of spectral and spatial filtering.

IR Transmission in the Atmosphere

18. Atmospheric Absorption. The periodic motions of the electrons in the atoms of a substance, vibrating and rotating at certain frequencies, give rise to the radiation of electro-magnetic waves at the same frequencies. However, the constituents of the Earth's atmosphere also contain electrons which have certain natural frequencies. When these natural frequencies are matched by those of the radiation which strikes them, resonance absorption occurs and the energy is re-radiated in all directions. The effect of this phenomenon is to attenuate certain IR frequencies. Water vapour and carbon dioxide are the principle attenuators of IR radiation in the atmosphere. Figs 5a, 5b and 5c show the transmission characteristics of the atmosphere at sea-level, at 30,000 ft and at 40,000 ft.

19. Scattering. The amount of scattering depends upon particle size and particles in the atmosphere are rarely bigger than 0.5μ , and thus they have little effect on wavelengths of 3μ or greater. However, once moisture condenses on to the particles to form fog or clouds, the droplet size can range between 0.5 and 80μ , with the peak of the size distribution between 5 and 15μ . Thus fog and cloud particles are comparable in size to IR wavelengths and transmittance becomes poor. Raindrops are considerably

larger than IR wavelengths and consequently scattering is not so pronounced. Rain, however, tends to even out the temperature difference between a target and its surroundings.

7-18 Fig 5 Atmospheric Transmittance vs. Altitude

Fig 5a Transmittance at Sea Level

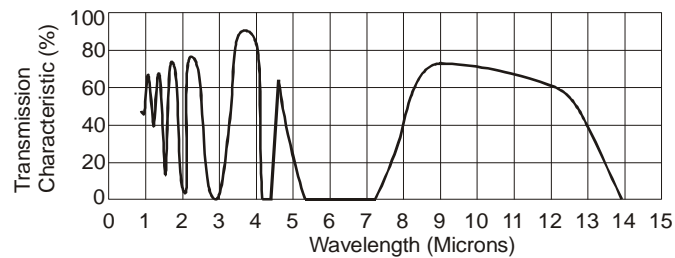


Fig 5b Transmittance at 30,000 ft

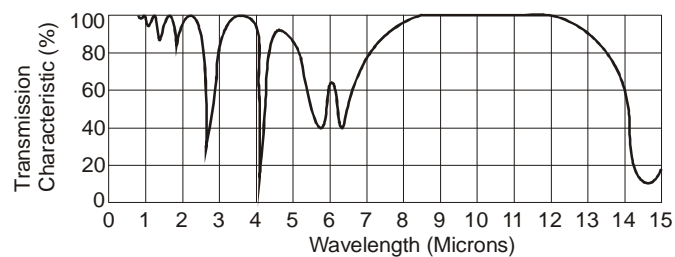
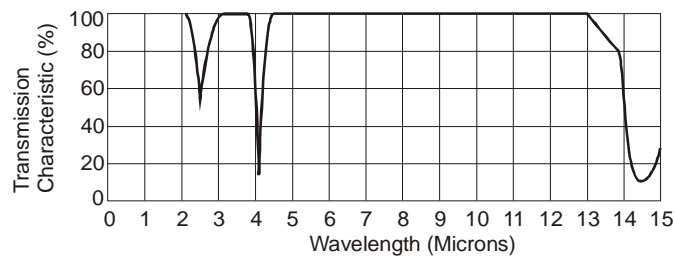


Fig 5c Transmittance at 40,000 ft



20. **Scintillation.** Where a beam of IR passes through regions of temperature variation it is refracted from its original direction. Since such regions of air are unstable, the deviation of the beam is a random, time varying quantity. The effect is most pronounced when the line of sight passes close to the earth and gives rise to unwanted modulations of the signal, and incorrect direction information for distant targets.

CHAPTER 19 - INFRA-RED SYSTEMS

Introduction

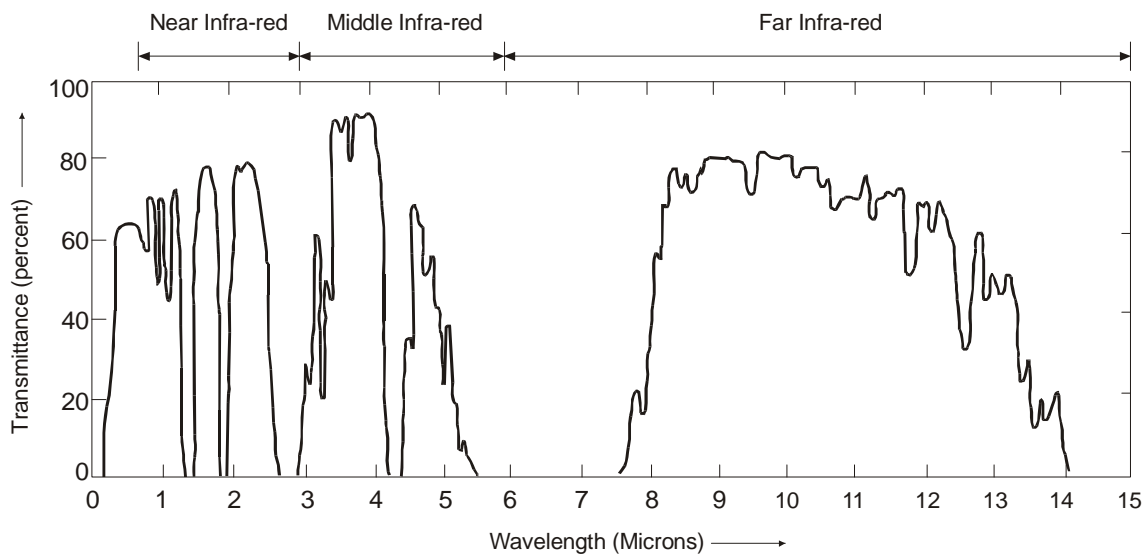
1. Objects with a temperature above absolute zero ($-273\text{ }^{\circ}\text{C}$) will emit infra-red (IR) radiation and, in addition, will reflect or absorb incident IR radiation to varying degrees. An IR sensing system can use these variations in emitted and reflected IR to form an image, in the same way that a sensor operating within the visible part of the spectrum uses variations in visual brightness to form an image. IR imaging systems, therefore, rely on detecting differences in IR intensity, rather than on measuring absolute values. Such systems have the advantages of being independent of natural or artificial visual illumination, and are not easily deceived by camouflage.
2. The basic physics of IR are covered in Volume 7, Chapter 18, where it is shown that the wavelength at which maximum radiation occurs (λ_{max}) is a function of the absolute temperature (T), and is governed by Wien's Displacement Law:

$$\lambda_{\text{max}} = \frac{2900}{T} \mu\text{m}$$

By substituting typical terrestrial temperatures (270 to 300 K) into this equation, it will be seen that the peak of radiation occurs in the far infra-red range between $9\text{ }\mu\text{m}$ and $11\text{ }\mu\text{m}$.

3. Fig 1 shows the relative transmittance of IR wavelengths in the atmosphere at sea level. Terrestrial IR falls within a transmission 'window' which covers wavelengths from about $8\text{ }\mu\text{m}$ to $13\text{ }\mu\text{m}$. IR imaging systems have, therefore, to be designed to operate in this band.

7-19 Fig 1 Relative Transmittance of IR Wavelengths in the Atmosphere (at Sea Level)



Atmospheric Attenuation

4. The transmission of IR energy through the atmosphere may be impeded due to scattering by suspended particles and absorption by constituent gases. If the suspended particles are small, as in haze and battlefield smoke, an IR system will not be affected, even in what appears as zero visibility to the naked eye. If the particles are larger, as in cloud, fog, mist, and rain, IR will be no better than the eye, and the range of the system will be limited to the visual range. In addition, some energy in the wavelengths between $8\text{ }\mu\text{m}$ and $13\text{ }\mu\text{m}$ is absorbed by gaseous water vapour.

Imagery Interpretation

5. Detecting and identifying IR images requires the same integration of size, shape, shadow, surroundings, and tone as is employed in the interpretation of conventional, visual-light imagery.

6. When considering size and shape, there are two aspects which must be considered. Firstly, the resolution of IR images is lower than that of visual images and some detail may be lost. Secondly, the true size and shape of hot objects may be exaggerated or disguised by blooming and halation effects. Blooming is the spreading of light around bright objects in an otherwise dark scene, while halation is the formation of 'haloes' around bright objects.

7. Thermal shadows often mirror their visible counterparts, and they are caused by areas being shaded from direct radiation. As with visual shadows, they are liable to change with changes in the direction of illumination, and to dissipate when the illumination lessens, eg after sunset or with cloud cover. The rate of dissipation will, however, vary with the physical characteristics of the shaded area (thermal conductivity and capacity), and with the meteorological conditions. Thermal shadows from movable objects (eg aircraft on the ground) can often be seen long after the object has moved.

8. The key factor in interpreting IR images is relative tone. The common materials encountered in night IR imagery are metal, pavement (e.g. runways), soil, grass, trees, and water and, although the appearance of these can vary with meteorological and physical parameters, some generalizations can be made (assuming the conventional cold is black, hot is white display):

a. **Metal Surfaces.** Thin, unheated, metal appears black, as metals have low emissivity compared to other substances. Although such surfaces are good reflectors and will reflect energy incident upon them from the sky, the intensity of radiation at night is quite low, so any reflected component will also be weak. Metal could reflect radiation from a nearby warm object, but this effect does not occur often enough to be significant.

b. **Pavement.** Pavement has relatively high emissivity and is in good thermal contact with the earth which acts as a constant heat source. It also has a high thermal capacity, and therefore retains any heat received from the sun during the day. These characteristics are generally true for all types of pavement, including concrete and asphalt, therefore they all appear light grey to white in IR imagery.

c. **Soil.** Soil, including most earths, sand, and rock, have the same characteristics as pavement and similarly appear in a light grey tone.

d. **Grass.** Grass has poor thermal contact with the ground and cools rapidly by radiation. It therefore appears black.

e. **Trees.** Trees appear medium to light grey. The tone is thought to be the result of a number of causes. There is some convective warming of the trees by the air in conjunction with the night temperature inversion (air temperature at night is usually lower at the surface than it is a few feet above the ground), some retention of heat from daytime solar heating, and some heat generation from the trees' life processes. In daytime imagery, the same leaves appear colder than the ground as temperature at tree-top height is cooler than at ground level.

f. **Water.** In night IR imagery, water ranges in tone from light grey to white, as a result of its high emissivity and good heat transfer properties. Conversely, by day, water appears dark; the

main source of IR energy in daytime is reflected solar radiation, and water is a poor reflector at IR wavelengths.

9. Man-made, heated objects, such as buildings and vehicles, are easily seen as IR images. However, the same applies to living things, such as farm animals, and it can be difficult to distinguish between these groups, other than at close range, due to the relatively poor IR resolution, and the blooming and halation effects referred to earlier.

FORWARD LOOKING INFRA-RED (FLIR)

Optical System Requirements

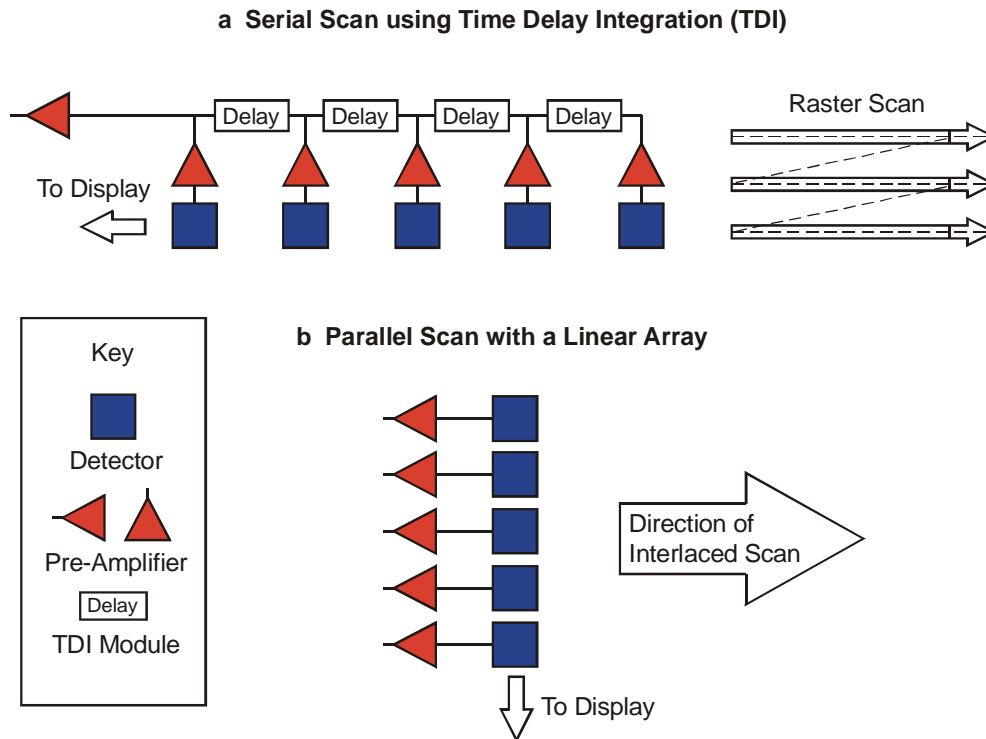
10. The optical system collects the IR radiation and focuses an image onto one or more detectors. The problems inherent in achieving this are more severe than those encountered in visible-light optics. The bandwidth in which FLIR operates (8 μm to 13 μm) is significantly wider than that in the visible band (0.4 μm to 0.7 μm), and this increases the optical design problems. Conventional glass is essentially opaque to IR at wavelengths in excess of 3 μm , thus special materials must be employed; these generally have higher refractive indices than glass. The combination of relatively wide bandwidth operation and high refractive index optics, results in much greater aberrations than have to be contended with in visible-light optics. Furthermore, many of the materials suitable for IR optics are not ideal for military and airborne applications, where they may be subject to pressure, vibration, shock, and extremes of temperature.

IR Detection

11. The detectors used in FLIR systems are made of semi-conductor material and are photo-conductive, i.e. their electrical conductivity increases in proportion to the number of incident IR photons. Cooling is essential for optimum performance. Two of the main characteristics of a detector are its time constant and its detectivity. The time constant is a measure of the time required for the detector to respond to radiation on its surface; detectivity is essentially a measure of sensitivity, ie the amount of incident IR energy necessary to generate an output signal over and above the detector noise.

12. Although IR Charge Coupled Devices (CCD) are under development, current FLIR systems use a rectilinear scanning action to examine the field of view. The two most common techniques are:

- a. Serial processing (Fig 2a), in which a single detector, or a linear array using time delay integration (TDI), is used to scan a scene line by line and element by element.
- b. Parallel processing (Fig 2b), in which an array of detectors is used to scan the scene element by element. Usually, the elevation field of view is covered by the line array and the scanning action covers the azimuth field of view.

7-19 Fig 2 FLIR Scanning Principle for Parallel Processing

13. Serial processing produces a good quality TV-like display, but requires very high speed scanning, calling for detectors of very short time constant or high detectivity. Parallel processing allows a slower scan rate, but requires multiple channels to handle the simultaneous line output signals, and the signal processing is complicated by the inevitable differences in performance of the detectors forming the array.

14. In practice, the detectors remain stationary and the scanning is achieved by mechanical movement of an optical system. Such electro-mechanical hardware carries a significant penalty in terms of reliability and weight, and much effort is being directed towards the development of non-scanning, solid state, devices, which, in addition, may dispense with the need for cooling.

Display

15. FLIR imagery is presented on a CRT as a real-time raster scan, in a similar manner to TV. The display is usually arranged such that an increase in relative temperature corresponds to a transition from black, through shades of grey, to white. This tonal correspondence can be reversed, if it is considered appropriate for particular applications.

INFRA-RED LINESCAN (IRLS)

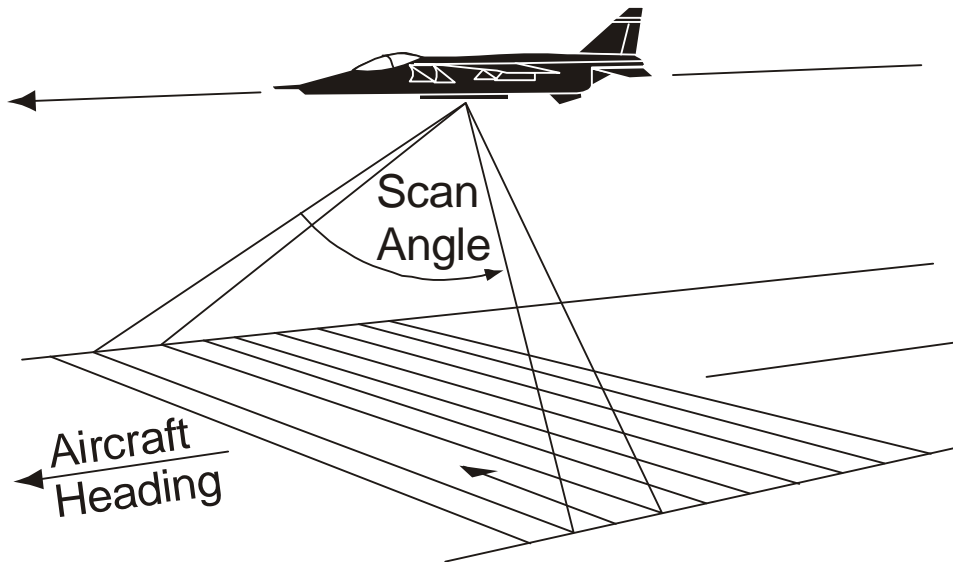
Introduction

16. Infra-red Linescan (IRLS) is a passive, airborne, infra-red mapping system that scans the ground along the flight path and produces a high resolution film map of the terrain. Radiated infra-red energy from the ground is received by an optical scanner in the IRLS and detected in the $8\text{ }\mu\text{m}$ to $13\text{ }\mu\text{m}$ band

Recording and Display

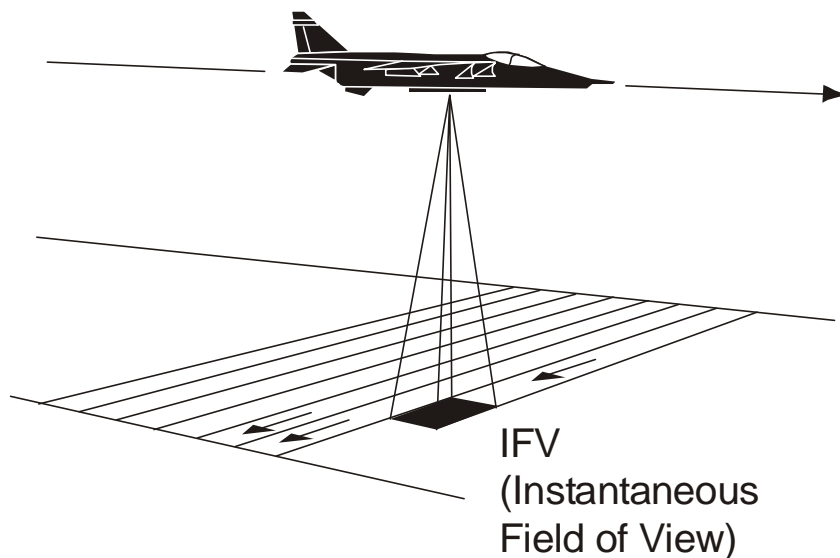
17. Instead of an instantaneous view of the entire area beneath the aircraft, narrow strips are scanned by a rotating mirror assembly. These scans are perpendicular to the heading of the aircraft. The forward motion of the aircraft accounts for the scanning strips being parallel to each other in the line of flight. This linescan procedure is shown in Fig 3.

7-19 Fig 3 Infra-red Linescan Techniques



The equipment scans an area of about 60° either side of the nadir (a total 120° scan). At any instant, the optical system is collecting IR radiation emitted from a small rectangular area on the ground. This area is called the 'Instantaneous Field of View' (IFV) (as shown in Fig 4). The IFV is an important factor in determining the resolution of the system, and is proportional to the altitude of the aircraft.

7-19 Fig 4 Instantaneous Field of View (IFV)



Height Limitation

18. The IRLS system is intended to be used at low levels, but is capable of being used at medium altitudes. The velocity-to-height (V/H) ratio is in the order of 0.06 to 1.6 kt per foot, equating to heights of between 200 ft and 5,300 ft at 320 kt.

Operation

19. **Detection.** The IR energy is focused, by the scanner optical system, on to two detectors which change the ground emissions into fluctuating electrical signals. Objects emit infra-red energy if they are above a temperature of absolute zero ($-273^{\circ}\text{C}/0\text{ K}$). In order to reduce interference from the inherent infra-red emissions coming from the detectors, they are cooled in a cryogenerator; helium is usually used to refrigerate them to as close to 0 K as possible. This action brings the sensitivity of the detectors within satisfactory limits. If the cooling system fails, the results become unacceptable in a very short time. One detector provides a wide, and the other a narrow, field of view of the ground along track. Infra-red energy in the $8\text{ }\mu\text{m}$ to $13\text{ }\mu\text{m}$ band, radiated from successive IFV, is received at the detectors in the IRLS. The size of the IFV at any given time is determined by which of the two detectors is selected. The selection is made automatically and depends on the V/H ratio.

20. **Video and Film Presentation.** The detected energy is processed in the receiver to obtain a video signal which is fed to a recorder where, after correction, the video is displayed on a CRT. The displayed video is optically focused onto film moving at a speed dependent on V/H. Flight data digital inputs (mainly aircraft position, height and heading) are projected by the optical system on to one edge of the film. Similarly, reference and event markers are projected onto the film between the video and data recordings. The result is an annotated continuous map of the terrain. An IRLS frame, with data markings suppressed, is shown in Fig 5, together with a comparative photograph.

