# DESIGN, ENGINEERING & TESTING OF GUIDED ROCKETS

## TABLE OF CONTENTS

# 1 INTRODUCTION

## 1.1 DEFINITION OF A GUIDED WEAPON, TERMS AND CLASSIFICATIONS

What is guided weapon?

*Guided Weapon = Sensors (Eyes & Ears) + Guidance logic (Brain) + Control & Propulsion (Muscles) + Warhead*

> *Definition of a guided weapon: A guided weapon is a weapon that can correct its course to hit a target.*

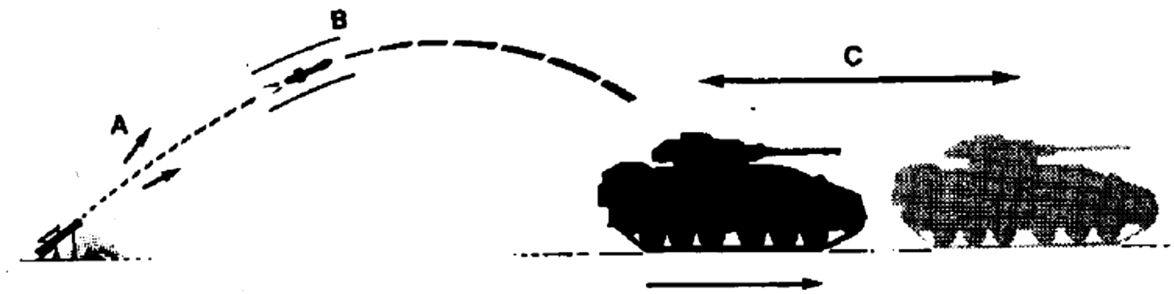


**Figure 1-1 Unguided missile flight – ballistic**

A – Dispersion of initial launching parameters

B – Atmospheric disturbances

C – Moving targets

A, B, C decrease efficiency of missile on the target

> *The main purpose of missile guidance and control is to increase probability of hitting a target with one shoot.*

The main characteristic of any guidance missile system is close loop automatic control. The guidance algorithm will produce guidance signal based on a target and a missile relative position (Figure 1-2).