7-19 Fig 5 Infra-red Linescan
Fig 5a Air-Ground Photograph of Line Feature

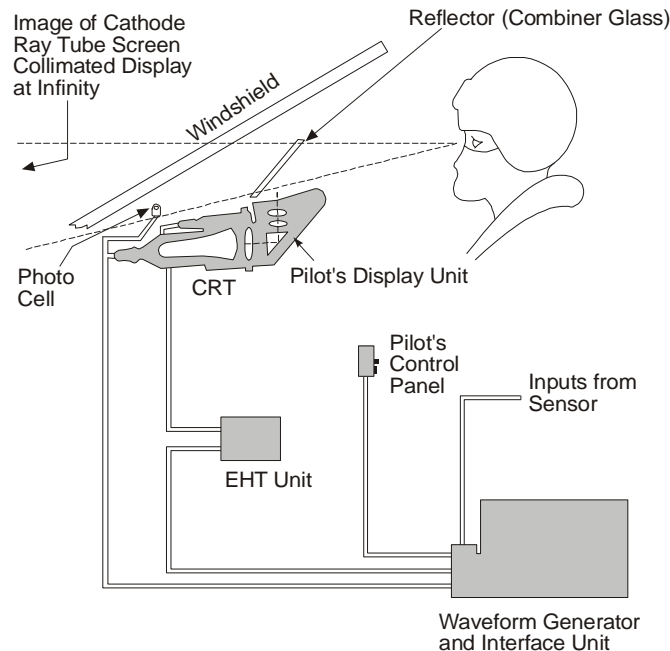


Fig 5b Infra-red Linescan Image of same Feature



1. Normal cockpit displays entail the pilot dividing his time between observing the outside world and reading the instruments. Thus, the pilot's eyes have frequently to switch between reading instruments situated at no more than a few feet away, and surveying the outside world, which is effectively at infinity. This requires not only a change of focus, but also an adjustment to light conditions which are often considerably different. It is a far more satisfactory arrangement if the instruments are read under the same conditions of focus and illumination as the outside world; this can be achieved by the use of head-up or helmet mounted displays.

2. The collimated head-up display (HUD) is a development of the gyro gunsight and is used to project an instrument display at the pilot's eye level. The symbols are produced in a waveform generator, displayed on a CRT, and reflected on a glass screen in front of the pilot. The symbols may be driven by a variety of aircraft sensors (eg IN, ADC, Radar, LRMTS) to provide aircraft attitude, altitude, and velocity, together with navigation and weapon aiming information. A control unit is provided to allow the pilot to select the appropriate symbols for any particular stage of flight. Initially, the display brightness can be adjusted manually by the pilot, after which it is controlled by a photocell to compensate for changes in the illumination of the outside scene. Fig 1 shows a block diagram of a typical HUD installation.



3. The pilot's display unit (PDU) incorporates a very bright CRT to ensure that the symbols can be viewed against a very high background brightness, equivalent to sunlight on cloud. The reflector, or combiner glass, is semi-transparent and reflects the CRT image while allowing the outside world to be viewed through it. The presented image is collimated, ie focused at infinity, so that the CRT symbols and the outside scene can be viewed as a composite image, without the need to change eye focus.

4. The optical system in the HUD may be either refractive (lenses and prisms) or diffractive (holographic); reflective optics have been used, but any advantages in terms of field of view (FOV) have been outweighed by considerations of size, cost, weight, and optical efficiency.

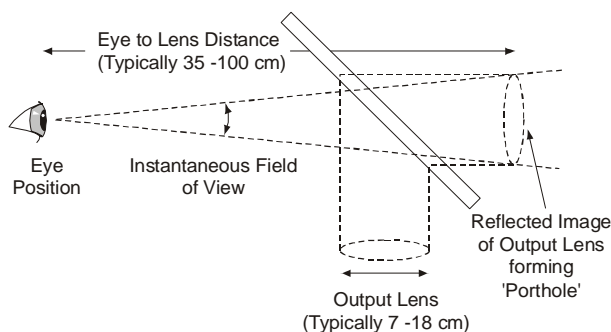
5. In addition to symbolic displays, the use of holographic technology has the potential to allow sensor imagery to be shown, such as LLTV, FLIR, or radar.

Refractive Optics

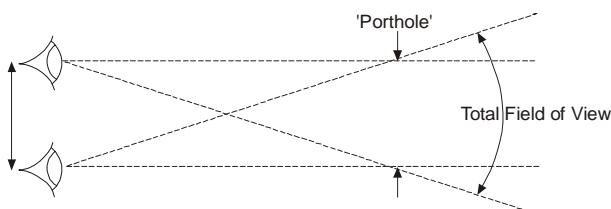
6. The use of refractive optics is still the most common technique although there are disadvantages in terms of restricted field of view, low optical efficiency, and bulky, heavy components.

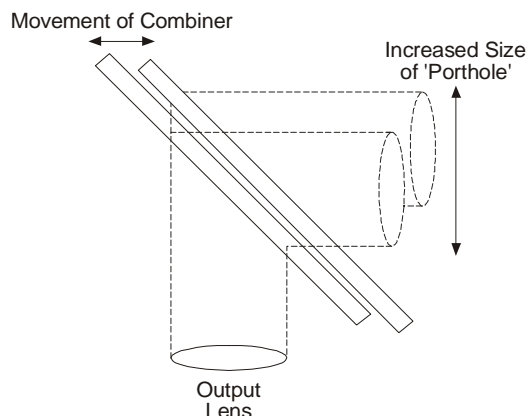
7. **Field of View.** The field of view of a conventional refractive HUD is determined principally by the size of the output lens, and the distance between it and the eye (via the combiner). The reflected output lens acts as a porthole through which the virtual image produced by the HUD is viewed (Fig 2). As an example, for a 12 cm diameter lens at an eye to lens distance of 70 cm, the single eye FOV will be approximately 10° . In practice, the total FOV in azimuth will be extended due to the separation between the pilot's eyes, and a further increase will result from small head movements (Fig 3). Some PDUs increase the vertical FOV by using a movable combiner glass. A servomechanism moves the glass, thus shifting the FOV in the vertical plane and increasing the total, but not the instantaneous, FOV (Fig 4). The major problem with a limited FOV is that of marking a target, or updating from a visual pinpoint, which is at a large angle-off from the aircraft centre line. In addition, the effect of the porthole and the resultant restrictions on head movement can be tedious for the pilot.

7-20 Fig 2 Single Eye Instantaneous Field of View



7-20 Fig 3 Increased FOV due to Binocular Vision and Head Movement



7-20 Fig 4 Increasing the Vertical FOV by Moving the Combiner

8. **Optical Efficiency.** In any optical system, there will be losses in light transmission. Typically, only 40% of the light produced by the CRT will reach the pilot's eye, and to compensate for this loss the CRT must be run at a very high output level, leading to a reduction in its life. Light entering from the outside scene may be reduced to about 70% which may cause a significant reduction in forward visibility.

9. **Size and Weight.** High quality lenses and prisms are heavy and expensive items, and since the output lens and the associated optics must be mounted on the pilot's side of the combiner, they tend to protrude into the cockpit. The equipment must be installed such that adequate clearance for ejection is maintained, while at the same time being close enough to the pilot's eye to yield an acceptable FOV.

Diffraction Optics

10. The trend in HUD construction is towards the use of diffractive optics in which a holographic element, tuned to the frequency of the CRT light output, is used as the combiner. Compared to the refractive system, the holographic combiner has a higher transmission efficiency, improved reflectivity, and variable geometry.

11. The combiner is produced by exposing a film of photosensitive emulsion to laser light under specific conditions. The recorded diffraction pattern in the emulsion has the property of acting as a mirror to light of the same wavelength as the laser used in production, while being transparent to light of other wavelengths. After development, the film is sealed between glass plates, and the resulting unit is used as the combiner glass.

12. The reflectance of the narrow band of CRT frequencies can reach 80%, while the transmission of other frequencies from the outside world is typically in excess of 90%. Thus, the technique allows CRTs to be run at lower power levels, with the attendant gains in life, and allows the outside scene to be viewed with only minimal reductions in brightness and contrast.

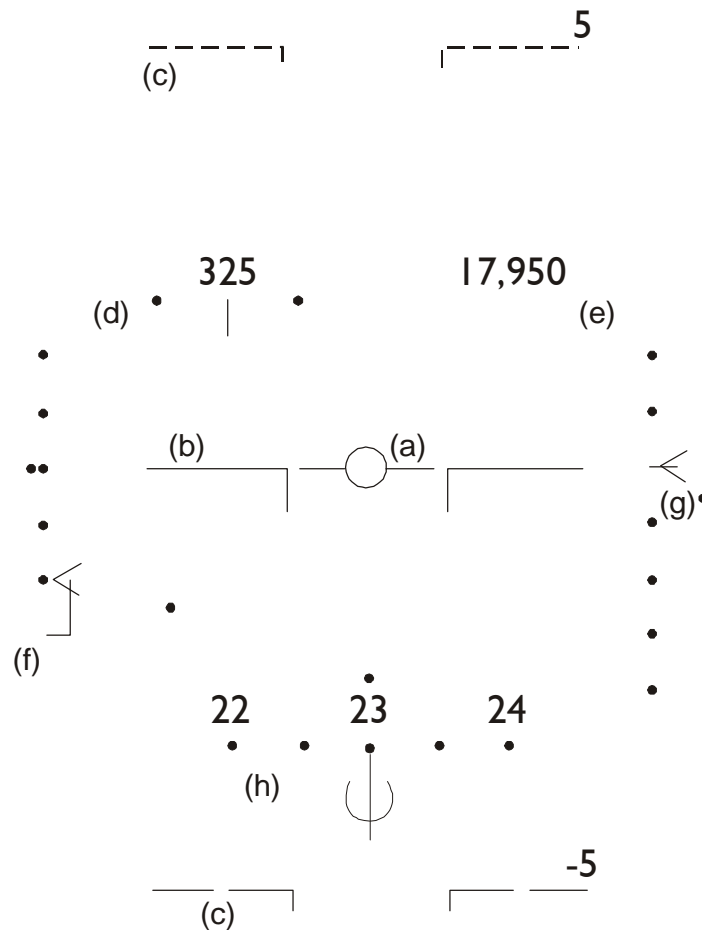
13. The element can be produced in either a curved or a flat form as necessary to fit the space available in the cockpit and this permits a wider FOV and less intrusion into the ejection line.

HUD Symbolology

14. A HUD can be designed to portray virtually any information in an infinite variety of formats. The format used will vary from manufacturer to manufacturer, and from aircraft type to aircraft type. Furthermore, the symbology may be amended during the lifetime of an aircraft as its role, or

equipments, change. It is not possible in this chapter to describe all of the displays available; rather a typical fast jet format will be illustrated in both a general and a weapon aiming mode.

7-20 Fig 5 Example of Typical UK HUD Symbolology - General Mode



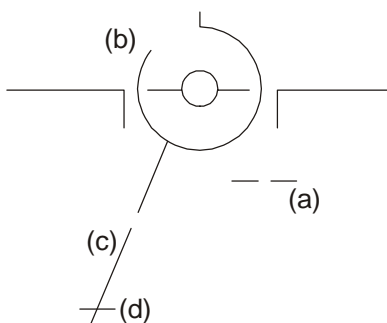
15. Fig 5 shows a typical HUD general mode which would be used during all stages of flight except for weapon delivery. The symbolology used is as follows:

- a. Aircraft symbol denoting either the fore and aft aircraft axis, the aircraft velocity vector, or some computed vector as required by a particular flight mode.
- b. Horizon bars, representing zero pitch.
- c. Pitch bars at 5° intervals with a 1:1 scaling.
- d. Airspeed indication, either IAS or Mach No, both as a digital read-out and as a pointer movement indicating rate of change.
- e. **Height.** As shown, the display indicates barometric height, but alternatively radar height may be shown, in which case the figures will be preceded by a letter 'R'.
- f. **Angle of attack.** The values associated with the scale will vary with aircraft type.
- g. **Vertical speed.** The values associated with the scale will vary with aircraft type.
- h. Heading (or track) scale with a superimposed steering bug (U).

16. A HUD will have a number of different modes and sub-modes, some of which will be selected automatically dependent on the mode of operation of the navigation and attack system, and others

which may be selected manually. An example of an air-to-ground weapon aiming mode is shown in Fig 6 with the following symbology:

7-20 Fig 6 HUD in Air-to-Ground Weapon Aiming Mode



- a. **Target bar.** The gapped target bar represents the system's computed target position. Once the pilot can see the target, delivery accuracy can be refined by changing phase and slewing the target bar, now a solid line, to overlie the target, where it will be stabilized by the system.
 - b. **Time circle.** The time circle unwinds anti-clockwise from 60 seconds to release (50 seconds to release illustrated).
 - c. **Impact line.** The impact line represents the track along which the weapons will fall, and the pilot's task is to fly the aircraft such that the impact line overlies the target position. The top of the line represents the minimum safe pass distance, and the gap $1\frac{1}{2}$ times the pass distance.
 - d. **Continuously Computed Impact Point (CCIP).** The CCIP represents the point on the ground where the weapons will impact if released at that instant.
17. The weapon is released, normally automatically, when the target bar and CCIP coincide. Until then, the pilot must ensure that the impact line overlies the target bar, and, for safe clearance, that the CCIP and target bar coincide before the target reaches the top of the impact line. In some systems and modes, additional symbols may be used, for example, to indicate LRMTS pointing and operation, air-to-air missile aiming, and gun aiming solutions, or to enable the navigation system to be updated by slewing the symbol to overlie a visual pinpoint.

HELMET MOUNTED DISPLAY (HMD) SYSTEMS

General

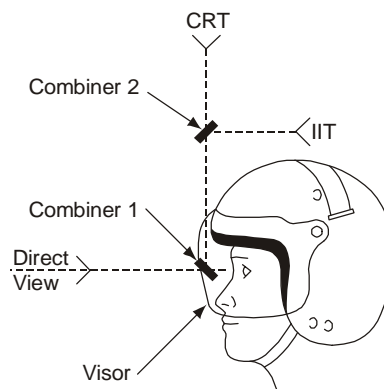
18. With the increasing complexity of airborne detection and display systems and the associated additional workload on the pilot, more and more designers are focusing on integrating sensor information into the flying helmet. This is aimed at removing the disadvantage of the Head-up Display (HUD) in that the display is only available to the pilot whilst he is looking at the HUD combiner and not when he looks away. Although not yet widespread in use, HMDS technology was first used operationally in attack helicopters where the need to meet ejection safety criteria did not exist. These HMD systems allow the pilot to benefit from displays of aircraft symbology superimposed, on demand, on his normal field of vision. HMD systems (often termed Integrated Helmet Systems (IHS)) become an inherent part of the aircraft avionics and weapons systems enabling off boresight weapon aiming, target designation, and pilot cueing, for example.

Displays

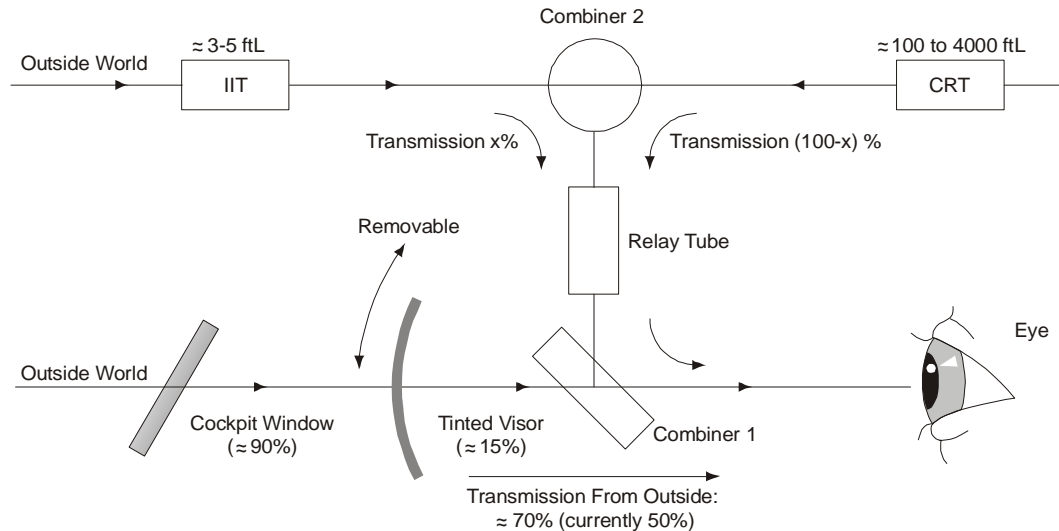
19. To permit aircraft to operate throughout the 24-hour spectrum, a HMD normally incorporates a miniature cathode ray tube (CRT) and an image intensifier tube (IIT). The display of thermal imagery

(TI) or output from other electro-optical (EO) sensors is provided to the pilot by means of the CRT. Miniature CRTs may present the forward-looking infra-red (FLIR) imagery and also provide flight and weapon aiming information in a similar manner to a conventional HUD. Alone, the TI may be dangerous for the 24-hour mission since the emissivity of natural materials will vary over the period. A so-called zero contrast (or washout effect) during rainfall is sometimes observed especially during twilight or at dawn. At these times, foreground is not detectable against background and, for example, pylons or cables become an extreme hazard. To overcome this, the IIT and TI may be combined. A HMD may be designed to allow the pilot to switch between IIT and TI at will, select both, or switch off the flight symbology altogether. In twilight or dawn periods, it might be better to present only one sensor at any one time. The IIT works on a different principle from the TI and is better suited to adverse weather conditions during night or twilight. Thus, a true IHS will be configured with the day and night capabilities combined as shown in Fig 7.

7-20 Fig 7 IHS Configuration



The TI and IIT images are integrated in Combiner 2 and the resultant image is superimposed on the direct view in Combiner 1. There are drawbacks, however. High brightness is required because of the complicated optical train that HMDs use whether the image is displayed on a combiner eyepiece or on the visor. Between source and projection, the pathway can attenuate both brightness and definition - affording a resolution of some 50% of that of the human eye. Moreover, once symbology is projected on to the eyepiece or visor, transmissivity to the real world is affected. The dichroic coatings necessary for image projection and the laser protection elements reduce real world transmissivity to about 70%. Clearly, some compromise and adjustment is necessary to provide the right balance of real world transmissivity and symbology brilliance. Fig 8 shows the combined optical paths and an example of their attenuation.

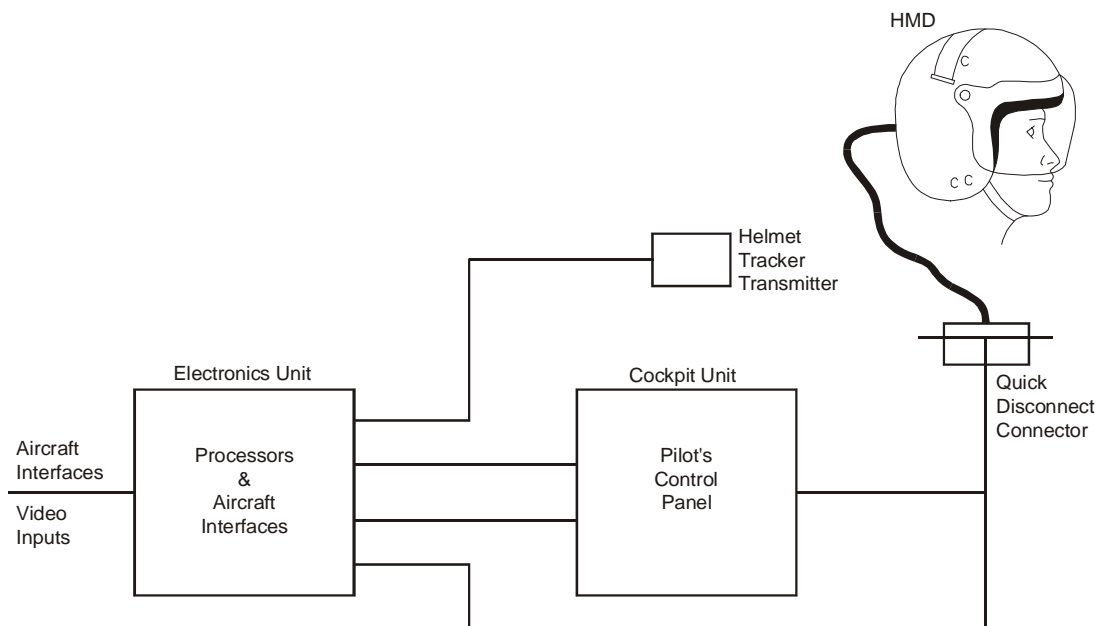
7-20 Fig 8 Optical Paths and Attenuation

Protection and Comfort

20. Wherever possible, all electro-optical parts are protected by the helmet shell. As few electronics as possible are actually located on the helmet. Rather they are mounted in a cockpit unit or main equipment bay electronic unit. A schematic diagram of the layout is shown in Fig 9. In most cases, the display module, containing the minimal electronic components, is clipped over the personalized helmet, thus allowing use by more than one pilot. The requirements of the display system have to integrate with the flying helmet in such a way that the fundamental properties of the flying helmet are not compromised. The aim is always to avoid an increase in weight whilst retaining helmet impact resistance. Therefore, equipment has to be positioned carefully to maintain the optimum helmet C of G and keep the helmet moment of inertia within acceptable limits. This is essential to avoid an increase in tiredness leading to loss of concentration and for safety in conditions encountered during ejection or during forced landings. Optical surfaces are either made of glass or optical plastics, the latter having the advantage of lower weight. The design must take into account the range of interpupillary distances and allow the eye to be positioned in the centre of the exit pupil with a correctly fitting helmet. The exit pupil is the optical 'window' through which the superimposed image is viewed. An exit pupil larger than 15 mm provides a very acceptable system in that if the helmet moves, the wearer does not suddenly lose the image. An increase in exit pupil necessitates an increase in weight so there has to be a sensible trade-off if comfort is to be maintained. Whilst helmet comfort is of paramount importance, in general the fitting requirements of HMDs assume more significance. The helmet fit, and therefore its stability, must be such as to maintain the eye(s) within the exit pupil(s). A visor (or visors) to attenuate glare and prevent eye damage from lasers is part of the helmet.

Tracking

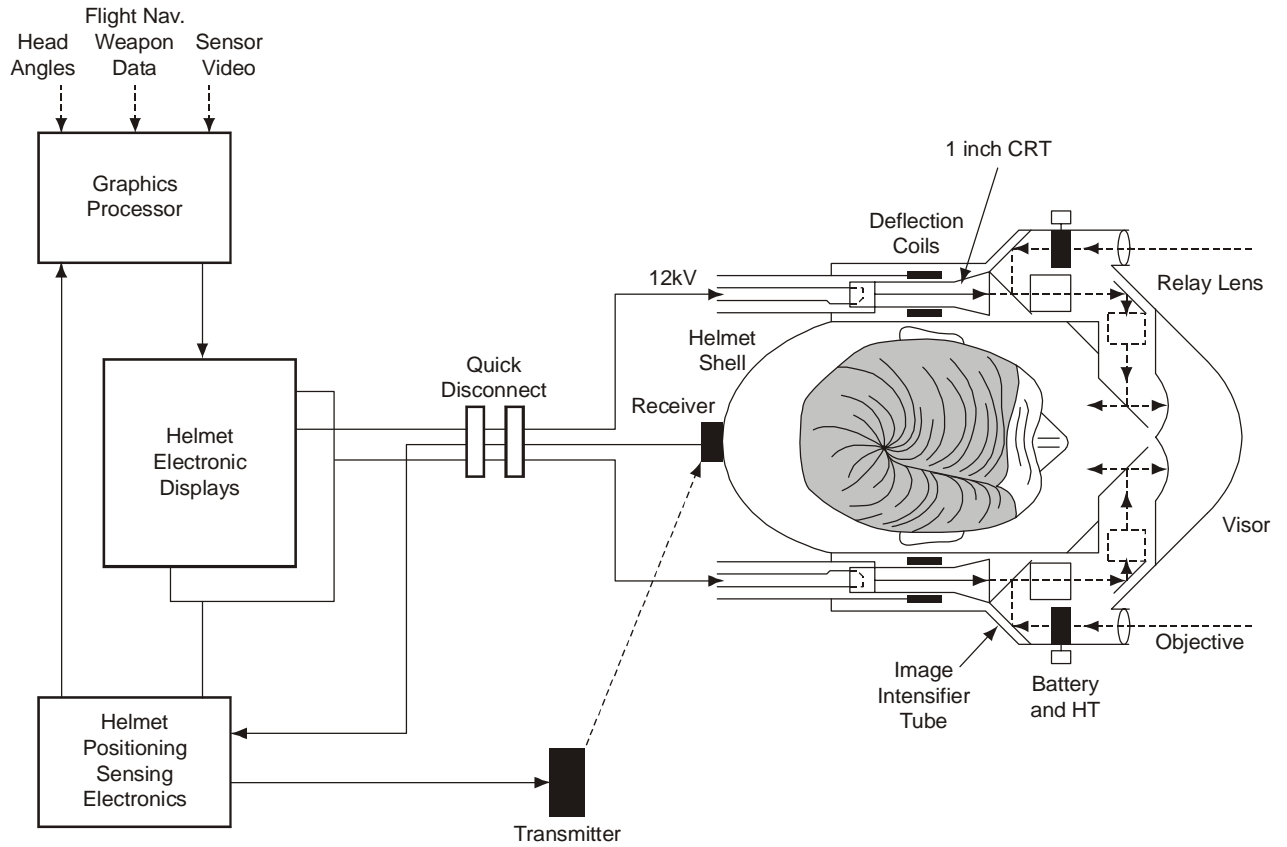
21. A HMD will not function without a helmet tracking system to determine the pilot's head position relative to the cockpit. Losses in the system which depend largely on processing power may result in the display lagging or jumping as the pilot moves his head. This is reduced by increasing the image refresh rate and introducing predictive algorithms. The potential of eye pointing has yet to be determined but it could provide a more natural method of designating objects. Furthermore, the natural stability of the eye could de-couple involuntary head motion (due to turbulence for example) from the aiming system.

7-20 Fig 9 Distribution of Components

Properties

22. To give binocular advantage and to cover for failure, two CRTs (for the thermal or other imagery) and two IIT units are usually fitted. When respective units are being employed separately, the single image is still viewed from two different sources. A binocular capability is preferred to retain depth perception although there are systems which project symbology to one eye only. Monocular systems are satisfactory for short-term tasks or during daytime. For enduring tasks, especially at night, such as en-route navigation, a binocular device overcomes binocular rivalry problems. Brightness and contrast are adjustable - or autocontrast can be selected to counter extremes of ambient light. A diagram of a typical IHS is shown in Fig 10. Overall, the HMDS requires the following properties:

- a. Parallelism of both IITs.
- b. No obscuration of IIT and CRT-based images.
- c. Low weight and correct CG for helmet.
- d. Parallelism and stability of combiners.
- e. Combiners preferably in one plane but must have high stability.
- f. Exact and easy adjustment of interpupillary distance if exit pupil is restricted.
- g. Large exit pupil for flexibility.
- h. Optimum adjustment of combiners should not change on switching between IIT and CRT channels.
- i. Field of view between 35° and 40°, although lower figures can be acceptable for specific tasks.
- j. Helmet tracker system with low image lag rates.

7-20 Fig 10 A Typical IHS

CHAPTER 21 – DIGITAL INSTRUMENT AND SENSOR DISPLAYS

Introduction

1. Conventional cockpit displays, involving banks of analogue instruments of all shapes and sizes, crammed into the available space on the instrument panel and adjacent consoles, have, for some time, given way to compact flat screen digital presentations in new aircraft. This follows the wide acceptance in the aviation industry of active matrix liquid crystal displays (AMLCD) (see Volume 7, Chapter 29 Para 12). Earlier attempts at applying TV screen technology employed bulky cathode ray tubes (CRT). These were found to be difficult to read in changing light patterns and too unreliable in that they were fragile and estimated to fail every 300 hours. Conversely, the AMLCD has a mean time between failures (MTBF) of some 15,000 to 20,000 hours, is extremely rugged, and offers true sunlight readability. The AMLCD also provides high resolution and a wide field of view. One disadvantage, fast being overcome, is the relatively high-power consumption of the AMLCD, the bulky power supply for which represents an undesirable weight consideration for small, tactical aircraft. Experiments with cholesteric techniques (no back lighting) have shown a considerable saving in power requirements and are cheaper than AMLCDs but they generally provide too little brightness. Flat display technology will continue to advance, and the results will tend to be incorporated into new aircraft types. However, whatever the method, the principle of the glass cockpit, described below, is only marginally affected.

THE 'GLASS' COCKPIT

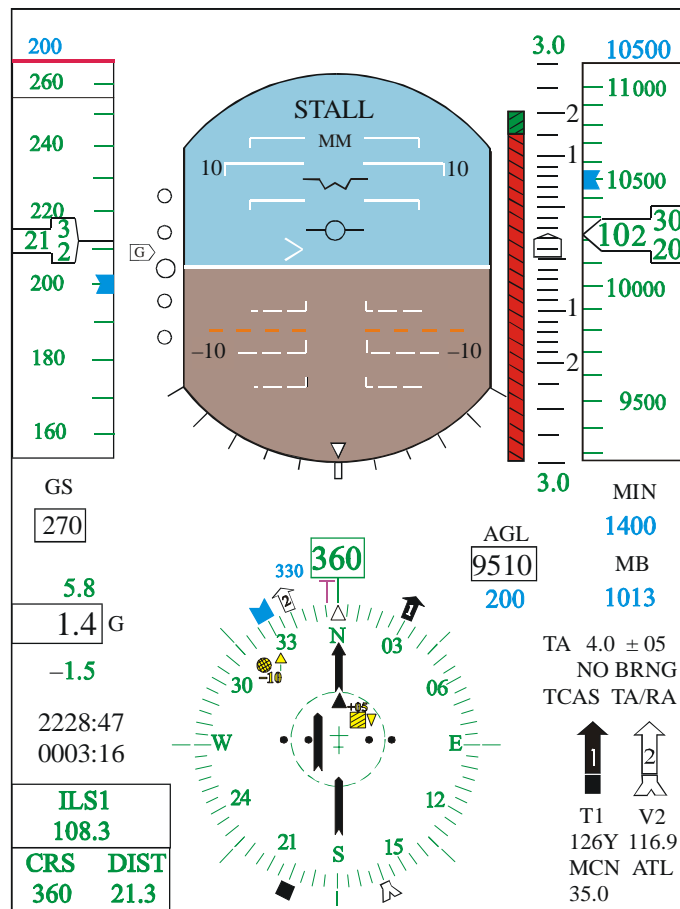
Principle

2. The term 'Glass Cockpit' describes the concept of having a number of multi-functional display (MFD) screens (each measuring some 25 by 20 cm) in front of the pilots rather than a conventional instrument panel. In fact, the panel becomes composite in nature in that it generally houses both the display screens and vital standby instruments and their associated switches and adjusters. The conventional instrument panel is, in the main, dominated by circular dials with pointers moving clockwise to increase and anti-clockwise to decrease the readings. In the glass cockpit, in order to make data recognition more intuitive, many such presentations have been superseded by the moving 'tape' symbology. In a tape display, the tape moves up or down and the pointer remains fixed. This, in development, led to many disagreements as to whether the tape should move up to increase the reading or down, the latter finally prevailing.

Typical Layout

3. Designers have striven with some success to standardize MFD presentations, not least in an effort to reduce aircraft type cross-training difficulties. Fig 1 shows a typical MFD in primary flight display mode which contains all critical flight instruments. Although the arrangement cannot be called standard, it represents the general consensus on where and what to display. Data is provided by the Air Data System (ADS) and other sensors and is controlled through an Avionics Management Unit (AMU). Invalid data is normally indicated by an X where the data should appear.

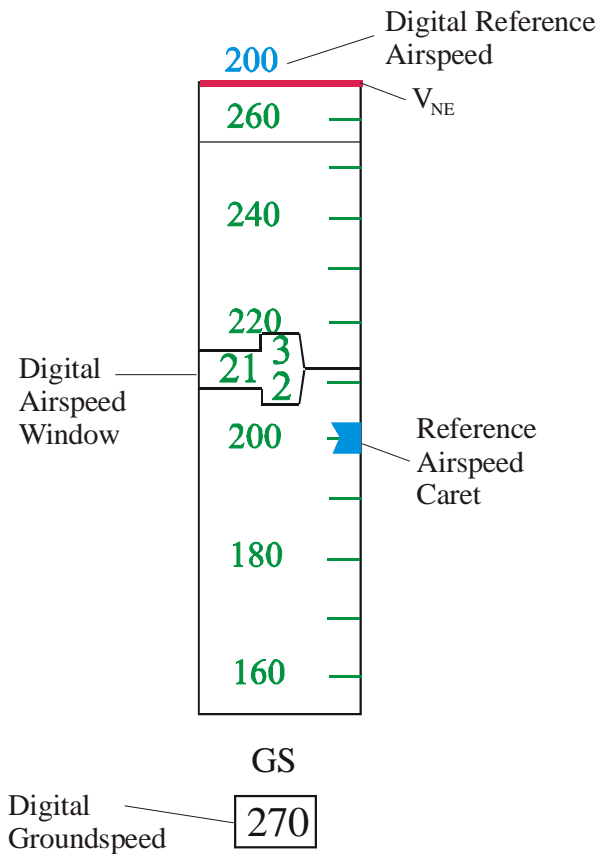
7-21 Fig 1 Typical Primary Flight Instrument Display



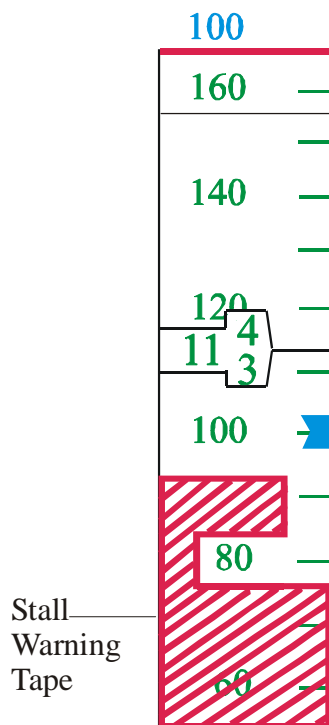
Airspeed Data

4. **Airspeed Indicator (ASI).** Calibrated Air Speed (CAS) or True Air Speed (TAS) may be selected for display and is represented by a tape as shown in Fig 2. As speed increases, the tape moves down and the (increasing) numbers are read against a fixed pointer. It can be interpreted from this intuitive presentation that the pointer is advancing, in the correct sense, upwards (or forwards). Similarly, reducing airspeed has the tape moving upwards with decreasing numbers passing the pointer. Thus, the pointer appears to move down or backwards as the speed falls. The pointer generally takes the form of an enlarged 'window' in which a more accurate reading of airspeed to the exact knot can be readily seen. On most systems, when CAS is selected, a reference CAS may be chosen and, is marked on the ASI by a coloured (usually cyan) caret or bug. The speed so selected also appears above the ASI tape in the same colour as the caret. If the reference speed disappears off the top or bottom of the display, half of the caret normally remains visible to the pilot. Once the reference speed reappears in the displayed range, the caret re-attaches itself and moves again with the tape. A red horizontal line marks the Never Exceed speed (V_{NE}). Groundspeed appears in a separate window just below the main display.

7-21 Fig 2 Airspeed 'Tape' Symbology



5. **Stall Warning.** Stalling speed is a function of weight, aircraft configuration, and engine power. A selectable stall-warning feature (see Fig 3) is usually configured to appear as a red and white diagonally striped area starting at a point just above the calculated stalling speed. The AMU will interrogate other systems, such as the Ground Collision Avoidance System and flap and undercarriage sensors in order, automatically, to configure the stall warning tape correctly.

7-21 Fig 3 Tape showing Stall Warning

6. **Take-off Speed Carets.** Carets may be displayed against the following critical take-off speeds:
- Decision Speed (V_1).
 - Rotation Speed (V_R).
 - Take-off Safety Speed (V_2).

The carets for V_1 and V_R are removed automatically once the main wheels leave the ground.

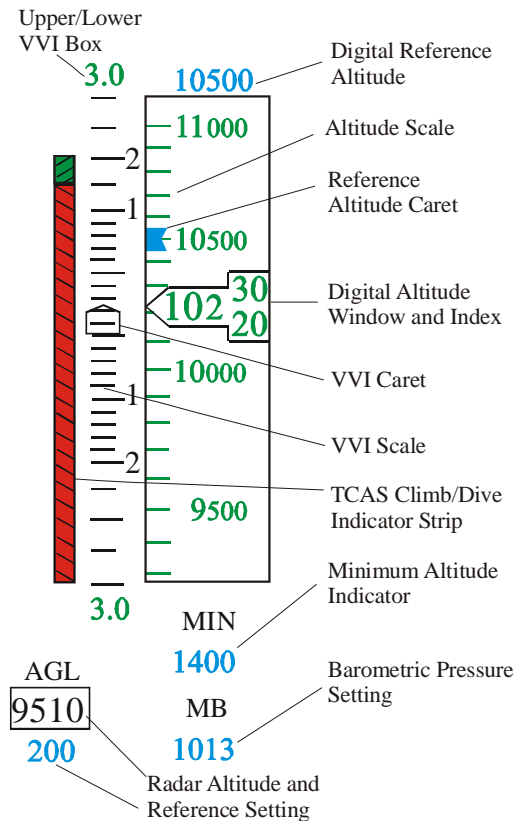
Altitude Data

7. **Altimeter.** The altimeter tape symbol is shown in Fig 4. It displays the barometric altitude of the aircraft derived from the ADS with invalid data denoted as Xs. Some, or all, of the following altimeter components may be selected for display, depending on aircraft type:

- Altitude scale.
- Digital altitude window and index.
- Reference altitude and caret.
- Minimum altitude indicator.
- Barometric pressure setting.
- Digital radar altitude.
- Reference radar altitude setting.

8. **Altitude Scale.** The altitude figures and scale tick marks appear in colour (green in Fig 4). The range is in the order of 1,000 to 50,000 ft with tick marks at 100 ft intervals. Increasing altitude is indicated by a downward movement of the tape which, like airspeed, may be interpreted intuitively as the pointer moving upwards for increasing, and downwards for decreasing, altitude.

7-21 Fig 4 Altimeter, VSI, and TCAS Displays



9. **Digital Altitude Window and Index.** A white pointer window depicts the exact barometric altitude in enlarged digits reading in 10 ft increments. Digits rotate downward to indicate increasing altitude and vice versa.

10. **Reference Altitude and Caret.** A reference altitude may be selected by positioning a coloured caret against the selected value. This altitude is then displayed digitally, and in the same colour as the caret, immediately above the altitude scale.

11. **Minimum Altitude Indicator and Barometric Pressure Setting.** The minimum altitude can be set and appears (shown as MIN and coloured cyan in Fig 4) above the barometric pressure setting and below the altitude scale. It also appears as a (cyan) line on the altitude tape. The barometric pressure setting can be selected to read in millibars (shown as MB) or inches of mercury (shown as IN HG).

Note: The hectopascal (hPa) is the standard unit of pressure although the millibar (mb) is still common in aviation. The hPa and the mb have equivalent values and so can be considered to be identical for all practical purposes.

12. **Digital Radar Altitude.** Digital radar altitude is presented as figures in a boxed read-out below and to the left of the barometric pressure setting and is labelled AGL.

13. **Reference Radar Altitude Setting.** A radar altitude reference setting, when selected, appears in cyan immediately below the digital radar altitude.

Vertical Velocity Data

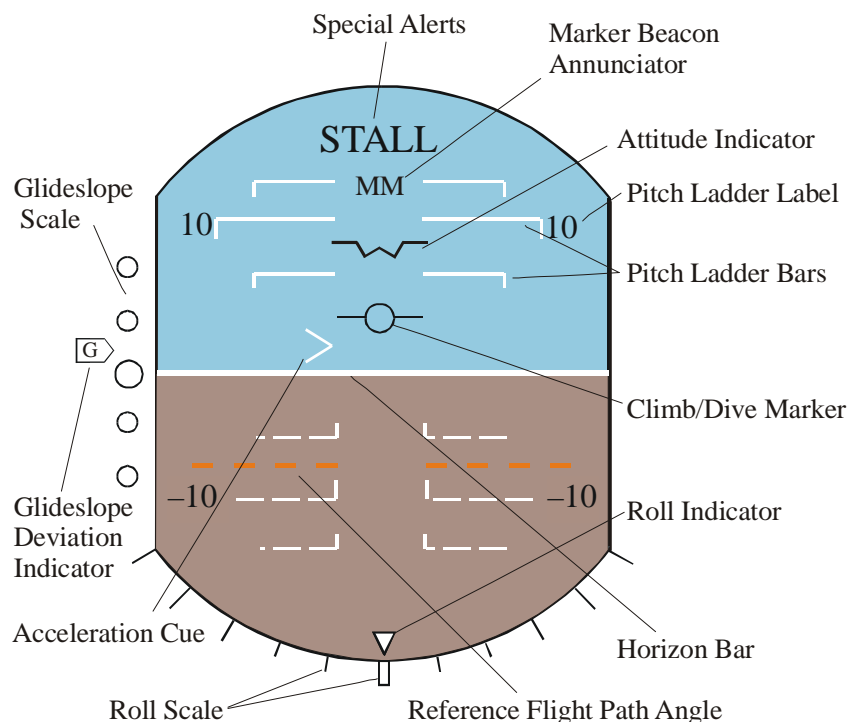
14. **Vertical Velocity Indicator (VVI).** The VVI (see also Fig 4) is a fixed scale located to the left of the altimeter. It displays rate of change of altitude using data from the ADS. The scale contains tick marks every 100 fpm between 1,000 and -1,000 fpm and every 500 fpm between 1,500 and 3,000 fpm and -1,500 and -3,000 fpm. The current velocity is indicated by a caret which expands from the centre of the fixed scale, upwards for positive vertical velocity and downwards for negative values. An upper and a lower VVI read-out are presented at the top and bottom of the scale respectively. This remains as a numeric 3.0 until the VVI tape caret becomes fully elongated at $\pm 3,000$ fpm or greater. The numeric 3.0 then becomes boxed in white, changes from green to white and displays the actual velocity, at 300 fpm intervals, up to a maximum of 9,900 fpm rate of climb or descent.

Flight Path Data

15. **Flight Path Director.** The Flight Path Director Indicator (FPDI) is positioned in the top centre of the MFD. Some of the selectable features are depicted in Fig 5. The FPDI is the primary instrument for displaying:

- a. The flight path - the actual trajectory of the aircraft through the air using data derived from the Inertial Navigation Unit (INU).
- b. Aircraft pitch and roll.
- c. Flight director information.

7-21 Fig 5 Flight Path Director Indicator



16. **FPDI Features.** The more important features available on the FPDI are:

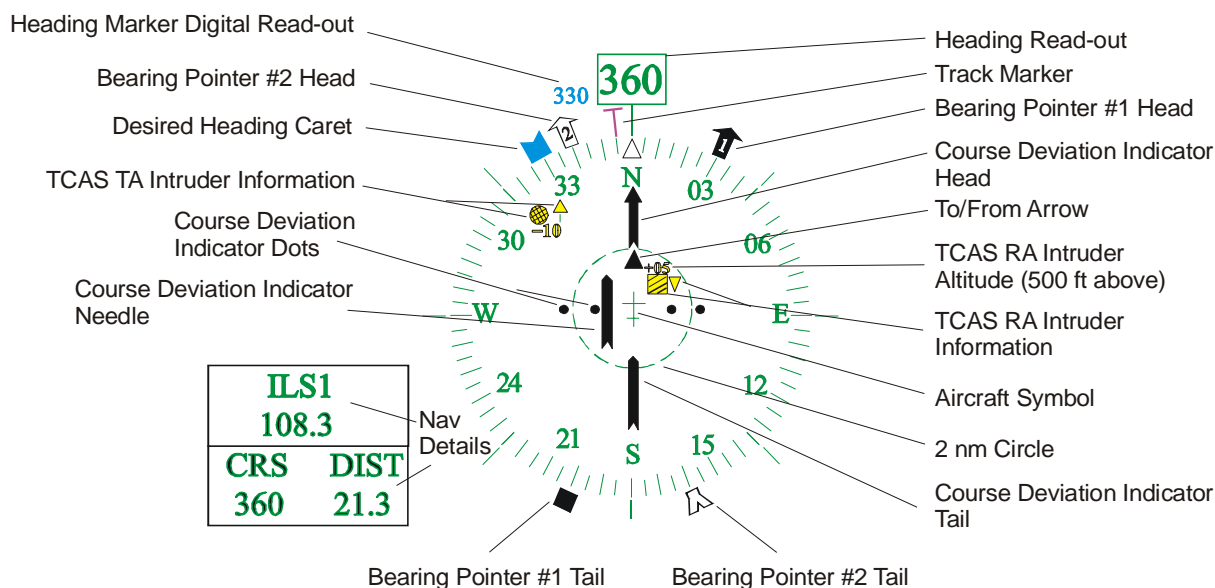
- a. **Climb/Dive Marker (CDM).** The CDM is represented by a circle with a horizontal line on either side except with the autopilot engaged when the circle changes to a diamond. The CDM indicates the vertical flight path angle through space.
- b. **Pitch Ladder Bars.** The pitch ladder bars indicate, relative to the CDM, whether the aircraft is in a climb or a dive. The bars are displayed at 5° increments, with the horizon representing the 0° vertical flight path.
- c. **Horizon Bar.** The flight path scale is divided by the horizon bar into two contrasting coloured areas to represent the sky and the ground. Nose up pitch is indicated by blue and nose down pitch by a brown background.
- d. **Roll Scale and Indicator.** The roll scale indicator shows the amount of aircraft roll. The scale is marked in 10° increments between 0° and 30° roll and 15° intervals between 30° and 60° roll. The roll pointer moves along the curved scale centred on the CDM. At angles above 60° roll, the pointer continues to traverse the vertical sides of the FPDI at the correct relative angle and is free to complete a circuit of the instrument if the aircraft continues to roll.
- e. **Aircraft Pitch/Attitude Indicator.** The aircraft pitch indicator is represented by a W with a horizontal line on each side, the W containing the angles of 30° and 60° for easy interpretation of bank. It represents the longitudinal axis of the aircraft and moves vertically. Pitch is indicated by the position of the pitch indicator relative to the horizon bar.
- f. **Flight Path Angle (FPA) Reference.** The FPA reference can be set from defined positive and negative limits, usually in 1° increments. The FPA reference is in the form of a split horizontal dashed line (shown in amber on Fig 5). Its position is in direct relationship to the value set by the pilot and is shown as a deviation from the horizon line.
- g. **Special Alerts.** Special alerts are centred at the top of the FPDI. Normally up to 20 characters can be displayed. Alerts may be sourced from any of the management or sensor systems.
- h. **Marker Beacon Annunciator.** Inner, middle and outer marker identification annunciations are displayed when the aircraft is in range of the appropriate transmitters.
- i. **Acceleration Cue.** The white acceleration cue appears alongside the left wing on the CDM. The cue is above the wing when the aircraft is accelerating and below the wing to indicate deceleration. It remains in line with the wing at constant speed.
- j. **Glideslope Scale and Deviation Indicator.** The glideslope scale gives raw glideslope data with the aircraft positioned in the centre of the scale. The pointer (a small arrow with a G inside) indicates the centre of the glideslope beam and is a 'fly to' indicator. The relationship between the scale and the pointer denotes the displacement of the aircraft from the centre of the beam. Thus, when the pointer is in the centre of the scale, the aircraft is on the glideslope. The glideslope indicator is only displayed when the navigation source is tuned to an ILS frequency; an invalid glideslope signal presents a red box with the letters GS inside at the bottom of the scale.