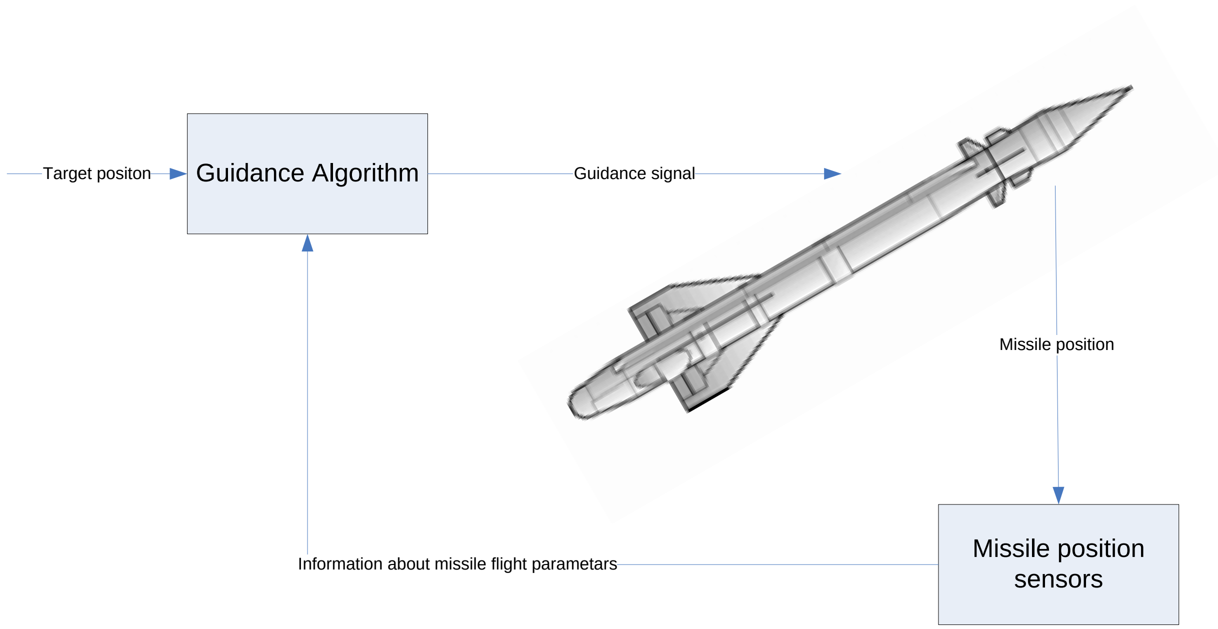


**Figure 1-2 Missile guidance automatic control close loop**

Classification of projectiles:

1. Unguided & unpowered (without propulsion) – bombs, artillery ammunition
2. Unguided & powered (with propulsion) – unguided air missiles, artillery missiles, based-bleed projectiles
3. Guided & unpowered (without propulsion) – smart bombs, smart ammunition
4. Guided & powered (with propulsion) – guided missiles

Classification of guided missiles based launch mode:

1. Surface-to-Surface guided missiles: A surface-to-surface missile is a guided projectile launched from a hand-held, vehicle mounted, trailer mounted or fixed installation. It is often powered by a rocket motor.

    a. Inertial guidance

**Figure 1-3 MGM-52 Lance missile**

**Figure 1-4 9P129 Tochka**

b. Anti-tank missiles


Figure 1-5 Milan anti-tank missile

Figure 1-6 HOT anti-tank missile

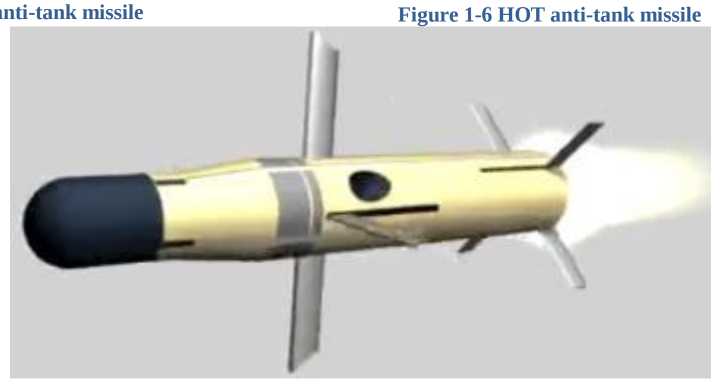

Figure 1-7 TOW anti-tank missile

c. Ship-to-surface

Figure 1-8 EXOCET

2. Surface-to-air missiles: A surface-to-air missile is designed for launch from the ground platform to destroy aerial targets like aircraft, helicopters and ballistic missiles. These missiles are generally called air defense systems.