Horizontal Situation Data

17. **Horizontal Situation Indicator (HSI).** The HSI (see Fig 6) is a development of the EHSI described in Volume 5, Chapter 13, Para 3. It is positioned centrally at the bottom of the MFD and displays heading, track and navigation information derived from the INU and other appropriate sensors:

- a. **Heading and Track.** Track and heading information displays are superimposed on a green compass rose. Heading is always indicated by the lubber line at the top of the HSI with a boxed green digital heading read-out immediately above. The compass rose is marked every 5° with a tick mark and every 30° with numerals. Green cardinal initial letters appear at the 90°, 180°, 270°, and 360° positions. As the card rotates, the numerals and cardinal letters remain upright with respect to the top of the display. A magenta track marker in the shape of a 'T' rotates around the outside of the compass rose and points to the calculated track azimuth.
- b. **Heading Marker.** A 'Set Heading' knob enables a cyan heading bug or caret to be rotated and positioned on a desired heading. When set, the caret rotates with the compass card as the aircraft is turned on to the desired heading. A cyan digital read-out of the setting appears to the lower left of the current heading read-out box.

7-21 Fig 6 Horizontal Situation Indicator



- c. **Course Deviation Indicator (CDI).** The CDI is split. It has a head and tail portion and a central needle which shows displacement to the right or left of the selected course. The needle traverses 4 CDI dots, 2 on each side of the aircraft symbol, and which indicate the extent of the deviation. The degree of displacement indicated depends upon the navigation aid selected.
- d. **Bearing Pointers.** There are three bearing pointers which rotate around the compass card. The heads numbered 1 and 2 (Fig 4) each have a corresponding tail depicted 180° removed. These pointers indicate the bearings of any two navigational aids selected, the details of which are placed in a read-out box to the bottom left of the HSI. If no aids are selected, neither the pointers nor the read-outs are displayed. The third, unnumbered, bearing pointer is shown as a white triangle beneath the lubber line. It moves within the compass card and points to the CDI source.
- e. **Aircraft Symbol.** A symbol representing the aircraft is placed in the centre of the HSI.

ASSOCIATED FEATURES

Other Symbolology

18. **Timekeeping Functions.** Time and stopwatch read-outs are placed on the left of the HSI. The upper digits represent the clock which displays current time in hours, minutes, and seconds. The stopwatch format below the clock has an identical format.

19. **Traffic Alert and Collision Avoidance System (TCAS).** The TCAS gives derived information in three areas of the MFD.

- a. Vertical resolution guidance is presented to the left of the VVI. The display may be set to give 'avoid' or 'fly-to' instructions. A typical indication is shown in Fig 4 where a 'fly-to' climb command is displayed by a red band extending from the –3,000 fpm mark on the adjacent VVI scale to +1,500 fpm followed by a green band from +1,500 fpm to +2,000 fpm. The pilot should respond smoothly to achieve the rate of climb which matches the green sector. When the threat ceases, the display is removed, and the original flight path may then be resumed.
- b. Text messages and modes are displayed in the space to the right of the HSI (see Fig 1).
- c. Traffic advisories (TA) or resolution advisories (RA) may be selected to appear on the HSI plan view (see Fig 6). Symbols vary between systems and are shown in the appropriate Aircrew Manuals. The details are positioned to show the bearing and distance of the intruder. The outer ring of the compass card represents a range of 6 nm and an inner, dashed circle denotes a range of 2 nm. Intruder altitude relative to own aircraft is shown in hundreds of feet with a positive (+) or negative (–) sign preceding the figure. Positive data tags are shown above, and negative tags below, the threat symbol. The display can be switched to show the absolute flight level (FL) of the threat in hundreds of feet above MSL. Intruder FL is always denoted by three digits but is still positioned above or below the threat symbol in relation to own aircraft FL.

20 **Safety Features.** An advantage in the use of an MFD is the feasibility of providing on-board surveillance of flying surfaces, undercarriage position or icing build-up, via miniature digital cameras. Many other services may be monitored electronically or by camera in this way but all remain selectable so as to relieve the MFD of unnecessary clutter. A g-meter readout can be selected to display in a white box to the left of the HSI. The maximum positive and negative g units recorded are displayed in green respectively above and below the g-meter box.

Head-up Display Functions

21. The Head-up Display (HUD) (see Volume 7, Chapter 20) repeats attitude and flight path symbology, essential basic flight symbology and navigation and approach data as selected on the MFDs. In addition, battle symbology and safety features can also be selected thus allowing a head-up view of practically all flight and battle parameters.

Electronic Warfare Symbolology

22. Instead of having a standalone electronic warfare scope, the targeting, weapons and engagement system displays can be put on any of the MFDs or the HUD to give an integrated view of the emitter battlefield. Emitter engagement rings can be included to show lethality zones and can be adjusted to compensate for own aircraft altitude and terrain masking considerations. This reduces the cockpit workload and makes the pilot's task of circumventing, suppressing, or destroying the threat much easier to accomplish.

Helmet-mounted Display

23. Flight reference data may be displayed on a helmet-mounted visor (see Volume 7, Chapter 20 Para 18 for details). This can be configured to permit target acquisition, weapon aiming and engagement at large off-boresight angles. The helmet may also incorporate night vision aids, using light intensification features, and FLIR imagery.

CHAPTER 22 - AUTOPILOT AND FLIGHT DIRECTOR SYSTEMS

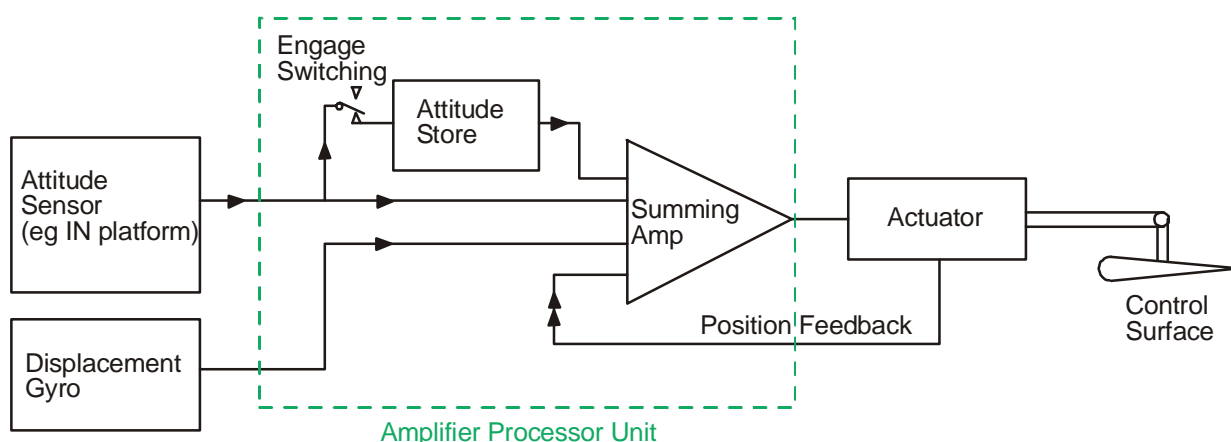
Introduction

1. Automatic flight control systems (AFCS) are discussed in Volume 4, Chapter 7. This chapter will examine the practical applications of autopilots to show how the AFCS can be used to alleviate pilot work load and to carry out tasks which, without autopilot assistance, would impose a considerable work load upon the pilot and in some cases would be impossible.

Autopilot Control of Aircraft Attitude

2. The control of aircraft attitude is essential to the manoeuvring of the aircraft by autopilot. Long term attitude monitoring is usually provided by a displacement gyro system (see Fig 1). Aircraft attitude information is passed to a memory unit in the amplifier/processor where it is stored. When attitude hold is selected, the input to the memory is disconnected so that the recorded attitude becomes a fixed datum. The attitude store output is compared with the direct attitude signal in a summing amplifier - the signal inputs are equal at the moment of engagement and the output of the amplifier is zero. If the aircraft deviates from its set attitude, the two signals are no longer equal and the summing amplifier produces an error signal. The error signal is passed to a demand actuator which moves the appropriate control to return the aircraft to its original attitude. A position feedback loop ensures that the control applied is proportional to the demand signal. A rate feedback loop controls the rate at which the aircraft responds to the demand signal, thus preventing over controlling and the possible overstressing of the aircraft. A three-axis autopilot has loops for pitch, roll, and yaw rate.

7-21 Fig 1 Attitude Hold Loop



Automatic Throttle Control

3. Complete automatic control of an aircraft requires an automatic throttle control system so that speed can be controlled during changes of altitude or whilst manoeuvring. The automatic throttle control system monitors airspeed and pitch rate against datum parameters set by the pilot or as a product of auto ILS, TFR, or weapons aiming and attack systems. The system can also control engine power to achieve ideal range or endurance speeds.

Autopilot Sensors

4. The attitude of an aircraft may be defined by its position in pitch, roll, and yaw. Datum information for roll and pitch can be provided by vertical gyros and yaw rate information can be provided by lateral

accelerometers. A heading reference may be a gyro-magnetic compass or an INS. Using these sensors, the autopilot is able to fly the aircraft straight and level on a constant heading.

Manoeuvring the Aircraft

5. Attitude demands may be pilot or autopilot initiated. Pitch, roll, and yaw demand signals are passed directly into the computer/amplifier/servo system. The autopilot responds by operating the appropriate controls to reduce the error signal as described in para 2.

- a. **Manual Control Facilities.** If the pilot wishes to enter attitude demands manually, he can do so by using switches or potentiometers to produce electrical signals which are fed directly to the autopilot as pitch, roll and yaw demands. The controls for entering demands manually may be on a control panel or on the control column of a fast-jet aircraft.
- b. **Automatic Control Facilities.** The outputs of various aircraft systems can be fed into the autopilot manoeuvring facility by selection. Typically, signals may be derived from:
 - (1) **Flight Instrument Systems.** The pilot may set a heading or track demand by moving an index marker on the horizontal situation indicator.
 - (2) **Radio Navigation Aids.** Inbound or outbound radials can be derived to steer the aircraft towards or away from VOR, TACAN, or ILS localizers.
 - (3) **Air Data Systems.** Datum signals can be produced to fly the aircraft at constant barometric height, airspeed, or Mach number.
 - (4) **Terrain Following Radars and Radio Altimeters.** Signals can be derived from terrain following radar or radio altimeters to fly the aircraft automatically at selected heights above the ground.
 - (5) **Navigation Computers.** Signals can be derived to steer the aircraft towards a navigation feature or turning point.
 - (6) **Weapon Aiming/Attack/Search Systems.** Signals from weapon aiming, attack or search systems can be used to fly the aircraft in predetermined search and attack patterns.

The ability to use these systems enables the pilot to select the appropriate inputs to the autopilot for a very wide range of flying activities from a relatively undemanding navigation task to very demanding low level navigation and attack mission, possibly at night or in bad weather.

Autopilot Safety

6. An autopilot must not be capable of endangering the aircraft or its crew. Autopilot safety is ensured by a variety of design features and devices to ensure at least a 'fail-safe' capability. Features and devices vary greatly but typical examples are:

- a. **Design Features.** Circuits are designed to be as simple as possible and components are used at a fraction of their rated values to ensure high reliability. Additionally, switching circuits are given clearly defined priorities to avoid inadvertent selection of dangerous flight configurations and to avoid selection of incompatible flight control modes.
- b. **Safety Devices.** The following safety devices are typical:
 - (1) **Pilot's Instinctive Cut-out.** The instinctive cut-out is positioned on the control column and can be quickly and easily operated to disengage the autopilot giving full manual control authority to the pilot.

(2) **Rate and Angle Limiters.** The rate and angle limiters prevent the overstressing of the aircraft by limiting the rate of response or angle achievable in any channel.

(3) **Control Limit Switches.** Control limit switches are microswitches which operate when a control reaches the end of its allowable travel. These switches are able to prevent any damage from servo runaway.

(4) **Excess Torque Devices.** Excess torque devices are used either to prevent overstressing of the aircraft or to detect excessive current demands such as might occur if an electrical servo was attempting to overcome a control restriction.

(5) **Monitoring Facilities.** Most autopilot functions are continuously monitored by a built-in test equipment system which is able to generate warnings and initiate automatic reversionary modes. Commonly monitored functions are power supplies, the accuracy of datum information on attitude and heading and the serviceability states of systems which provide inputs to the autopilot.

Flight Information System

7. Autopilots include a flight information system which provides aircrew with an integrated presentation of:

- a. Primary flight information showing attitude and heading.
- b. Flight director information showing indices and markers which indicate the horizontal and vertical control required to regain a demanded flight path.

Flight information systems range from simple displays to fully processed electronic head-up or head-down displays. The system enables the pilot to fly the aircraft manually to meet the autopilot demands, or to check that the autopilot is following the demands correctly.

CHAPTER 23 - UHF HOMING

Introduction

1. For the purpose of this description, only azimuth Direction Finding (D/F) will be considered, although, the principles are also relevant to elevation D/F. UHF Homing is used in conjunction with the UHF Transmitter/Receiver (T/R) to provide relative azimuth indications from sources of Continuous Wave (CW), Modulated Continuous Wave (MCW) or Radio/Telephony (R/T) transmissions on any selected frequency in the 225 MHz - 400 MHz band.

Leading Particulars

2. The equipment is used during SAR operations to enable an aircraft to home towards a Personal Locator Beacon (PLB). The effective range is 100 nm maximum, decreasing with decreasing altitude.
3. The relative azimuth of a signal is shown by the deflection of the vertical pointer of an indicator.

Power Requirements

4. The power supplies for the equipment are obtained from the UHF T/R and from the aircraft 28V DC supply.

Equipment Components

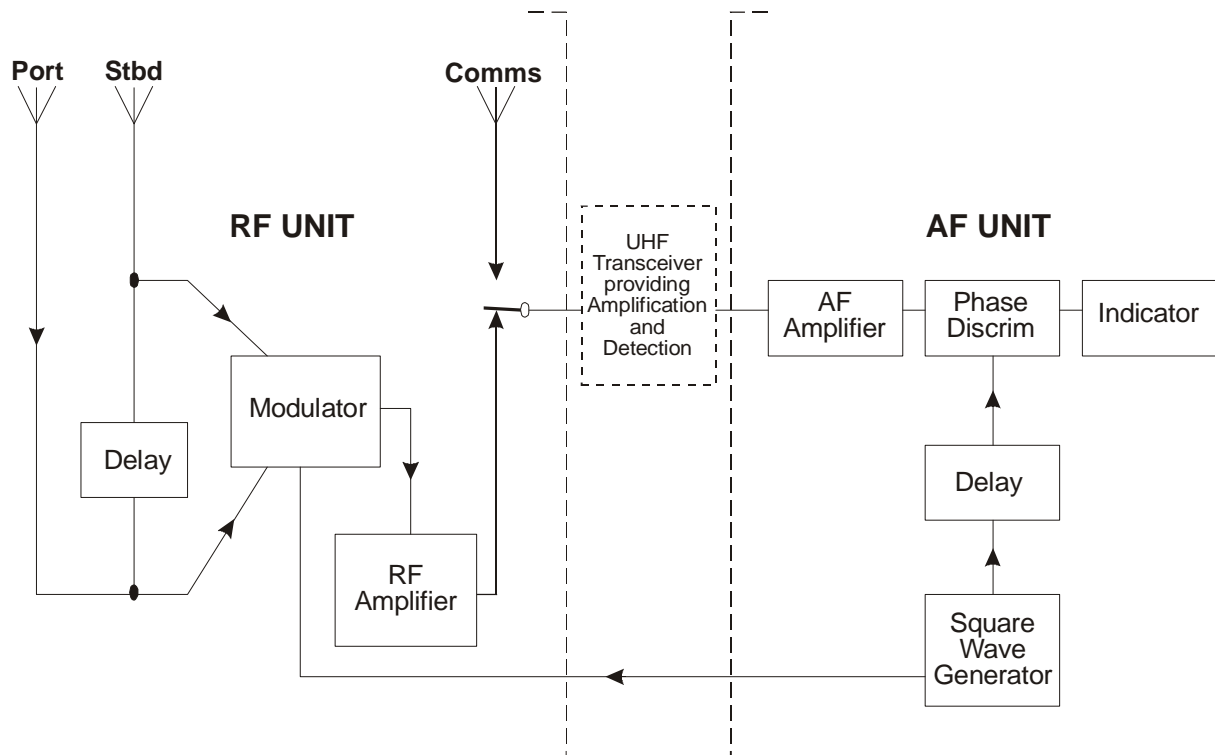
5. The units comprising a typical UHF homing system are:
 - a. Radio Frequency (RF) unit.
 - b. Audio Frequency (AF) unit.
 - c. D/F aerials.
 - d. Indicator.
 - e. Control switches.
6. **RF Unit.** The RF unit converts the phase difference of the two received signals into an amplitude modulation of the carrier wave.
7. **AF Unit.** After amplification and de-modulation in the UHF T/R the AF unit converts the signals into a form suitable to operate the indicator's pointer.
8. **D/F Aerials.** Port and starboard aerials mounted close together on the upper surface of the fuselage are used for azimuth D/F.
9. **Indicator.** Either the Horizontal Situation Indicator (HSI) or Radio Magnetic Indicator (RMI) will provide azimuth indications during homing operations. The 'OFF' flag appears when homing signals become unreliable.

Principle of Operation

10. Depending on the relative bearing of a transmitting PLB, there will be a small time lag - and hence a phase difference - between the signals received in the port and starboard aerials.
11. The signals are processed electronically to produce a DC voltage to the indicator which is a function of the displacement of the signal source from the fore and aft axis of the aircraft. The polarity of the DC voltage depends on the phase relationship of the incoming signals, and determine which way the pointer deflects.

12. If the two signals are received at the same time then the transmitting source is directly ahead of or directly behind the aircraft. Fig 1 shows a simplified block diagram of a typical UHF homing system.

7-23 Fig 1 Block Diagram of a Typical UHF Homing System



Operating Procedures

13. Selection of the homing function on the UHF T/R control unit brings the UHF homing into operation.

14. It should be noted that signals received from astern the aircraft displace the HSI lateral deviation bar in the same sense as those ahead.

15. Operation of the transmit button causes the D/F aerial relay to be de-energized and the previously selected UHF aerial to be reconnected.

CHAPTER 24 - CENTRAL COMMUNICATIONS SYSTEMS

Introduction

1. In multi-crew aircraft, crew members are able to communicate with each other by means of an intercommunication (intercom) facility. Where this facility also enables the crew to select the required aircraft radio and communication equipment, it forms a 'communications control system' (CCS). By use of a CCS, most crew positions are able to select radio transmitter/receiver facilities, whilst other crew positions may have 'receive only' services. In addition to the main intercom facility, independent subsidiary and conference intercom circuits may be available.

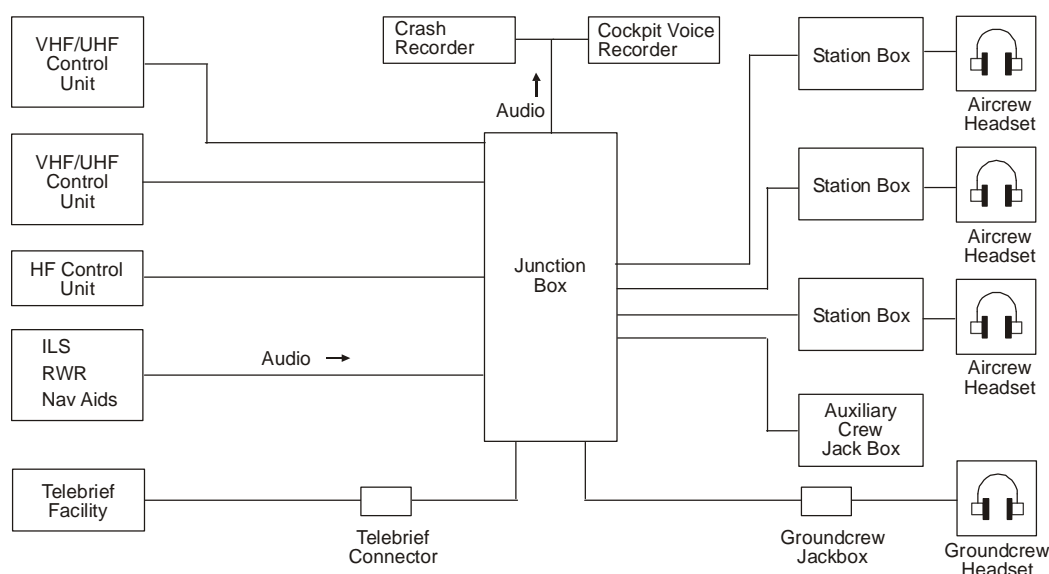
General Description of System

2. Since the detailed requirements of individual aircraft may vary considerably, there is no standard CCS installation. However, the control system normally consists of a number of units, each with different functions, which can be interconnected in various ways to provide the facilities required. The underlying principles are the same, irrespective of how the system has been connected.

3. Fig 1 shows the layout of a typical intercom installation. This example provides intercom and radio facilities for the aircrew, and external ground inputs from groundcrew and telebriefing facility. The major components are:

- a. **Junction Box.** All the services are connected through the junction box. The various flight crew, cabin crew and ground crew circuits radiate from it to the different aircraft positions. The system is powered by a 28V DC supply.
- b. **Intercom Station Box.** An intercom station box is provided at each of the principal crew positions. This unit permits individual selection of transmitters and receivers as required. The crew member's headset is connected to the station box.

7-24 Fig 1 Intercom System - Simplified Block Schematic

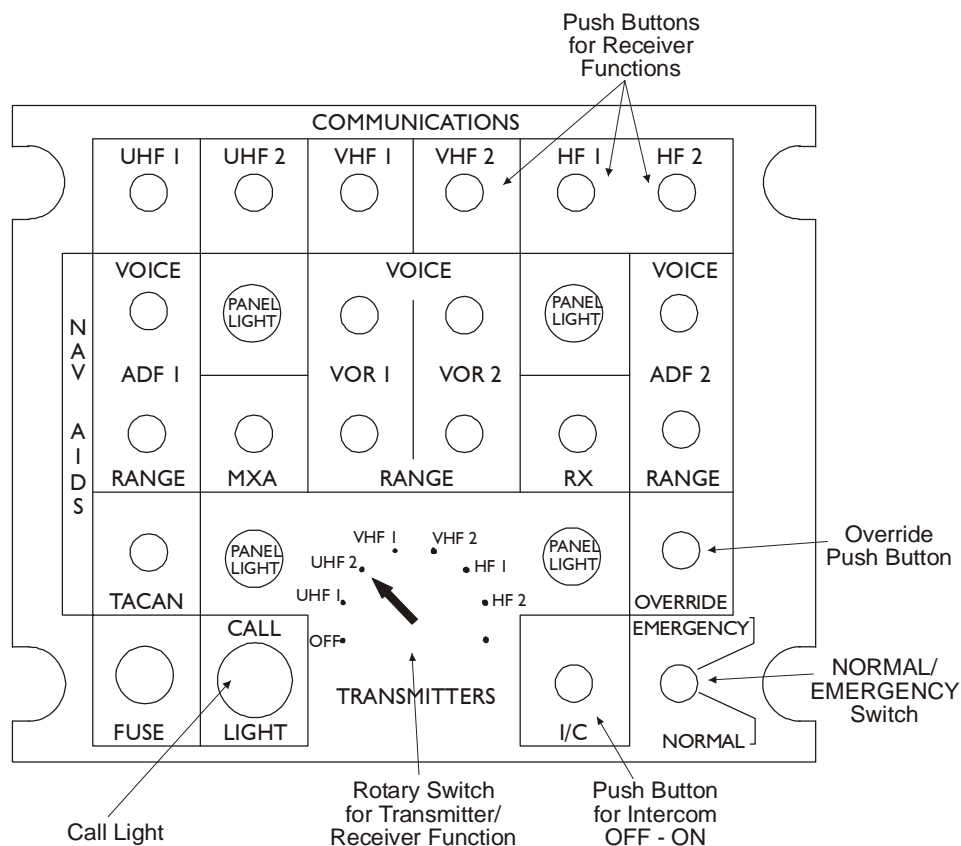


- c. **Radio Transmitter/Receiver Control Units.** The intercom system provides the connection from the crew's headsets to the VHF, UHF and HF radio transmitter/receiver control units (see Volume 7, Chapter 1 and Chapter 7 Chapter 3).

- d. **Auxiliary Jack Boxes.** Cabin/ground crew can use the normal intercom by plugging a headset into a convenient jack box. This action closes a switch, which provides a path for the 28V DC supply to the relays in the crew amplifier.
 - e. **Telebrief Facility.** The telebrief facility gives a secure briefing capability whilst the aircraft is on the ground.
4. To increase reliability, the systems are decentralized with each intercom station containing its own transistor amplifier and emergency operation facility. To reduce the risk of failure due to damaged components, the intercom stations are usually wired in a continuous loop or ring main circuit between the main distribution boxes. By this means, single unit or cable failure will not affect the operation of other stations.

Intercom Station Boxes

5. The intercom station box provides crew members with a convenient means of selecting and controlling the various services available (see Fig 2). The main facilities provided by the station box include:
- a. **Crew Intercom.** The intercom facility is selected by means of the Push-ON, Push-OFF button, which remains latched (partially depressed) when ON. The button incorporates a rotary volume control.
 - b. **Receiver Push Buttons.** Audio from radio facilities can be selected by the appropriate pushbutton. Any combination of facilities can be listened to by each crew member, simply by pressing the required button ON and rotating it to the desired volume level.
 - c. **Radio Transmitter Selection.** A multi-position rotary switch allows the crew member to select a specific transmitter. The position of the switch determines which service the operator can transmit and receive on. The corresponding receiver button does not need to be depressed, but is still used to adjust volume.
 - d. **Override Facility.** The OVERRIDE pushbutton or switch permits high-priority messages to be fed at high volume to all other intercom stations in the system, irrespective of the services which they have selected.
 - e. **Call Light Facility.** When any call light is pressed, all call lights on all station boxes illuminate. This serves to attract the attention of all crew members, even those not currently listening to the intercom.
 - f. **Voice/Range Buttons.** Voice/range filters are used with the aircraft automatic direction finding (ADF), VHF omni-range (VOR) and Instrument Landing System (ILS) installations to separate the range and voice elements of the received signals. The installation provides a choice of voice, range, or voice and range audio signal inputs.

7-24 Fig 2 Typical Intercom Station Box

6. The unit usually contains a two-section amplifier. This amplifies incoming radio and intercom signals before feeding them to the headset. It also provides the appropriate level of output from the crew member's microphone, before feeding it to the selected transmitter.

7. The NORMAL/EMERGENCY switch can be set to restore intercom and receiver services in the event of a failure in the two-section amplifier. The switch works in two stages:

- a. The intercom microphone signals are switched to the transmitter selector switch.
- b. The receiver signals, once selected by the appropriate pushbutton, are switched direct to the headset, bypassing all volume controls.

Intercom may then be restored by utilizing the sidetone output of a convenient transmitter. All other crew members requiring intercom with the faulty station must set their own selector switches to the same transmitter.

Intercom Discipline

8. To maintain smooth and efficient communications, an intercom discipline is essential. The following basic rules normally apply:

- a. The operator's headset microphone should be switched off unless actually speaking, as background noise will cause interference. (Some aircrew headsets have microphones that are speech-activated.)
- b. The operator should monitor the radio in use before, during and after speaking, to avoid interrupting radio transmissions to and from the aircraft.

- c. In a large multi-crew environment, crew procedures should be followed. Usually, the operator will nominate to whom they wish to speak, state who is speaking, and ensure that the recipient acknowledges before passing the message.
- d. Crew members should not switch off audio, or leave intercom, without first informing other crew members.

CHAPTER 25 - RADIO AIRBORNE TELETYPE (RATT)

Introduction

1. In some circumstances it is preferable to communicate between aircraft, and between aircraft and surface stations, using a teleprinter system rather than voice. The system in use in the RAF is known as the Radio Airborne Teletype (RATT), and both the equipment, and the operating techniques and procedures, are compatible with other national systems within NATO. The UK Glossary of Joint and Multinational Terms and Definitions, JDP 0-01.1, defines RATT as Radio Teletype. This definition is modified to include 'Airborne' within the RAF, while the Royal Navy and Army use the term Radio Automated Teletype.

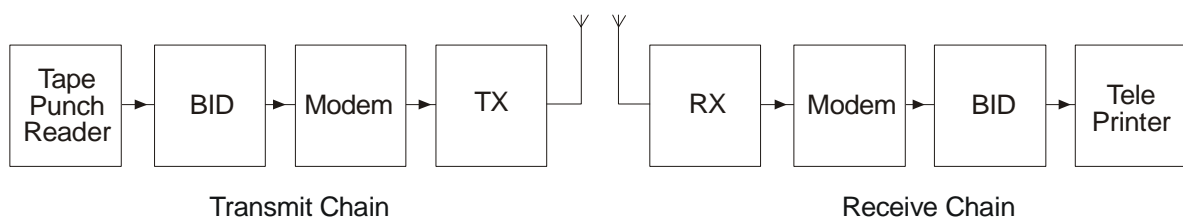
2. The equipment incorporates an automatic on-line encryption/decryption facility and is thus capable of combining a high level of communications security and accuracy. The use of RATT reduces the transmission time of messages compared to other methods, reduces errors in coding and decoding, and since hard copies of messages are available to both operators, subsequent enquiries and corrections are easily resolved. Fully formatted messages can be sent between, and are readily understood by, all NATO members which helps to avoid ambiguity and overcome any language problems.

Equipment and Operation

3. A RATT system uses a basic radio modified for telegraphy and incorporating a crypro unit (BID).

4. A simplified diagram of a RATT system is shown in Fig 1. Switching on the MODEM (modulator/demodulator) initiates a tone transmission containing no data. A Phase Indicator and Message Indicator (PIMI) superimposes a signal on this tone which enables remote receivers to lock on to the phase of the transmitter, using a Pseudo Random Sequence (PRS) technique, and so synchronize their equipment ready to receive the message.

7-25 Fig 1 RATT System – Schematic



5. Prepared messages are released via the Tape Punch Reader (TPR), automatically encrypted by the BID, converted into Murray code, and then sent by modulating the transmission with tones corresponding to the marks and spaces.

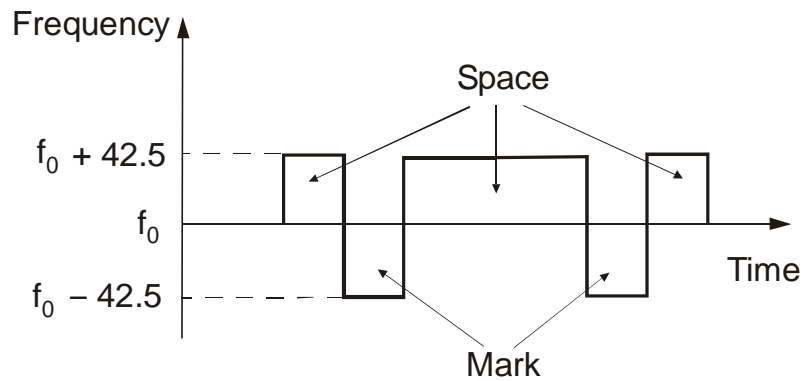
6. The RATT can operate in the HF, UHF, and LF bands:

- a. **HF.** In HF the operation is simplex, i.e. transmission and reception is on the same frequency and so cannot be accomplished simultaneously. The transmission uses single sideband (upper) suppressed carrier techniques, with the mark being represented by a tone of 1575 Hz and the space by 2425 Hz.

b. **UHF.** UHF operation is also simplex but uses double sideband transmission. The mark is represented by a 700 Hz tone and the space by a 500 Hz tone.

c. **LF.** LF is a receive only mode using frequency shift telegraphy. There is an 85 Hz shift between mark and space with the mark low, i.e. $f_0 - 42.5$ Hz, as shown in Fig 2.

7-25 Fig 2 LF Frequency Shift Telegraphy



7. Switching the transmitter to standby at the MODEM returns the equipment to the receive state, which is automatically activated by incoming PIMIs. The receive chain is almost the reverse of the transmit chain with any correctly phased traffic being printed out in clear at the teleprinter; any other traffic will be indecipherable.

CHAPTER 26 - VOICE RECORDERS

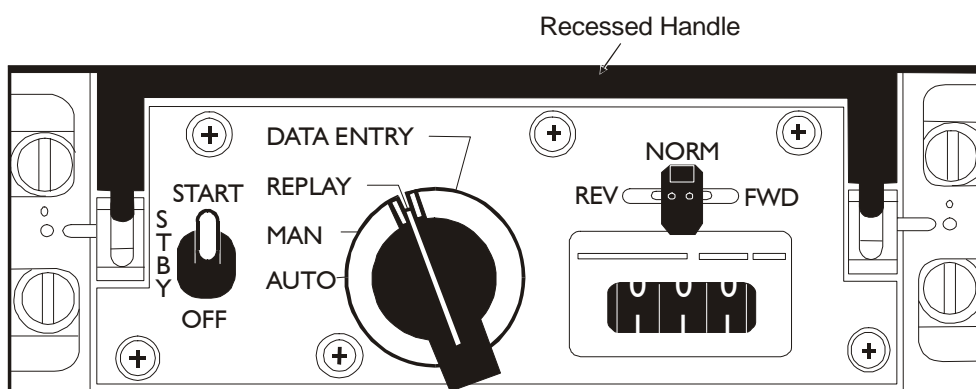
Introduction

1. A typical cockpit voice recorder (CVR) operates as an audio flight log for crew members and also provides a rapid data entry (RDE) facility for the insertion of data into the main computer (MC). The CVR works on the same principle as a domestic tape recorder. The following description is for the Tornado installation.

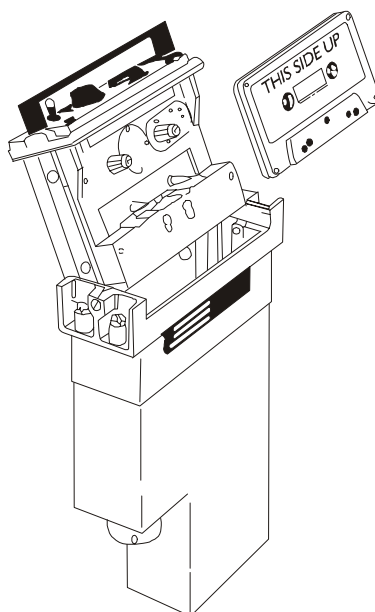
Control Panel

2. The control panel, shown in Fig 1 controls the operating mode, power supply and position of the tape. The cassette housing is located behind the control panel and is accessed by pulling the panel up, using the recessed handle (see Fig 2). This permits the cassette to be inserted or changed.

7-26 Fig 1 CVR Control Panel



7-26 Fig 2 Cassette Location



Operation

3. **Rapid Data Entry (RDE).** Before take-off, pre-recorded information can be entered into the MC via the CVR, ie operational flight program (OFP), map definition data or mission data. The information is

used by other avionics systems during the flight and, if necessary, further data can be entered as the mission progresses.

4. To replace or insert a cassette, the operator should select **REPLAY** and then **OFF** on the CVR control panel, before opening the unit. Once the tape is loaded, standby (**STBY**) should be selected and using the reverse/normal/forward (**REV/NORM/FWD**) control, the tape wound to the required position, watching the tape position indicator. Prior to selecting **START** and **DATA ENTRY** on the CVR control panel, the **MC** and **TV Tabular Displays (TV Tabs)** need to be switched to **ON**.

a. **Operational Flight Programme.** When loading the **OFP**, the **LOAD** light on the main computer control panel (**MCCP**) will illuminate. Should a failure occur, the **FAIL** light on the **MCCP** will illuminate and the **LOAD** light will extinguish.

b. **Map Definition/Mission Data.** While loading map definition data or mission data, one of the following messages will be displayed on the **TV Tabs**:

- (1) **RDE COMPLETE**, which indicates that the data has been loaded correctly.
- (2) **RDE FAULT**, which indicates one, two or three faults in the data loaded.
- (3) **RDE FAILED**, which indicates there are four or more faults in the data loaded.

5. **Automatic Voice Recording.** The CVR records automatically with **START** and **AUTO** selected on the control panel and when audio signals exist. Recording ceases approximately five seconds after the audio input ceases. To facilitate separate recordings for each crew member, the tape has two tracks, each track designated to a crew position.

6. **Manual Voice Recording.** With **START** and manual (**MAN**) selected on the control panel, the CVR will record continuously (the length of time depending on the length of the cassette tape). The CVR rotary/push switch on the intercom box must also be selected.

7. **Replay.** With **START** and **REPLAY** selected on the CVR control panel and the CVR control on the intercom box adjusted to give the required volume, previously recorded information can be replayed. A separate switch on the intercom box, track 1/off/track 2 (**TRK1/OFF/TRK2**), is used to enable the crew member to select which track they replay.

8. **Faults.** If a failure occurs in data loading, or if a recording sounds garbled, the CVR heads should be cleaned using the prescribed head cleaner.

CHAPTER 27 - AIRBORNE COMPUTERS

Introduction

1. Since 1975, computer technology has made rapid advances in the fields of speed, memory capacity, and reliability, while at the same time there have been reductions in physical size, power consumption, and cost. This chapter, however, deals specifically with computers in the airborne environment and reviews the tasks undertaken and the types of computer available. It also examines the computer architectures best suited to meet the requirements, and briefly discusses the various peripheral devices which are commonly used for the input and output of data. Since computers are often required to control and integrate data obtained from a variety of disparate sources, the manner in which data is transmitted between equipments will also be discussed.

Airborne Computer Tasks

2. Airborne computing tasks can be broadly divided into three main groups:

- a. Navigation and weapon aiming.
- b. Control and management.
- c. Data processing.

3. **Navigation and Weapon Aiming.** Examples of the navigation and weapon aiming problems which are normally solved by computer are:

- a. Control of inertial, doppler and satellite navigation systems, including the use of Kalman filtering techniques to provide mixed solutions.
- b. Co-ordinate conversion (eg Lat/Long to Grid, geoid to geoid).
- c. Air data processing.
- d. Weapon aiming calculations, including:
 - (1) Ballistics
 - (2) Continuous Computed Impact Points.
 - (3) Offset and range calculations.
 - (4) Target recognition and tracking.

4. **Control and Management.** The control and management functions carried out by computers include:

- a. Flight control systems, eg fly-by-wire.
- b. Fuel and engine monitoring and control.
- c. In-flight recording.
- d. Equipment self-test routines.
- e. Data transmission control and management.
- f. Electronic warfare (EW) equipment management.
- g. Health and Usage Monitoring Systems (HUMS), eg fatigue life monitoring.

5. **Data Processing.** Examples of data processing applications are:

- a. Image processing (eg IR and television).
- b. Radar data processing.
- c. EW data processing.
- d. Digital land mass data manipulation.
- e. Data collation tasks.

6. **Data Updates.** Many of the functions carried out by the computer could be completed automatically with no human intervention. However, it is normally desirable that the crew should be able to maintain some measure of control over the computer and contribute to the decision-making process. In addition, mission-specific data (such as flight plans and target information) can be prepared before flight and saved onto the appropriate magnetic or optical medium, to be loaded into the aircraft computer during start-up.

Airborne Computer Types

7. There are three types of computer currently being used in airborne applications:

- a. Analogue.
- b. Hybrid (mixed analogue and digital).
- c. Digital - General Purpose and Special Purpose.

A further type, the optical computer, is still at the development stage. The characteristics of analogue and hybrid computers will be summarized, but the remainder of this chapter will be concerned with the digital computer.

8. **Analogue Computers.** Analogue computers accept and process data as continually varying quantities, represented by physical parameters, eg voltage or shaft angle. In the early days of digital computers, analogue machines had the advantages of avoiding the sampling errors associated with digital techniques, and of being inherently 'real time'. However the analogue computer is inflexible in its applications, and does not have the ability to store large quantities of data. The development of digital computers has been such that their sampling errors are now generally much lower than those generated by the mechanical tolerances in the analogue computer. In addition, modern digital systems are now able to operate at speeds which make them essentially 'real time'.

9. **Hybrid Computers.** Hybrid computers use a mixture of analogue and digital techniques. They were originally used to overcome the slow speed of digital machines where real-time operation was required, while still retaining memory, accuracy, reliability, and programming flexibility. In addition to suffering to some extent from the drawbacks of analogue machines, they also require analogue/digital and digital/analogue conversion devices. Hybrid computers are still sometimes used in older inertial navigation systems, but they have largely been superseded by digital computers.

10. **Digital Computers.** Digital computers are in widespread use for airborne applications. There are two main types:

- a. The general-purpose computer which can be adapted for a variety of uses by suitable programming.
- b. The special-purpose computer which is designed by the manufacturer to perform a specific task.

The general-purpose computer is much cheaper to produce and easier to upgrade, particularly when commercial off-the-shelf (COTS) components are used in the manufacturing process. The characteristics of the special-purpose type are optimized for the task in hand and it therefore tends to be more efficient. However, this advantage must be weighed against additional cost of production, options for growth, and how quickly the unit becomes obsolete.

11. The Benefits from Digital Computers. The programmability and versatility of digital computers has enabled interconnectivity between aircraft state sensors (speed, angle of attack, etc) and sensors which provide situational awareness (target range, terrain elevation, collision avoidance, etc). Digital processing provides precise solutions over a large range of flight, weapon and sensor conditions. The reliance on specific hardware has reduced, and software has become more important in controlling the interaction between avionics components.

The Airborne Environment

12. Ground based computers usually operate in clean air-conditioned surroundings with little chance of physical damage, whereas the airborne environment is essentially hostile to electronic equipment which may be subjected to large temperature changes, vibration, and acceleration forces. Thus, high standards of hardware ruggedness are necessary. It is clearly advantageous for the computer to have low size, weight, and power requirements, but although developments in integrated circuitry have been beneficial in these respects, the dense packing of components in airborne systems has increased the problems of heat dissipation and the provision of essential cooling.

Precision

13. The precision to which a digital computer can work is a function of wordlength and the required wordlength will be determined by the quality of the various sensor inputs, and the requirements of the systems that the computer is required to drive. Whereas an 18-bit word is sufficient for most navigation, weapon aiming, control, and management functions, other systems (such as imaging tasks) require greater precision, and the current trend is away from 18 or 24 bit words to 32 or 64 bit words.