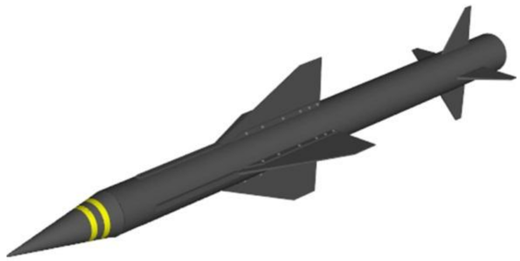

Figure 1-9 RAPIER

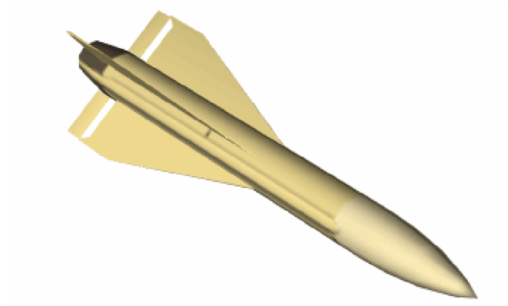

Figure 1-10 HAWK


Figure 1-11 PATRIOT

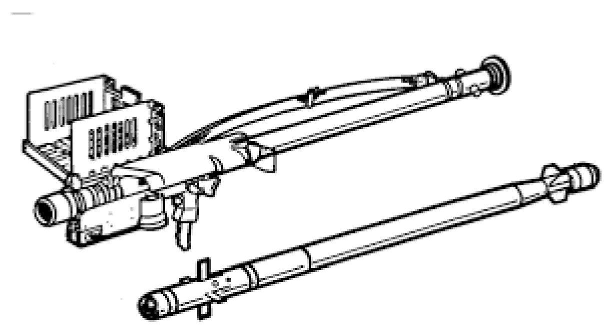

Figure 1-12 STINGER

Figure 1-13 MISTRAL

3. Air-to-surface missiles: An air-to-surface missile is designed for launch from aircraft and strike ground targets on land, at sea or both. Maverick hellfire


Figure 1-14 Maverick

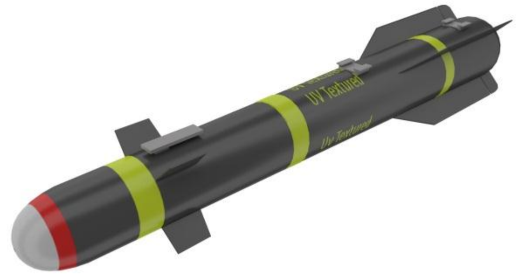


Figure 1-15 Hellfire

4. Air-to-air missiles: An air-to-air missile is designed for launch from an aircraft to destroy the enemy aircraft. The missile flies at speed of over 4 Mach. Sparrow sidewinder

Figure 1-16 Sparrow

Figure 1-17 Sidewinder

5. Submarine launched missiles
    a. Ballistic
    b. Cruise
    c. Anti-ship

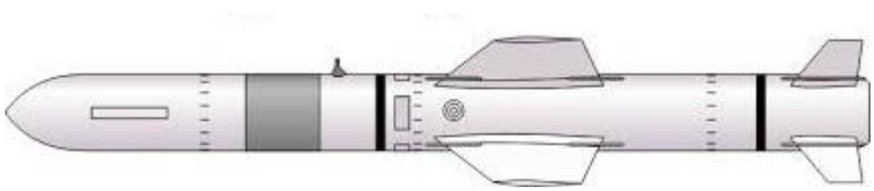

Figure 1-18 Harpoon

## 1.2 DESIGN OF GUIDED MISSILES

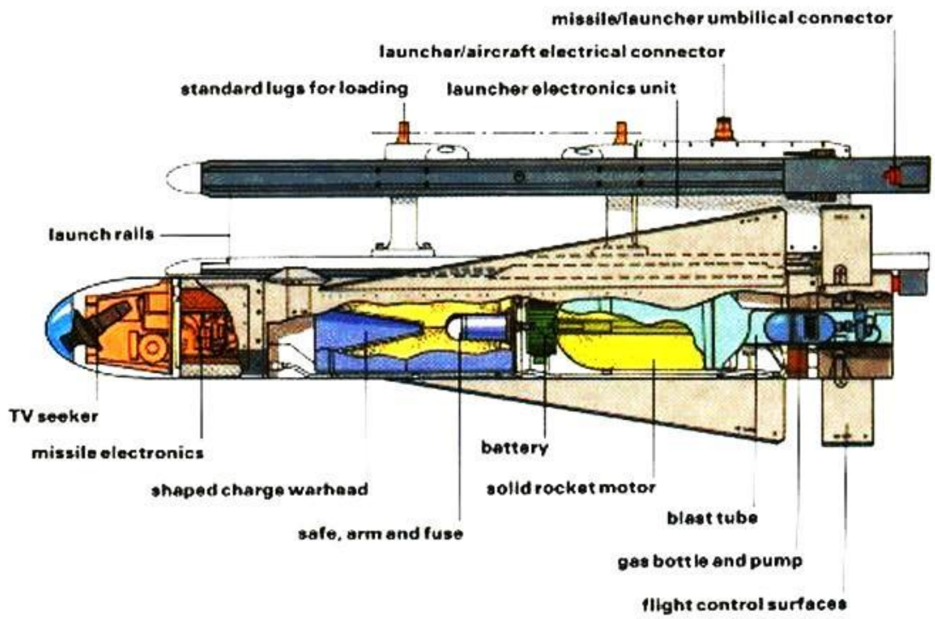


Figure 1-19 Maverick missile subsystems

**Main subsystems of guided missiles:**

1. WARHEAD, which is main purpose of any missile
2. PROPULSION SECTION:
    a. Rocket motor
    b. Jet engine
3. SAFETY ARMING UNIT: Usually SAU is electromechanical subsystem. Main purpose of SAU is to prevent premature detonation of the warhead and to ensure safety near launching place and during transportation.
4. GUIDANCE SISTEM: The most complex and the most expensive subsystem of missile. Key features:
    a. Target detection
    b. Comparison of target and missile position
    c. Calculation of guidance signal which are further sent to control section.

    Design of guided missile usually starts form guidance and control section.

5. CONTROL & STABILISATION SECTION: this section convert guidance signal to control surface deflections, it consists of autopilots and actuators.

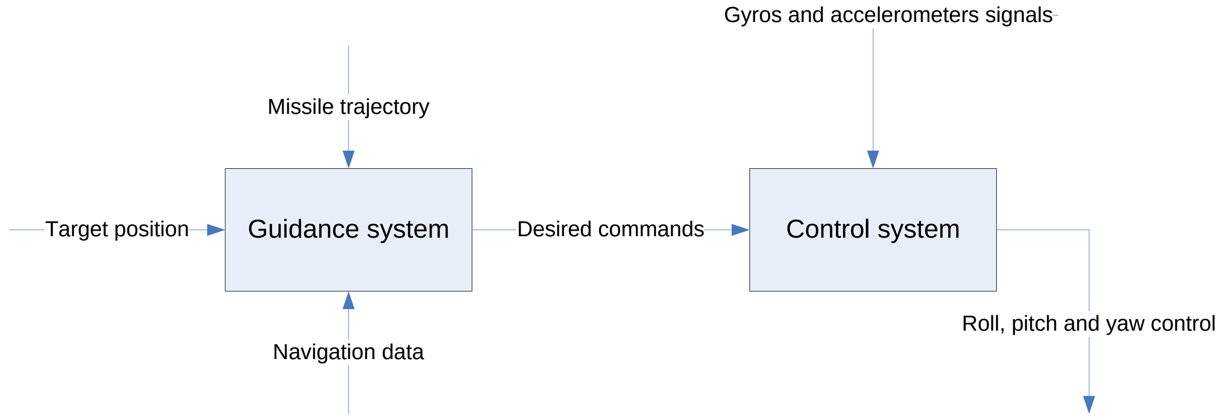


**Figure 1-20 Relationship between guidance and control system**

Missile guidance concerns the method by which the missile receives its commands to move along a certain path to reach a target. On some missiles, these commands are generated internally by the missile computer autopilot. On others, the commands are transmitted to the missile by some external source. The missile sensor or seeker, on the other hand, is a component within a missile that generates data fed into the missile computer. This data is processed by the computer and used to generate guidance commands. Sensor types commonly used today include infrared, radar, and the global positioning system. Based on the relative position between the missile and the target at any given point in flight, the computer autopilot sends commands to the control surfaces to adjust the missile's course.