Input/Output Devices

14. The processed information from an airborne computer will be needed either by other aircraft systems, or by the crew. Where it is necessary for the crew to input or receive data it must be in a form, which is readily interpreted, rather than as a digital data stream. Various devices are used in aircraft to accomplish this, so enabling the crew to interact with the computer. In addition to simple warning lights and flight instruments, the following devices may be encountered:

a. **Storage Media.** Types of storage media include:

(1) **Magnetic Tape.** Magnetic tape may be used both to input and output information. The principle is essentially the same as that used in domestic recording equipment. Because magnetic tape is a serial device it tends to be rather slow in operation.

(2) **Ruggedised Hard Disk.** Hard disks may be used both to input and output information. The 'ruggedised' hard disk is similar in all respects to the hard drive used in personal computers, but has been adapted and strengthened for in-flight conditions. Hard drives have very large data storage capabilities, and are much quicker than tape because information is more accessible and easier to find.

(3) **Optical Disk.** Optical disks are high capacity storage devices that can be used in a similar manner to hard drives. Optical disks do require highly stable conditions in which to operate, as the lasers used to read disks are susceptible to vibration.

(4) **Solid-state Memory Devices.** Recent development in solid-state memory devices has seen marked improvement in storage capability and speed of data access. Current procurement trends are towards solid-state memory devices due to their rugged nature, high capacity and flexibility. The main disadvantage with this form of memory is its cost.

b. **Printers.** Printers are output devices and may be classed as impact or non-impact. Impact printers operate by means of a print head striking an inked sheet or ribbon overlaying the paper. These printers are noisy, relatively slow (typically 1000 lines/minute), but generally cheap. Non-impact printers are typified by ink-jet and laser printers, which, although significantly faster (24,000 lines/minute) and much quieter, are more expensive.

c. **Control and Display Unit (CDU) and Keyboard.** A CDU can be either a high quality cathode ray tube (colour or monochrome screen) or a high resolution Liquid Crystal Display (LCD). With suitable software it can display both alphanumerics and diagrams, and data can easily be edited. The CDU is normally associated with a keyboard to enable manual entry of data by the crew. This may be either a standard QWERTY type as found on a typewriter, or a special type designed to fulfil a specific function. In some systems there is a 'soft' keyboard in which the function of a key is dictated by the computer software according to the mode of operation, and is displayed on the CDU adjacent to the key. Some modern systems use touch-sensitive screens which are generally more accurate than keyboards and are favoured by some operators who find them faster and easier to use.

d. **Direct Voice Input (DVI).** In a DVI system the computer is programmed to recognize a limited vocabulary of command and data words, having first been taught the operator's speech characteristics. The system is inherently faster than keyboard entry.

e. **Manual Input Devices.** A variety of hand controllers (joysticks and roller balls), and switches may be used to input data. Hand controllers are typically used to move cursors on a radar, moving map display, or HUD.

DATA TRANSMISSION

Introduction

15. Once a digital computer is installed in an aircraft there will be a need to transmit data between the computer and other systems such as sensors and displays, or indeed other computers. There is a requirement for this data to have a high degree of reliability and integrity. Reliability refers to ensuring that the message arrives at the intended recipient, and is of the same format that was transmitted. Integrity refers to the accuracy of the message itself and the fact that errors can be detected and corrected if required.

16. **Digital Data Transmission.** With digital data, there are three transmission protocols which can be followed:

- a. **Simplex.** Simplex is one-way flow of information down a transmission media.
- b. **Duplex.** Duplex has a simultaneous two-way flow of information across a transmission media.
- c. **Half-duplex.** Half-duplex protocol is similar to duplex, in that data flows in both directions, but can only do so in one direction at a time. Half-duplex systems will use some form of controlling mechanism to achieve directional flow.

17. **Analogue Data Transmission.** Analogue data transmission methods are still widely used in aircraft systems (see Volume 5, Chapter 24). When analogue systems are present in a digital

network, they must be connected to digital transmission systems by means of analogue/digital converters. This creates problems for the digital system, as analogue data must be digitally sampled. This can produce errors associated with the size of the sample and sampling rate.

Transmission Media

18. **Copper Wire.** Copper wire is still the commonest form of transmission medium, and may be in the form of a screened twisted pair, or co-axial cable. Binary signals are represented by electrical pulses. Cables of this type are relatively inexpensive and easily handled, but have the following disadvantages:

- a. They are bulky and heavy.
- b. They permit only a limited bandwidth.
- c. They create fields which give rise to electro-magnetic interference (EMI) and are also prone to EMI from other sources.

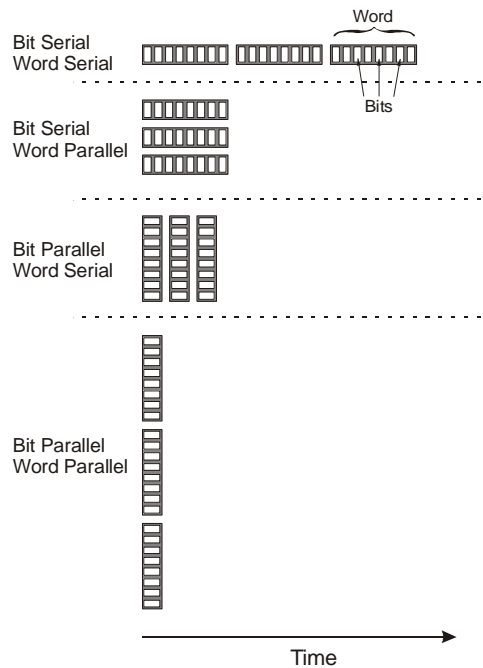
19. **Optical Fibre.** Optical fibres are in common use as transmission media. They consist of lengths of glass fibre, usually clad in plastic, along which binary signals are transmitted in the form of pulses of light. The light sources are either light-emitting diodes (LEDs), or injection laser diodes (ILDs), which operate in the near infrared region of the spectrum. Photodiodes at the receiving end of the cable convert the light signals back into electrical signals. Because optical fibres use light waves to transmit signals they do not suffer from electrical interference caused by high voltages, radio frequencies, magnetic fields, lightning, or electro-magnetic pulse. Similarly they are themselves non-radiating and therefore do not interfere with other electronic equipment. They have inherent security, and are safer to use in a potentially explosive environment. The bandwidth of the medium far exceeds that of copper wire; LED sources can operate at bandwidths up to 100 MHz, and ILDs at up to 100 GHz. In summary, compared to copper wire, optical fibres can transmit greater amounts of data over longer distances, are impervious to electro-magnetic interference, and are physically much smaller, lighter, and less cumbersome.

Channel Configurations

20. A channel is the connection needed to transmit a data word. The number of wires required depends on the form in which data is to be transmitted. The four options available are illustrated in Fig 1 which shows three data words of 8 bits each. An explanation of each of the options is given in the following sub-paragraphs:

- a. **Bit Serial Word Serial (BSWS).** In BSWS format, both the words in the data stream, and the constituent bits of a data word are transmitted serially.
- b. **Bit Serial Word Parallel (BSWP).** In BSWP format, the words within the data stream are transmitted in parallel but the bits within a word are transmitted serially.
- c. **Bit Parallel Word Serial (BPWS).** In BPWS format, the bits within a data word are transmitted in parallel but the words in the stream are transmitted serially.
- d. **Bit Parallel Word Parallel (BPWP).** In BPWP format, both blocks of words and their constituent bits are transmitted in parallel.

Serial transmission requires fewer wires and is relatively cheap and light, but slow. Parallel transmission is more expensive, complex, and heavier, but faster.

7-27 Fig 1 Channel Configurations

Transmission Control

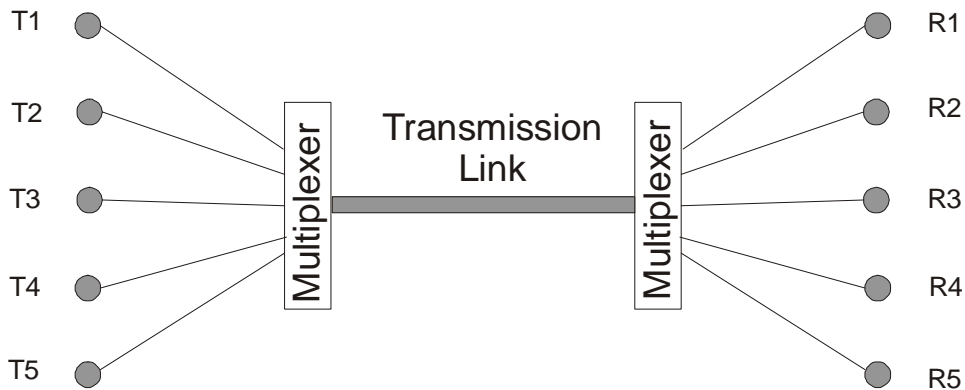
21. In most systems, data is transmitted as a stream of data words, preceded by a command word. The command word contains such details as the number of words in the ensuing stream, and the destination address. The data will often have inbuilt error detection and correction by using bits within words (eg Parity Checking) or additional words (eg Checksum Words).

22. The transmission of data streams must be controlled to ensure that the appropriate information reaches the correct destination. There are two types of control, synchronous and asynchronous:

a. **Synchronous.** Under synchronous control, peripherals will be accessed in a strict sequence under some form of central control, normally based on a clock. Each peripheral is connected for the length of time necessary to pass the maximum number of permitted data words regardless of how much information is actually transmitted. Therefore, although such a system is relatively simple to design and construct, it tends to be slow, inflexible, and inefficient.

b. **Asynchronous.** In an asynchronous system, when a peripheral has information to transmit it tells the processor which arranges a connection, and maintains it until the message has been passed. On completion, the connection is broken, allowing another peripheral to use the line. An asynchronous system is more complex than a synchronous one, but is faster and more efficient. Asynchronous control is becoming more widely used as experience and technology improves.

23. **Multiplexing.** Multiplexing provides a means of reducing the amount of hardware required by sharing transmission channels. The concept is illustrated in Fig 2 where T1, T2, etc are transmitters and R1, R2, etc are receivers. There is a single shared link through which signals between each pair of transmitters and receivers are synchronized and processed. This link can be 2-way. There are three main types; frequency division multiplexing (FDM), time division multiplexing (TDM) and code division multiplexing (CDM)

7-27 Fig 2 Multiplexing

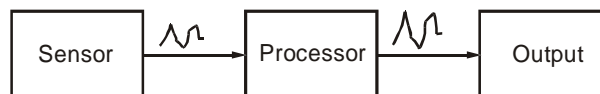
COMPUTER ARCHITECTURE

Processing Options

24. The decision as to which computer arrangement is appropriate in any aircraft will depend on the scale of the computing task and the number of systems to be controlled or integrated, and on whether real-time operation is required. The four main architectures currently in use are:

- a. Single processing.
- b. Dual processing.
- c. Multiprocessing.
- d. Distributed processing.

25. **Single Processing.** In a single processing arrangement all tasks are performed in a single computer. This organization was favoured when computers were first installed into aircraft. The arrangement often has very low integrity, as there is no redundancy or reversionary capability; failure of the single processor results in the loss of all computing capability (Fig 3).

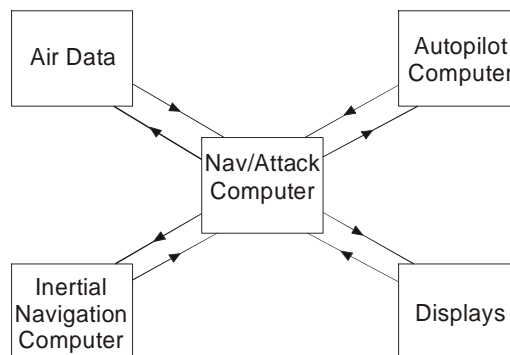
7-27 Fig 3 Example of Single Processing

26. **Dual Processing.** In a dual processing organization, two digital computers work independently, sharing the same function. It is possible for this arrangement to provide better integrity than the single system if essential programs and data are stored in both computers, such that if the primary processor for any particular function fails, the other can take over the task (this may of course entail the loss of some less essential capabilities). Dual systems can offer a limited real-time performance, but, if this capability is required, a multiprocessing or distributed organization is much to be preferred.

27. **Multiprocessing.** In a multiprocessing arrangement two or more central processing units (CPUs) operate with one memory. An elaborate supervisory program allocates processor time according to predetermined priorities. The multiprocessing system has high integrity and good real-time performance, but at the cost of complex and difficult programming.

28. **Distributed Processing.** In a distributed system, separate computers are used for the various tasks, but with one of these exercising some control over the others. The arrangement is shown in Fig 4. The controlling computer is used to reduce operator workload by performing some of the switching functions needed for the management of the system and to provide centralized control of reversionary routines in the event of equipment failure. The distributed system can have a good real-time performance and there is less of a programming problem compared with a multiprocessing arrangement. Failure of a dedicated computer in such a system would probably entail the loss of that element and critical tasks may therefore have to be protected by the provision of redundant machines. System integration will suffer if the controlling computer fails; however, the dedicated computers will continue to operate and a well-designed system will make their information available, even if in a degraded mode.

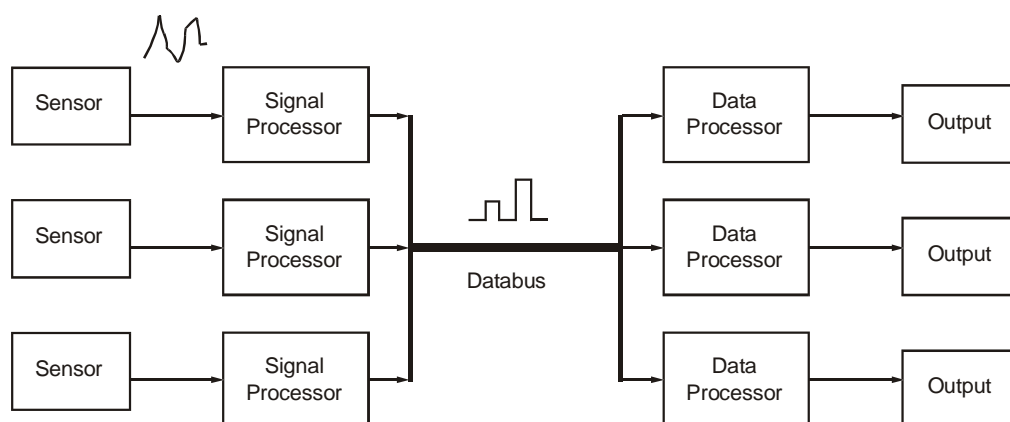
7-27 Fig 4 Distributed Processing System



Federated Avionics

29. Major development of the distributed processing methodology took place during the late 1970s and early 1980s. Processors and sensors became smaller. Heavy and cumbersome point-to-point wiring connections within aircraft avionics were replaced by the use of shared data 'highways', now referred to as Data Buses. This architecture has become known as 'Federated Avionics' and is typical of most current military systems (see Fig 5).

7-27 Fig 5 A Federated Avionics System



Integrated Modular Architecture

30. Integrated Modular Architecture (IMA) systems are now being developed, to provide improved processing flexibility, and increased mean time between failure (MTBF). Within an IMA system, the peripheral sub-systems are connected to centralized processors via a high-speed data bus. However,

the central processors each now consist of multiple, common computing modules. The tasks of the computing modules are software-defined and, if one module fails, any other module can be reconfigured to take over its tasks. This design therefore provides a very flexible redundancy capability and almost instantaneous recovery.

Open Systems Architecture

31. Military avionics systems have traditionally been 'closed' and largely aircraft-unique. The reasons for this are largely historic, in that different manufacturers produced differing designs which were not compatible with equipment from rival companies. This is not a problem when a complete system is produced by the same manufacturer. However, if a sensor manufactured by Company A was used with a display manufactured by Company B, then problems may well arise. To enable the two sides of the system to work together, an interface box would be needed so that the output from the sensor could be translated into a form that the display would recognise as input. This interface would be an expensive item to produce, as it would be specific to the two pieces of equipment concerned. This same situation existed in the field of Personal Computers (PCs) until a few years ago. However, things have changed rapidly and now most peripherals can be plugged in to any PC and they will work straight away. This 'plug and play' technology is possible through comprehensive agreement on interface specifications, now universally applied by all manufacturers. This is 'Open Systems Architecture' and the avionics industry has now realised the benefits which can accrue from it and are actively pursuing the methodology.

DATA BUSES

Introduction

32. Data buses are arrangements whereby multiple electronic devices are connected to a common data busbar, and the flow of data is controlled by a bus controller using a predetermined protocol. The bus controller will be informed that a peripheral wishes to transmit and the appropriate receiver will be commanded to receive. The command will be acknowledged and the data transmitted.

33. Avionics data buses are designed to accept system expansion or modification efficiently, without the costly and time-consuming exercise of changing the aircraft wiring. A change of equipment would involve connections to, or disconnections from, the bus, and modification to the bus controller software.

34. **The Bus Controller.** The bus controller is the most important, complex, and costly part of the system and its major function is to ensure that information is routed correctly between remote terminals. In addition, it monitors the status of remote terminals and if, for example, one source of information failed, the bus controller would automatically arrange for systems requiring that information to receive it from a secondary source, if available.

35. **Data Rates.** The rate at which digital data may be transmitted is measured in 'bits per second' (bps). Modern data buses can transfer data at rates from 1 Mbps (Mil-Std-1553B) to 1 Gbps (Fast Ethernet).

36. **Weight and Space.** The use of a data bus in an aircraft, in place of conventional wiring looms, will save a large amount of weight, even for a small fighter aircraft. They are less prone to chafing, making them inherently more reliable. Data buses are also less bulky than the equivalent wiring looms and, therefore, save space.

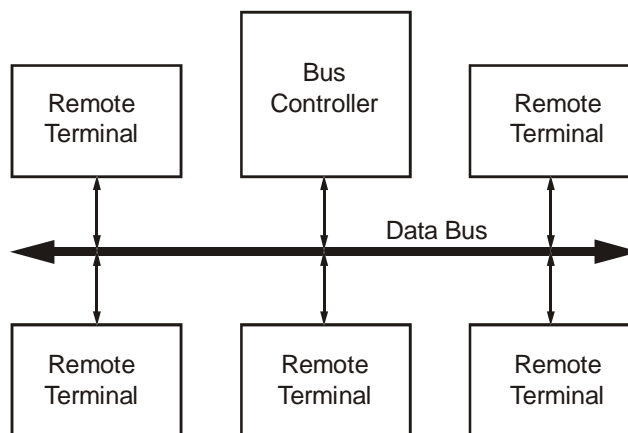
Data Bus Examples

37. There are a number of different databuses in use in avionics applications, with slight differences between civil and military design. However, high development costs may bring them more closely aligned in the future.

38. Flight safety is the key consideration for avionics databuses. At present, well-proven and reliable avionics databuses include ARINC 429, Mil-Std-1553B and Mil-Std-1773. These three databuses are described briefly in the following paragraphs.

39. **Mil Std 1553B (NATO STANAG 3838) Data Bus.** The Mil Std 1553B data bus (known as "Fifteen Fifty-three") was introduced in 1973 by the US Department of Defense as a standard format for aircraft data buses and all new US aircraft were to employ the system. The system utilizes time division multiplexing, and is organized such that up to 30 remote terminals can be connected to a common data bus (Fig 6). A remote terminal can be embedded in a particular avionic component, or can stand alone and service up to five avionic systems. The disadvantages of the system are its complexity and slow speed. The transmission medium is a twisted pair of copper wires, which limits the bandwidth to 1 MHz and suffers from the other disadvantages of electrical transmission (see para 18).

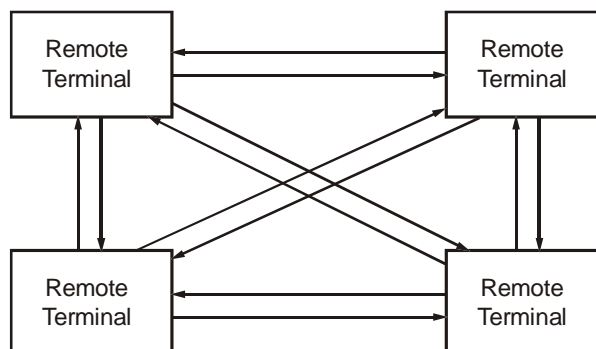
7-27 Fig 6 Example of Mil Std 1553B Architecture



40. **Mil Std 1773 Data Bus.** The Mil-Std-1773 ("Seventeen Seventy-three") introduces the use of fibre optics as the transmission medium for the Mil-Std-1553B. This provides military aircraft with an EMI-resistant data bus. Differences between the two standards were kept to a minimum to permit Mil-Std-1553B compatible equipment to be upgraded to a fibre-optic transmission medium. The actual fibre-optic bus architecture is not specified and Mil-Std-1773 allows several acceptable bus configurations. Optical power levels, wavelength and power distribution are left to the designers.

41. **ARINC 429.** The ARINC 429 standard was agreed by the commercial airline industry in 1977-78. It is a single-source, multiple-receiver data bus constructed on uni-directional flows of data (see Fig 7).

7-27 Fig 7 ARINC 429 Architecture



ARINC 429 uses simple, reliable hardware which is relatively inexpensive. Because its bus data rates are relatively slow, and the bus is uni-directional, there is no requirement to incorporate sophisticated bus access protocols to prevent transmission collisions. Consequently, it finds favour in commercial aviation since flight certification is easier. The data bus uses two single wires to transmit 32-bit words. No more than 20 receivers should be connected to a single bus. Since each terminal must be connected to each of the others in order to communicate, expansion of the system is complicated, requiring a significant number of extra connections.

CHAPTER 28 – REAL-TIME PROGRAMS

Introduction

1. A real-time system is a combination of computer hardware and software which has the ability to process data sufficiently quickly that it can keep pace with events and influence or control responses with minimal time lag. In the case of airborne systems, the acceptable time lag will be in the order of milliseconds.

2. A system's ability to operate in real time depends principally on the amount of Central Processor Unit (CPU) time available and, as suggested in Volume 7, Chapter 27, multiprocessing or distributed processing systems are normally used. In a multiprocessing system, the supervisory software allocates tasks to the CPU based on priorities. Despite the complex software involved, this system is very flexible and is probably the most suitable when tasks occur at random times. The distributed arrangement allocates specific tasks to dedicated computers. This avoids the need for a complex supervisory program, but requires careful initial design and accurate forecasting of system workload. Communication between machines may cause delays, especially if dissimilar computers are used, as different word lengths and input/output characteristics dictate the need for complicated interface units.