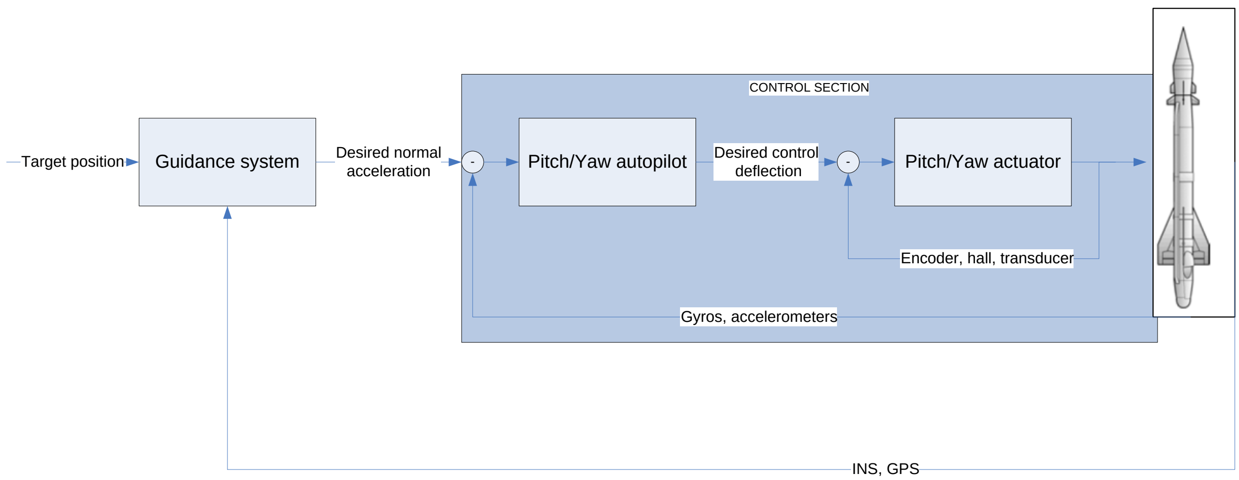


**Figure 1-21 Guidance and control closed loops**

## 1.3 BASIC GUIDANCE LAWS

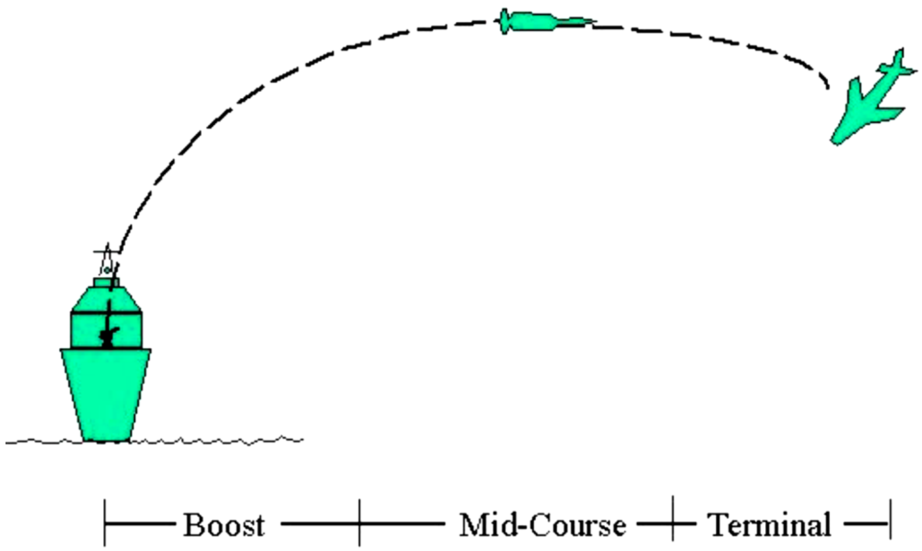


**Figure 1-22 Phases of missile guidance**

In many missiles, the guidance system is divided into three phases, as pictured above. The first is a launch or boost phase in which the guidance system is usually disabled to allow the missile to safely travel away from the launch platform. The majority of the flight is flown using midcourse guidance, during which the missile makes slight adjustments to its trajectory allowing it to reach the vicinity of the target. The final phase is terminal guidance when the missile uses a highly accurate tracking system to make rapid maneuvers for intercepting the target. Many missiles use a different type of guidance in the midcourse phase than in the terminal phase.

### LINE OF SIGHT (LOS) GUIDANCE

> ***In LOS guidance the missile follows the line of sight (LOS) from an external tracker to the target***

### Beam Rider Guidance

The beam rider concept relies on an external ground- or ship-based radar station that transmits a beam of radar energy towards the target. The surface radar tracks the target and also transmits a guidance beam that adjusts its angle as the target moves across the sky.

The missile is launched into this guidance beam and uses it for direction. Scanning systems onboard the missile detects the presence of the beam and determine how close the missile is to the edges of it. This information is used to send command signals to control surfaces to keep the missile within the beam. In this way, the missile "rides" the external radar beam to the target.

Beam riding was often used on early surface-to-air missiles but was found to become inaccurate at long ranges. Limited improvement was possible using two different surface-based radar beams, but the beam rider method has been largely abandoned.

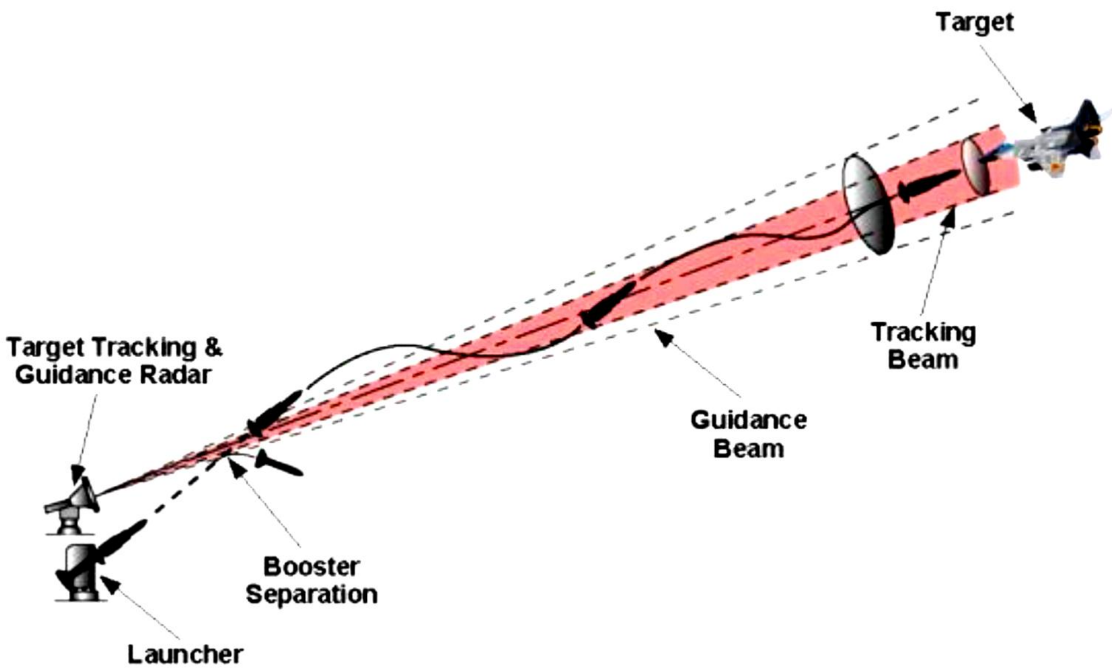


**Figure 1-23 Beam rider guidance**

### Command Guidance

Command guidance is similar to beam riding in that the target is tracked by external radar. However, second radar also tracks the missile itself. The tracking data from both radars are fed into a ground based computer that calculates the paths of the two vehicles.

This computer also determines what commands need to be sent to the missile control surfaces to steer the missile on an intercept course with the target. These commands are transmitted to a receiver on the missile allowing the missile to adjust its course. An example of command guidance is the Russian SA-2 surface-to-air missile used against US aircraft in North Vietnam. Also note that command guidance is not limited just to radar. Another method that falls under command guidance is the use of wire guided systems. In this technique, commands are sent to the missile through a conventional wire or fiber optic cable that reels out from the missile back to its launcher. Wire guidance is often used on anti-tank missiles like TOW, which can be launched from both ground vehicles and helicopters. Many naval torpedoes fired from submarines also use wire guidance.