Iteration Rates

3. An airborne computer will usually have several different programs to run. Each program will take a certain amount of time and must be repeated at certain intervals. The number of times each program is repeated in 1 second is termed the iteration rate and is expressed in Hertz (Hz). Thus, for example, an iteration rate of 10 Hz means that the program must be completed every 100 ms. The iteration rate will be determined by consideration of the maximum error in a variable that can be permitted and the maximum rate at which the variable can change. For example, it may be that a certain aircraft navigation system cannot tolerate an error in pitch of greater than 1° if it is to meet the specified accuracy. If the maximum pitch rate of the aircraft is 20° per second, the iteration rate must be at least 20 Hz, and, in practice, a higher rate would be chosen to give a safety margin, provided that sufficient CPU time is available. Some typical iteration rates for various airborne computing tasks are shown in Table 1.

Table 1 Typical Iteration Rates

| <i>System</i> | <i>Iteration Rate (Hz)</i> |
|-----------------------|----------------------------|
| Air Data Computer | 20 |
| Autopilot (stability) | 100 |
| Autopilot (control) | 50 |
| Head-up Display | 50 |
| Weapon Aiming | 50 |
| Routine Navigation | 10 |

Priority and Interrupt System

4. In order to achieve the required real-time performance with a digital computer, the tasks will be grouped into a number of priority levels with the most important tasks (generally those with the highest iteration rates) having the highest priority. A series of interrupt pulses generated by a real-time clock will be used. Interrupt signals from a peripheral, such as a navigator's control unit, may also be used. For convenience of organization, the iteration rates of programs on the same level may be changed to

ensure that they are multiples or sub-multiples of each other. As an example, it may be necessary to run three programs, A, B, and C at the same level, with program A requiring a rate of 5 Hz, B a rate of 7 Hz, and C a rate of 9 Hz. In this situation, it would be more convenient to run A at 5 Hz as needed, but to increase the iteration rates of both B and C to 10 Hz. Iteration rates would not normally be reduced as this would, in most cases, entail either a lower safety margin or decreased accuracy.

5. The timing, initialization, control, and scheduling of work is accomplished by a supervisory program, which also handles the input and output of data and the servicing of interrupts. The heart of the supervisory program is the main scheduler routine which determines the order in which processing is done, and allocates resources to the various programs.

6. The sequence of events is as follows:

- a. After each instruction is complete, a check will be made on the contents of an interrupt status word (ISW) in a special register. If a particular bit is set to 0, this indicates that an interrupt has not been generated and the computer will go on to the next instruction. If, however, a 1 is present, an interrupt has been generated.
- b. If a 1 is present, the rest of the ISW will be examined to determine the priority level of the interrupt.
- c. If the interrupt is of lower priority than the program currently running, the interrupt will be ignored. If it is of a higher priority, the contents of various registers, such as the accumulator, will be stored, and the location of the first instruction of the new program will be loaded into the program counter.
- d. The new program will then be run until it is complete or is, in its turn, interrupted by a still higher-level program.
- e. When a program is complete, and there are no further programs to be run at that level, the last instruction will cause the registers to be loaded with the values pertinent to the next most important program, which will then be run until it is complete or interrupted.

7. As an example, consider a computer being used in a nav/attack system, having the following four program priority levels:

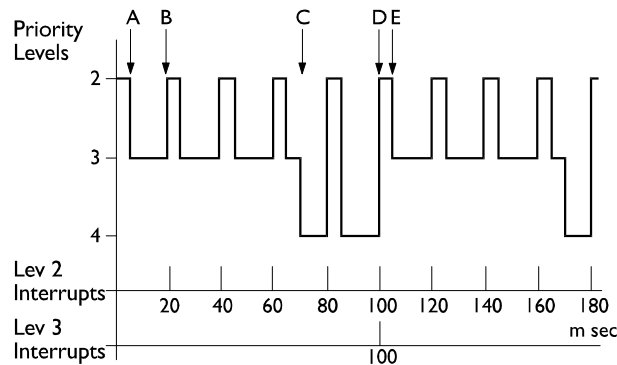
- a. Level 1 is the highest priority level and is used only for switch on, switch off, and fault conditions.
- b. Level 2 is used for programs requiring an iteration rate of 50 Hz, such as an Inertial Navigation Schuler loop, or weapon aiming calculations. An interrupt signal is generated every 20 ms to ensure that a rate of 50 Hz is achieved.
- c. Level 3 services routine navigation equations and the generation of display information. The iteration rate required is 10 Hz and so interrupts are generated every 100 ms.
- d. Level 4 is used for self-test routines and programs are run only when time is available in the CPU after the tasks at the higher levels have been completed.

8. The nav/attack system may be operated in several modes and, in this example, it will be assumed that the computer is operating in the routine navigation mode in which Level 2 programs require 5 ms per iteration and Level 3 programs require 50 ms per iteration. The allocation of computer time and the associated hardware interrupt signals is illustrated in Fig 1 and described below:

- a. At $t = 0$, the CPU begins the Level 2 program.

- b. At point A ($t = 5$ ms), the Level 2 program is complete and the last instruction causes an automatic reversion to the next lowest priority level - Level 3.
- c. At point B ($t = 20$ ms), a hardware interrupt signal is generated which demands that the Level 2 program is serviced again. Before the computer leaves Level 3 the address of the current instruction and intermediate data results are automatically stored in protected memory locations. Only 15 ms of the 50 ms needed by the Level 3 program has been made available at this stage.

7-28 Fig 1 Priority System



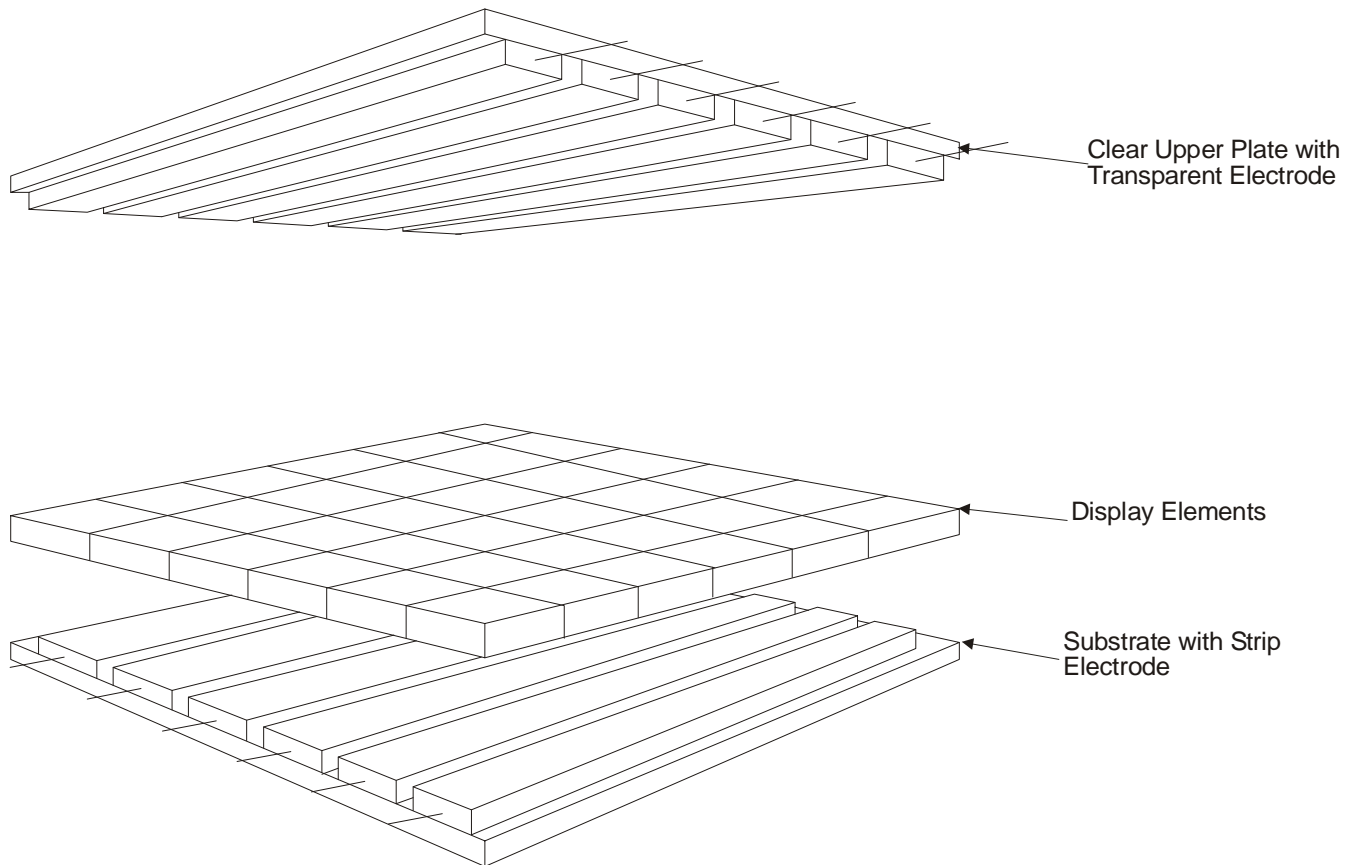
- d. The sequence now repeats with the Level 2 program being serviced every 20 ms to maintain the 50 Hz iteration rate. The time available between Level 2 iterations is spent at Level 3 where the stored data and instruction addresses are used to ensure continuity.
- e. At point C ($t = 70$ ms), the Level 3 program has been completed and time is available for Level 4 programs. Program running will now alternate between Levels 2 and 4.
- f. At point D ($t = 100$ ms), interrupt signals are received for Levels 2 and 3. Level 3 must be serviced again to achieve the required 10 Hz iteration rate, but the Level 2 has priority and the Level 3 interrupt is stored until the Level 2 program is complete (point E), when the Level 3 program can be commenced. The whole cycle is then repeated.
9. The time spent at various levels will vary with the mode of operation. Suppose that the operator carries out an attack using a weapon-aiming mode which requires 5 ms of calculation at Level 2. The total time required at Level 2 is now 10 ms. In every 100 ms period the time required by Levels 2 and 3 is now 100 ms (5×10 ms at Level 2 and 1×50 ms at Level 3). Thus, no time is available to service the self-test programs at Level 4 and these must be dropped for the duration of the attack.
10. It may be necessary to adjust the Level 3 tasks at some stages of flight. Suppose the weapon-aiming mode selected requires 7 ms of time at Level 2, giving a total of 12 ms per iteration. In every 100 ms period the Level 2 program now requires 60 ms (5×12 ms) leaving only 40 ms available at Level 3. As Level 3 requires 50 ms in every 100 ms to achieve a 10 Hz iteration rate some adjustment must be made, either by accepting a lower iteration rate for the Level 3 programs, or by reducing the Level 3 tasks, ie during the period of the attack some less important facility or information will be lost. In most cases, this will be possible without significantly degrading the overall system performance.

CHAPTER 29 - FLAT DISPLAYS

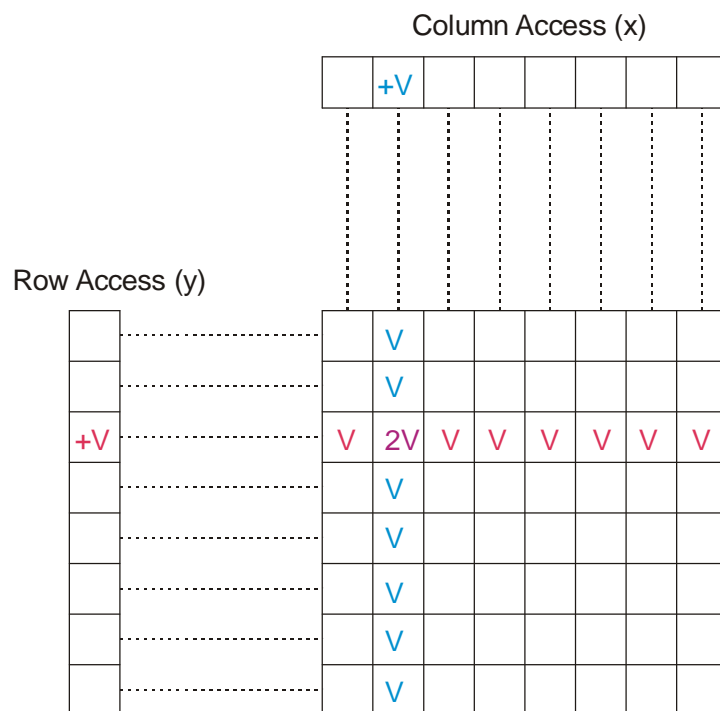
Introduction

1. Although CRTs will have a place in airborne displays for the foreseeable future, they have a number of disadvantages. They are bulky - in particular requiring considerable depth behind the display face; they are heavy; they operate at very high voltages; and interfacing them with digital equipment is complex. These disadvantages can be overcome by the use of flat panel displays.
2. Flat panel displays normally consist of a matrix of individual elements and the display resolution will be defined by the number of these elements. For example, a display of comparable resolution to a 625 line TV picture would require a matrix consisting of 585×704 elements - a total of 411,840 elements. The problem of controlling the voltages across such a large number of individual elements is usually overcome by using an X-Y (Cartesian co-ordinate) addressing procedure.
3. The simplest addressing procedure is called a 'Passive Matrix' in which the element at the intersection of two power strips is energized. The passive matrix display is constructed as a three layer sandwich. The middle layer comprises the display elements; the top and bottom layers are strip electrodes, set mutually at right angles. The top electrode layer is transparent so that the display elements can be viewed through it (Fig 1). Any individual element can be addressed by a signal passing through one electrode strip in each layer (Fig 2).
4. This arrangement is suitable for binary signalling and, for example, an array of $1,024 \times 1,024$ elements can be addressed by two ten-digit X and Y inputs. Such systems can be scanned in a raster manner as in a conventional CRT, or elements can be randomly addressed by means of their unique X, Y label. A problem with this type of system is that for a typical display of 10^6 pixels, and with a refresh rate of 50 Hz, each element can only be addressed for $(50 \times 10^6)^{-1}$ secs during each frame and so the ideal element will have a very short 'turn-on' time and will remain on until extinguished (inherent memory).
5. An alternative to the passive matrix is to directly address individual elements (called an 'Active Matrix'). The main difficulty with an active matrix is the number of individual connections. An active matrix display with $1,024 \times 1,024$ elements will need 1,048,576 connections, whilst the passive matrix would only require 2,048 connections. The principal advantages of the active matrix are that each element can be switched on and off directly. Also, any connection failure will affect only one element and not the entire row, as would be the case with a passive matrix.

7-29 Fig 1 Construction of a Matrix Display



7-29 Fig 2 Matrix Addressing



Element is 'ON' if voltage across it exceeds 1.5V

Display Types

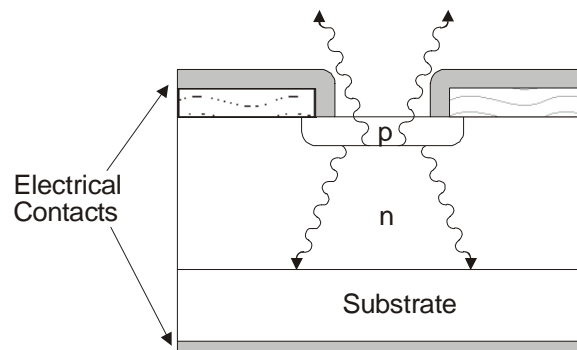
6. Five types of flat panel displays are currently in use in avionic systems:

- a. Light Emitting Diode (LED).
- b. Liquid Crystal Display (LCD).
- c. Active Matrix LCD (AMLCD).
- d. Plasma Display.
- e. Electroluminescent Display.

Light Emitting Diode (LED)

7. A light emitting diode is a semiconductive junction which emits light when a current is passed through it. Fig 3 shows the construction of a typical LED, in which a shallow p-n junction is formed. While electrical contact is made to both regions, the upper surface of the p material is largely uncovered so that the flow of radiation from the device is impeded as little as possible. The primary materials used are gallium arsenide, gallium phosphide, and gallium arsenide phosphide. LEDs are most suitable for 'on-off' displays rather than in applications requiring a grey scale. LEDs have no inherent storage and, if addressed with a passive matrix, the display must be refreshed at a rate fast enough to avoid flicker. LEDs are lightweight, have good brightness, low power requirements and long life; they are currently replacing most CRT displays in Helmet Mounted Displays.

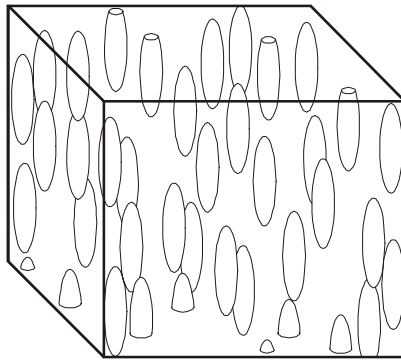
7-29 Fig 3 Construction of Typical LED



Liquid Crystal Display (LCD)

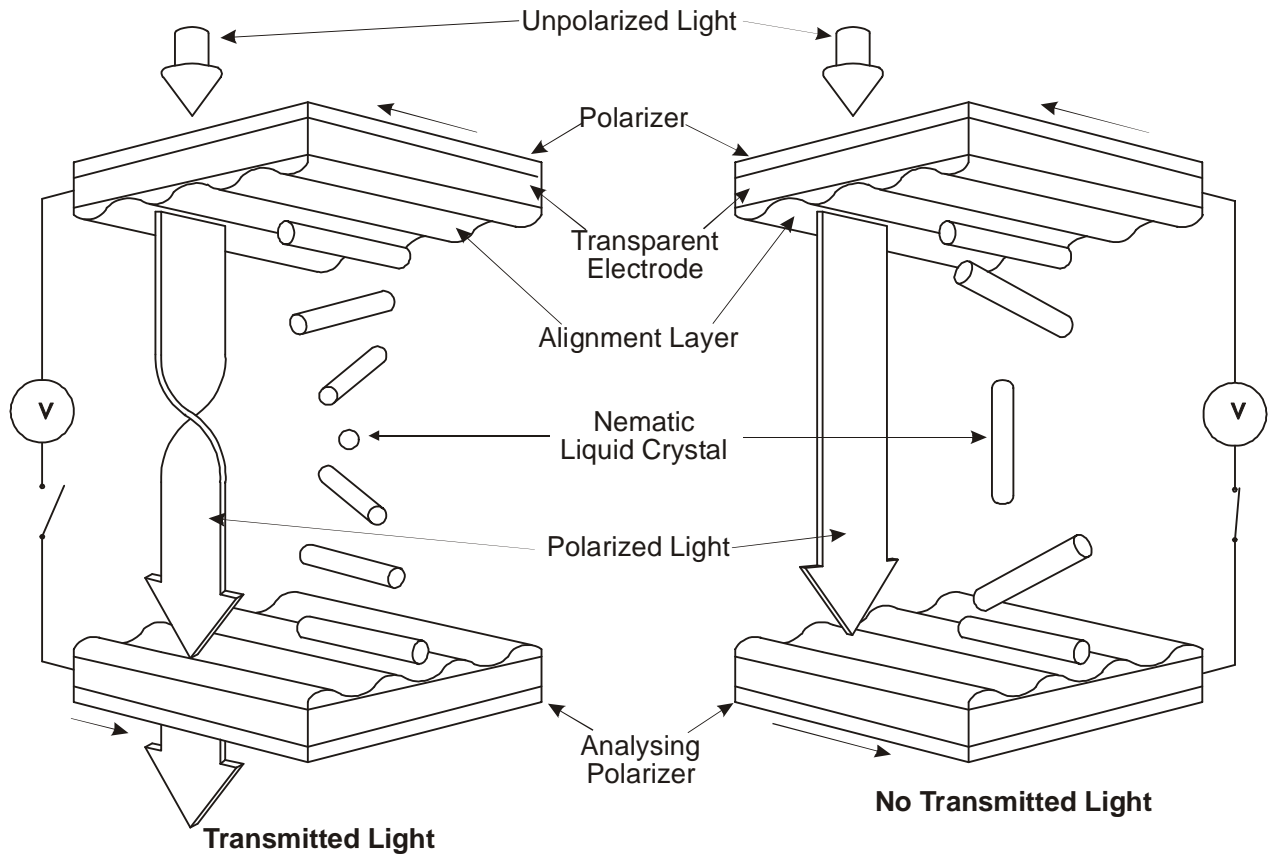
8. Liquid crystal displays are unlike other flat displays in that they are not light emitters, but rely on an external light source for their operation. Often the light source is placed behind the LCD, as is the case with laptop computer displays.

9. Liquid crystal is an organic compound which, while having the physical characteristics of a liquid, has a molecular structure akin to a crystalline solid. There are three classes of liquid crystal which vary in their molecular structure and, although all three have been used in LCDs, the structure known as 'nematic' is by far the most common. In this structure, the elongated, rod-shaped molecules are aligned parallel to each other but not in regular layers (Fig 4).

7-29 Fig 4 Nematic Molecular Structure

10. In the display cell, the inner surfaces of the top and bottom glass or perspex walls are grooved, with the top grooves aligned at 90° to the bottom grooves. The grooves induce a corresponding alignment of the molecules so that their alignment within the liquid crystal twists through 90° (the twisted nematic structure). The top and bottom of the display cell is covered by linear polarizing plates such that the plane of polarization of one plate is at 90° to the other.

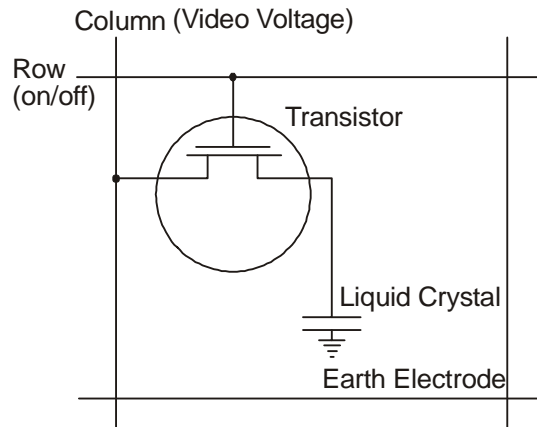
11. When light passes through the initial plate it is polarized, and the twisted nematic structure causes the plane of polarization to be rotated through 90° so that the light is able to pass out through the second polarizer unimpeded; the cell therefore has a transparent appearance. When a voltage is applied across the cell, the molecules tend to align themselves with the field thus destroying the twisted structure. The polarized light entering the cell will no longer have its plane of polarization twisted through 90° and it will not therefore be transmitted by the second polarizer and the cell will appear dark. When the field is removed, the molecules return to the original twisted nematic structure. The structure and operation of a twisted nematic LCD is shown in Fig 5.

7-29 Fig 5 Structure and Operation of Twisted Nematic LCD

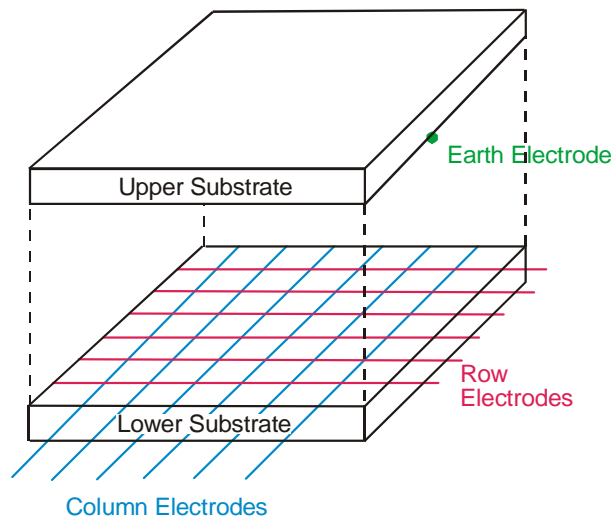
12. The normal display is one of dark characters on a light background although this can be reversed by arranging the polarizers parallel rather than at 90°. Coloured displays are possible by adding dyes to the liquid crystal material or by the use of colour filters. Full colour displays are now easily produced and available for most applications. LCDs must be either front or back lit and do not operate at very low temperatures (heating is required for aviation displays).

Active Matrix LCD (AMLCD)

13. AMLCDs are more common in aviation than passive matrix LCDs. In an AMLCD, the voltage applied on each element is actively controlled by a transistor, as shown in Fig 6, ensuring that the liquid crystal receives the correct voltage during the address time, and is isolated from stray voltages when it is switched off.

7-29 Fig 6 AMLCD Element

14. Fig 7 shows a cross-section of an AMLCD. In this example, rows and columns of the matrix are disposed on the lower substrate, the upper substrate carrying the earth electrode. An element (or pixel) is addressed by applying the video voltage corresponding to the signal to be displayed on the column, and a voltage to energize the element on the appropriate row. The element is then turned off while the other rows of the display are successively addressed.

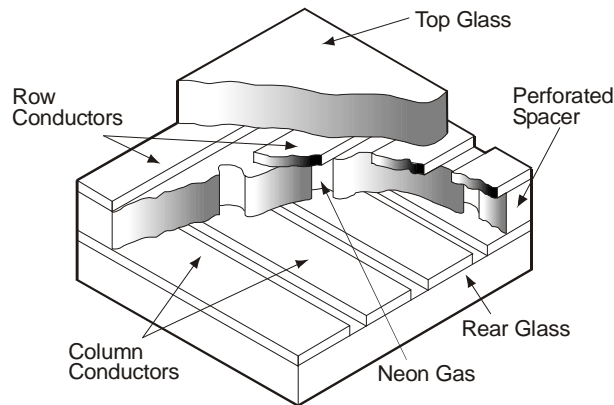
7-29 Fig 7 Representative Cross-section of AMLCD Display

15. A limited grey effect can be obtained by modulating the amplitude of the input video voltage. The display needs to be refreshed periodically due to leakage currents.

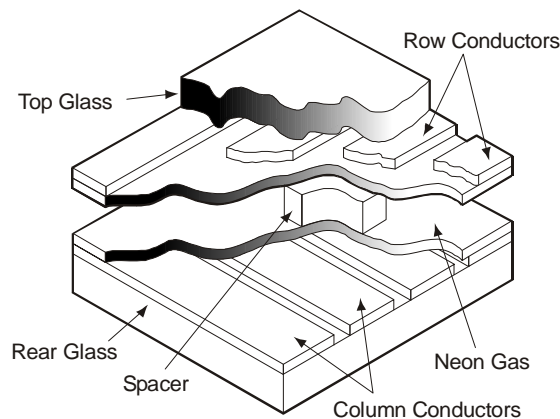
Plasma Displays

16. Plasma (gas discharge) displays use an electrical discharge in a gas to produce light; both DC and AC systems are available.

17. A DC plasma display consists essentially of a gas filled space between two electrodes (Fig 8). When the DC potential across the electrodes exceeds a certain value (typically 180V), which depends on the gas type, pressure and the electrode gap and type, the gas molecules ionize and emit light. The DC technique has no inherent memory and therefore requires constant refreshing.

7-29 Fig 8 DC Plasma Display

18. An AC plasma display has a similar basic concept, but the electrodes are insulated; the structure is illustrated in Fig 9. A voltage is applied to all the row electrodes and is antiphase to all column electrodes. The field generated across the gas is insufficient to strike a discharge and, in order to light a particular pixel, the AC voltages on the appropriate row and column are increased for one half cycle. This causes a capacitive current to flow and build up a charge at the insulating layers. Subsequent cycles are at the normal AC voltage, but this is sufficient to maintain the discharge previously created. To switch off the pixel, row and column voltage must be selectively lowered. AC types of plasma displays have inherent memory for each element.

7-29 Fig 9 AC Plasma Display

19. Older plasma (gas discharge) displays are not generally suitable for producing grey scales and are primarily available in neon orange colour for use in on-off displays.

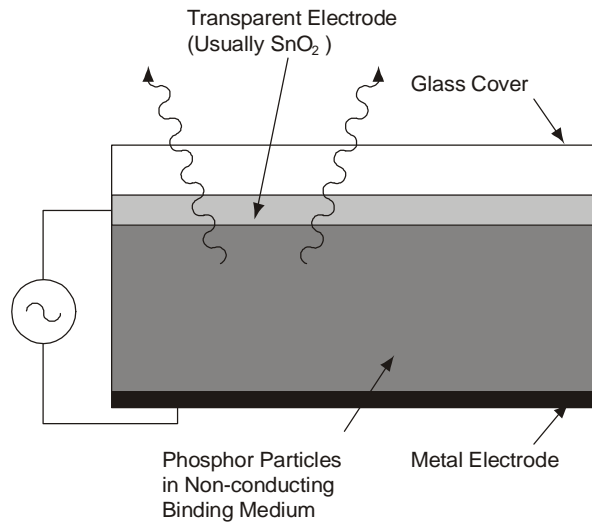
20. Plasma displays are now available in full colour and operate on a principle much like the CRT. The display is made up of an array of individual cells, each containing plasma gas, coated with red, green or blue phosphorous. Energizing the cell excites the plasma which emits Ultra-Violet (UV) radiation. The UV radiation, in turn, energizes the phosphorous which glows red, green or blue, with each pixel on the display made up of one red, one green and one blue element. The advantages of the plasma display are its brightness and wide viewing angle.

Electroluminescent Displays

21. Electroluminescent displays consist of a layer of phosphor, sandwiched between two electrodes, which glows when an electrical field is applied across it. Displays may be either AC or DC driven and the structure of each type is somewhat different.

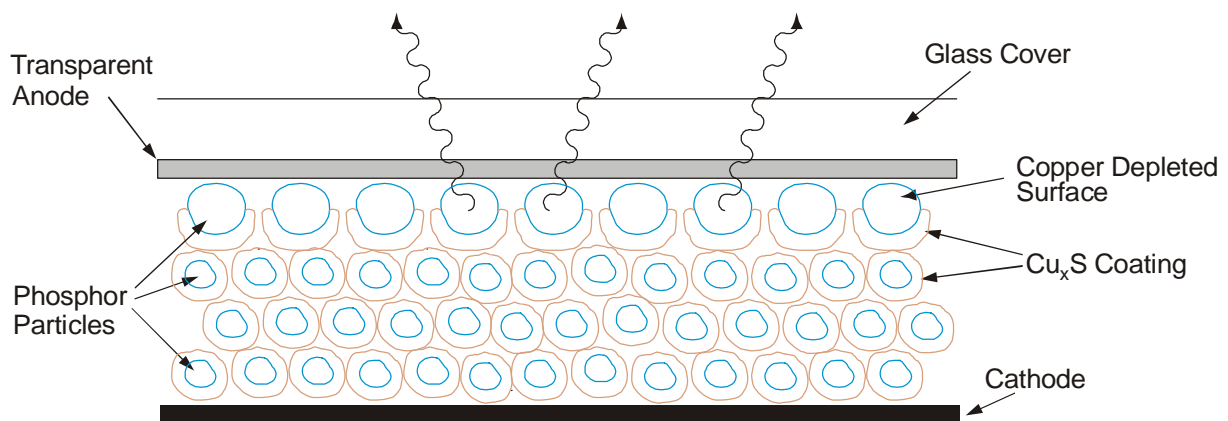
22. Fig 10 shows the structure of an AC device in which phosphor particles are suspended within a transparent insulating medium (thick film technique) and sandwiched between two electrodes, one of which is transparent. As an alternative, the phosphor can be deposited, normally by evaporation, as a thin layer onto a dielectric base (thin film technique). The phosphor particles emit light when an AC voltage is applied.

7-29 Fig 10 Structure of Thick Film AC Electroluminescent Device



23. Fig 11 shows the construction of a DC device. The phosphor particles have a coating of either Cu₂S or Cu₃S (generally termed Cu_xS) which is removed from the anode side of the particles in contact with the anode by the application of an initial high current pulse. Light is emitted from the Cu_xS depleted particles when a normal DC voltage is applied.

7-29 Fig 11 Structure of DC Electroluminescent Device



24. Electroluminescent devices have full video capability and have potential as a replacement for CRTs in Head Mounted Displays. All colours are available, dependent on the phosphor selected, and a full colour display is available in very small devices.

CHAPTER 30 - PROJECTED AND ELECTRONICALLY DISPLAYED MAPS

Introduction

1. The most widely used navigation aid for low level VMC operations is the topographical map. Such a map allows very accurate pinpoints to be obtained, and also presents information about the aircraft's position in relation to its surroundings in a relatively easily assimilated way. However, the use of conventional maps, covering large areas, presents handling problems in small aircraft cockpits. Moving and electronic map systems were devised to overcome these difficulties.

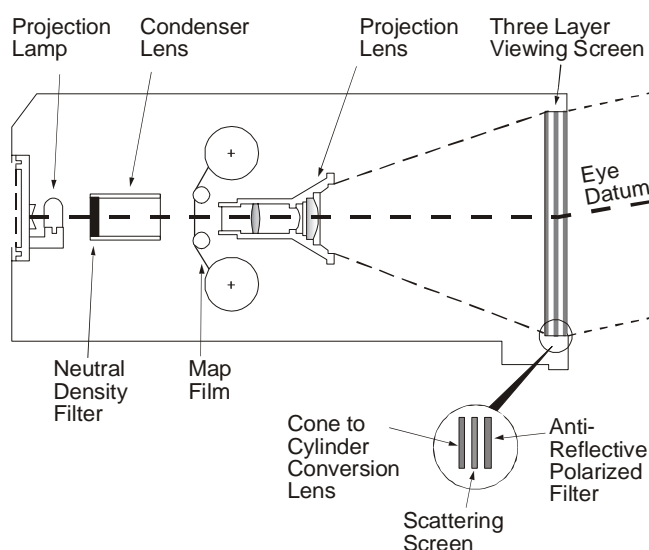
2. Early systems used strips of paper maps wound on rollers with an overlying cursor to indicate position; the rollers and cursor were driven by outputs from a doppler or radio navigation aid. However, because the map was cut to suit the planned route, these early systems offered limited flexibility when faced with changes to route during flight. The second generation of moving map displays solved this problem by projecting a 35mm film strip onto a display screen. These offered much improved tactical capability although, being analogue systems, they had minor problems in accuracy. The latest generation of moving maps use electronic data and are rapidly replacing the projected systems.

Projected Map Displays

3. The projected map display (PMD) employs a coloured map transferred to 35mm film. This is back-projected, using conventional optics, to give a bright image on a translucent screen. The map is driven by an inertial or mixed inertial navigation system, such that the drive mechanism keeps up with the aircraft's ground position. The map projection can be orientated with aircraft track or true north.

4. A typical PMD system is illustrated in Fig 1. The screen is designed to concentrate the image luminance within a limited field of view, matched to the observer's eye, in order to increase the resistance to strong ambient light. It will be seen that the screen has three layers. The first (inner) layer is a Fresnel lens which converts the light cone from the projection lens into a light cylinder in the plane of the operator's eye datum. The image is formed on the second layer which is designed to minimize hot spots towards the centre, and image degradation towards the circumference. The third (outer) layer is a polarizing filter which eliminates reflections from both inside and outside the PMD which might otherwise obscure the image.

7-30 Fig 1 Schematic of a Projected Map Display



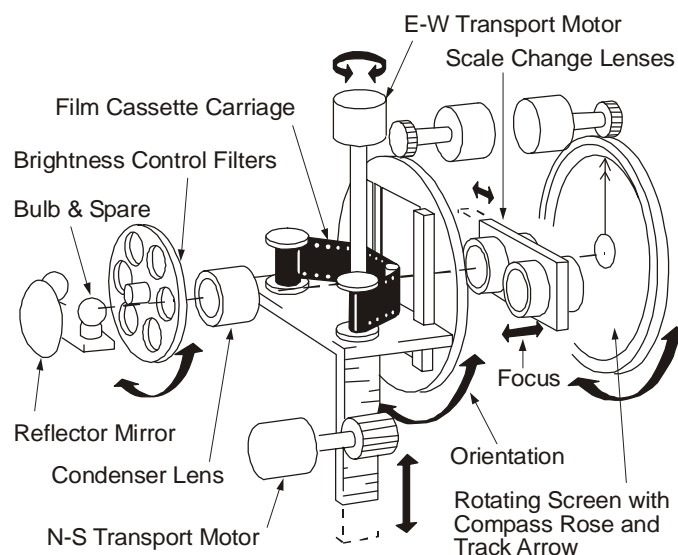
5. In a typical PMD, the map is photographed in segments onto 35mm film. Up to 4 million square miles of map coverage, at a scale of 1:500,000, can be reproduced on a 20 metre strip of 35mm film. In practice, some area coverage will usually be sacrificed in order to have a selection of map scales available, and there may also be check lists and terminal charts included. In normal operation, the change over from one frame to the next is automatic and is usually accomplished in under three seconds. The life of the film strip tends to be limited by the currency of the map rather than by fading, or wear and tear.

6. Within a PMD, scale change can be accomplished by either:

- a. Increasing magnification by lenses.
- b. Changing between map segments on the 35mm film.

Fig 2 shows a simplified diagram of the internal construction of a typical PMD. In this example, the magnification option is employed.

7-30 Fig 2 Simplified Diagram of a Typical PMD



7. The accuracy of a PMD is governed by:

- a. The accuracy of the navigation system driving it.
- b. The manufacturing tolerances in the electro-mechanical projection system.

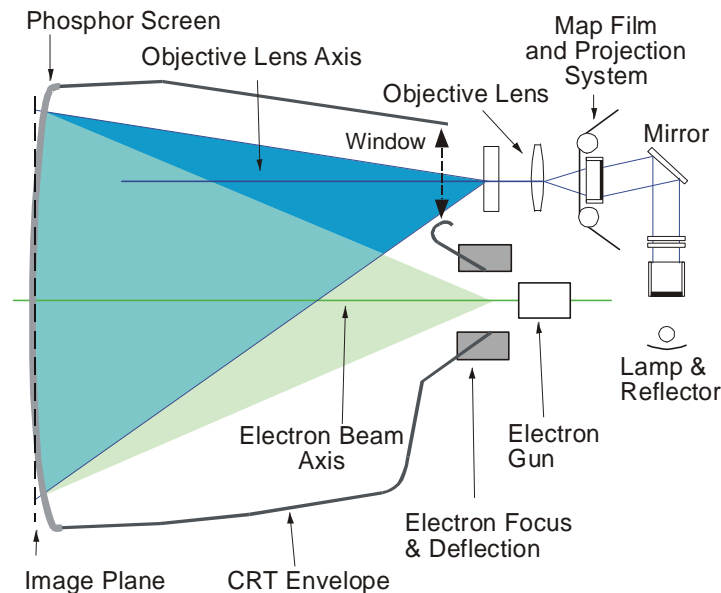
Errors due to map scale and convergency limitations are reduced to relatively insignificant levels by automatically applying a correction to the map drive system, or by applying a correcting distortion during the photographic process. Typical values for the overall accuracy of the system are $\frac{1}{4}$ nm on a 1:500,000 map and 50 metres on a 1:50,000 map.

8. **Limitations of the PMD.** Although reasonable levels of reliability have been achieved with PMDs, the fairly complex electro-mechanical system remains a drawback. The use of many moving parts can lead to mechanical wear and failure. A further disadvantage of the PMD is that the film strip is not normally produced at unit level and cannot be easily amended to reflect late changes, or to incorporate up-to-date tactical information.

9. **Combined Displays.** The PMD can be combined with the display from other electronic systems, principally radar. Two techniques are used:

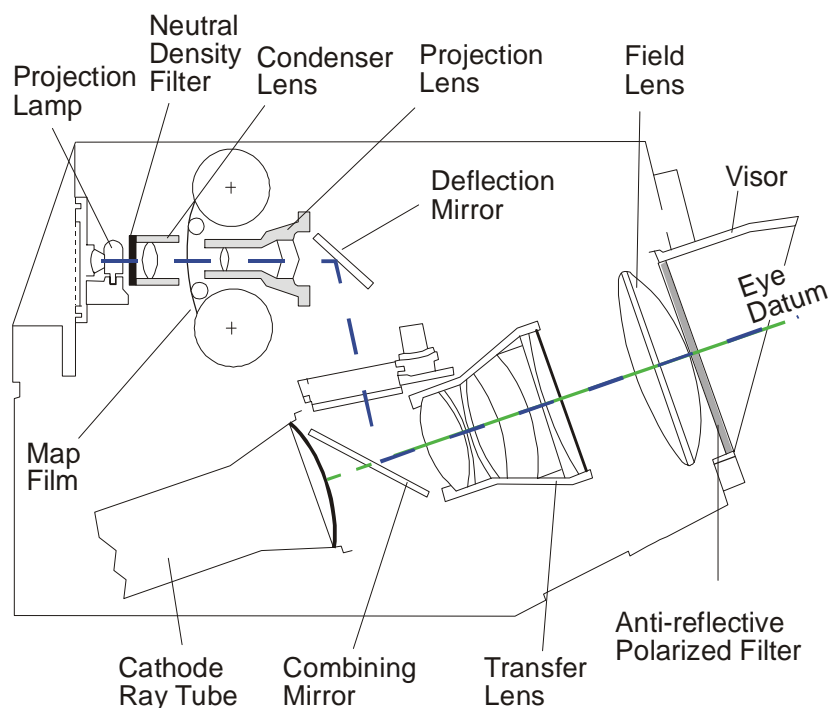
a. **Ported CRT.** The ported CRT is similar to a conventional CRT, but has a transparent window in the envelope. The film image is optically back-projected through the window, on to the CRT's phosphor screen. A simplified diagram is illustrated in Fig 3. The phosphor must be selected so that it can be used both as a back-projection screen and for writing the electron beam. Since the electron gun and the projector cannot both be on the optical axis of the system, one of the images must be distorted to allow them both to be correctly harmonized for simultaneous viewing.

7-30 Fig 3 Combined Display using a Ported CRT



b. **Optically Combined Display.** The optically combined display combines the optical and electronic images using conventional optics and semi-reflecting surfaces. This technique overcomes the distortion problem and makes the phosphor independent of the projection system. The arrangement of an optically combined display is shown in Fig 4.

7-30 Fig 4 Arrangement of an Optically Combined Display



10. Both techniques are capable of over-writing the topographical map with cursive symbology, thus allowing routes, danger areas, and tactical information to be added or amended.

Electronically Displayed Maps

11. Combined display systems overcome some of the problems of keeping the map up to date, but the problems inherent in an electro-mechanical device remain. A preferred solution is to replace the projected map by an electronically derived version in one of two forms:

a. **Digitized (or raster) Maps.** Most of the displayed maps are digitized versions of the familiar paper originals. An electronic scanner is used to scan the paper map in a raster fashion, either as a complete entity or as separate overlays. For example, the air information (including airfields, restricted airspace and obstructions) can be scanned separately from the base topographical map, to facilitate more efficient updating. Digitized maps (which are also referred to as 'raster' maps) have the advantage that the source data is already available, in the form of paper maps, at the scales and coverage required. Disadvantages of digitized maps include:

- (1) They are expensive in terms of digital storage requirements.
- (2) They require complex algorithms to process and reduce the data for display on small CRTs whilst retaining the required detail.

b. **Digital (or vector) Maps.** Digital maps are those that contain separate feature types as data sets in a digital database. Features such as coastlines, contours, railways, roads, woodland, etc are constructed in vector format and combined on the display to provide a usable map. Digital (or vector) maps are more efficient in terms of data storage requirements. They also offer great flexibility in what is displayed on the screen, since unwanted data sets can be de-selected to reduce clutter on the display. Intervisibility techniques allow for line of sight information and radar shadow areas to be displayed in a dynamic fashion as an aircraft moves across the map. The major disadvantage of the digital map concept is that, currently, only limited database production has taken place. Examples of existing digital map products include:

- (1) Products from the US National Imagery and Mapping Agency (formerly the Defence Mapping Agency).
- (2) The eight-nation Digital Land Mass Simulation (DLMS) product. This combines Digital Terrain Elevation Data (DTED) with Digital Feature Analysis Data (DFAD). DTED is used in combination with digitized or digital topographical maps to provide relief information on electronic displays (see also Volume 7, Chapter 15).

12. Both types of maps allow superimposition of latitude/longitude grids and other mission data. Mission data can be produced in a ground-based facility and transferred to the aircraft at initialization, or inserted manually by the aircrew at any stage throughout the mission.

13. Various data storage techniques have been investigated. These include holographic, magnetic, bubble, optical disc and random access memory (RAM). Both optical disc and RAM have been used in UK airborne applications for data storage. Both of these can now provide sufficient data storage for a fast jet ground attack mission, and can be manufactured with adequate ruggedness to survive the harsh airborne environment.