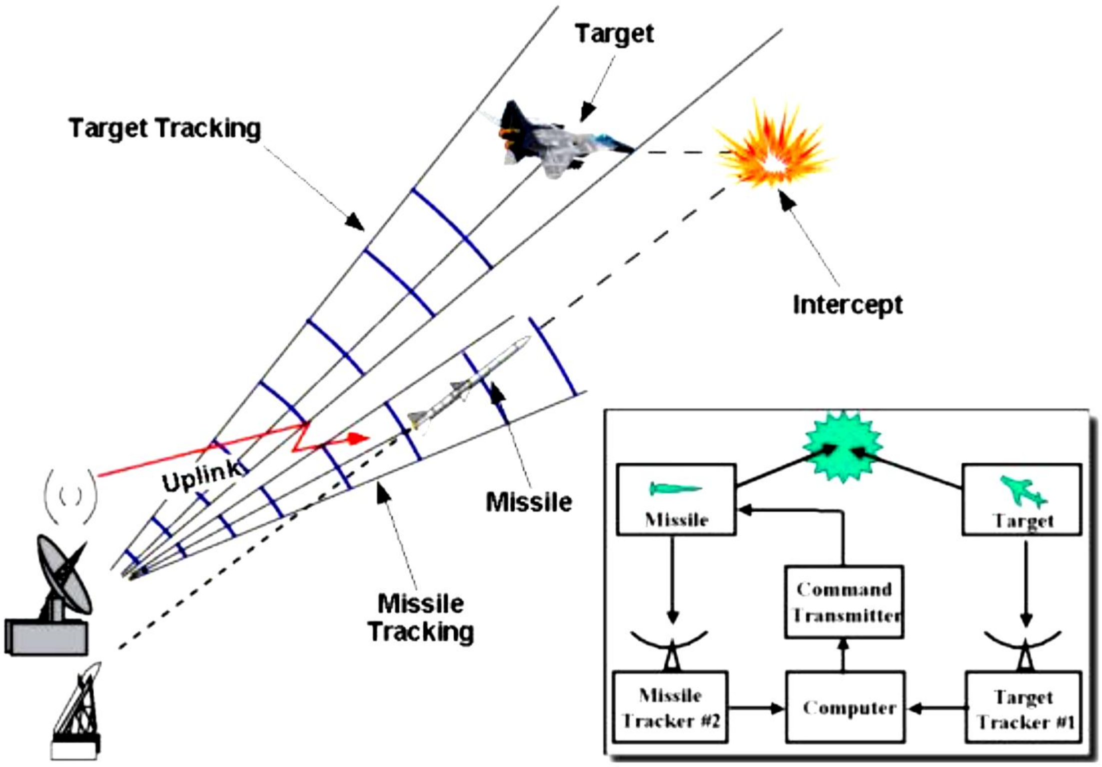


**Figure 1-24 Command guidance**

## HOMING GUIDANCE

> ***In homing guidance the missile steers itself towards a target using an onboard seeker which is able to detect some distinguishing characteristics of the target.***

Homing guidance is the most common form of guidance used in anti-air missiles today. Three primary forms of guidance fall under the homing guidance -semi active, active, and passive.

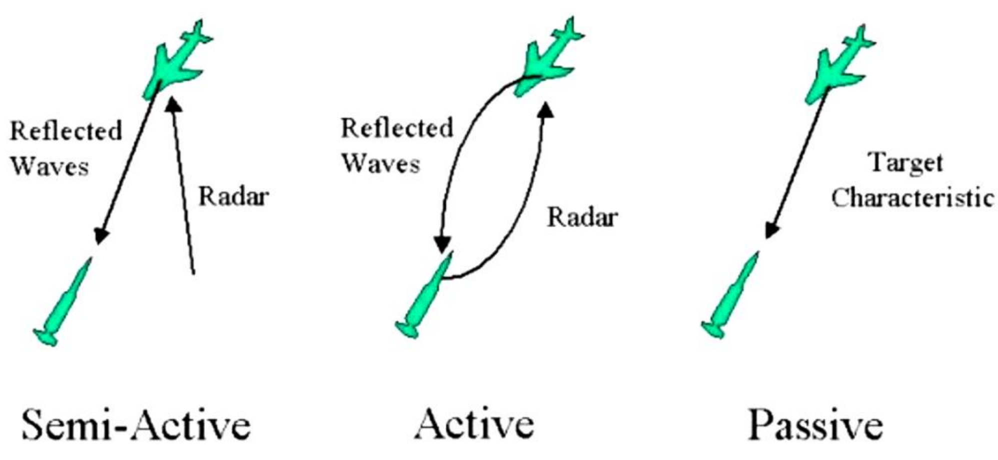


**Figure 1-25 Homing guidance**

### Semi-Active Homing Guidance

A semi-active system is similar to command guidance since the missile relies on an external source to illuminate the target. The energy reflected by this target is intercepted by a receiver on the missile. The difference between command guidance and semi-active homing is that the missile has an onboard computer in this case. The computer uses the energy collected by its radar receiver to determine the target's relative trajectory and send correcting commands to control surfaces so that the missile will intercept the target.

The example shown on figure below illustrates the guidance method used on an air-to-air missile like Sparrow. This missile relies on radar energy transmitted by the launch aircraft to track and home in on the target. However, it should be noted that semi-active guidance is used by other types of seekers besides radar. Laser-guided weapons like the Paveway series can also be considered semi-active weapons because the laser energy these bombs track as they steer towards a target is supplied by an external source. The source could be a laser designation pod on the launch aircraft, on a second aircraft, or aimed by a soldier on the ground.

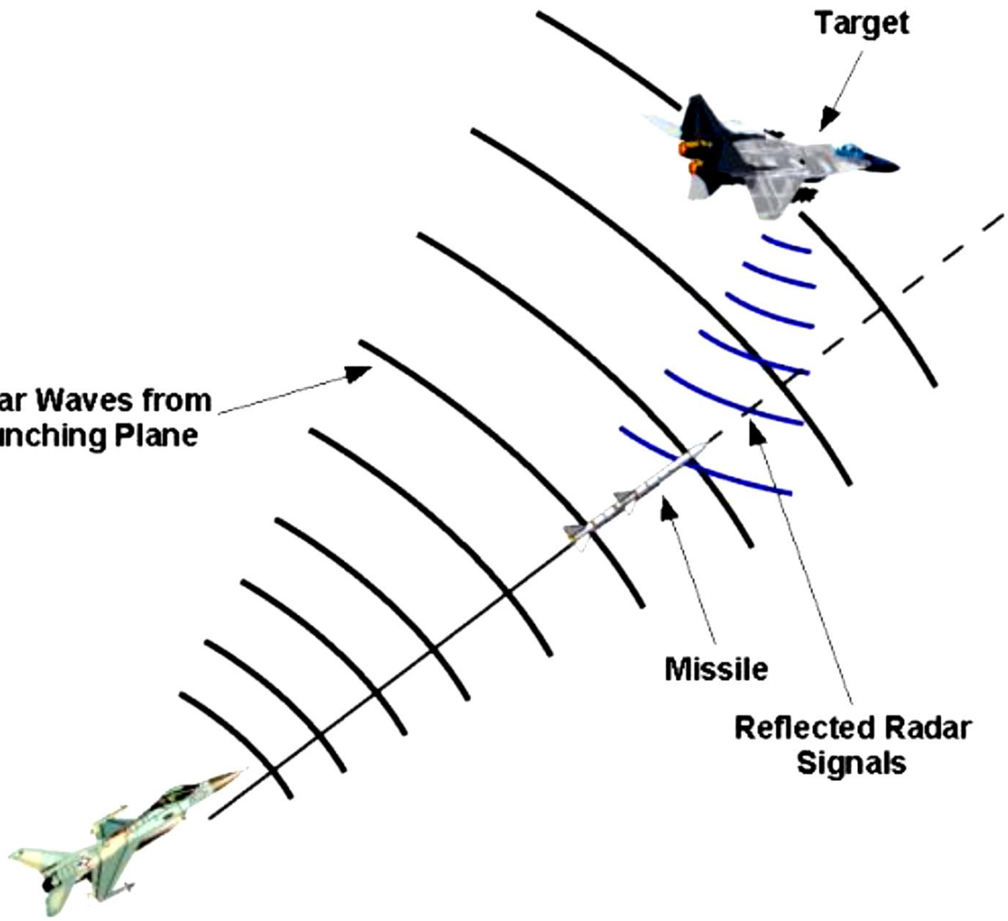


**Figure 1-26 Semi-active homing guidance**

### Active Homing Guidance

Active homing works just like semi-active except that the tracking energy is now both transmitted by and received by the missile itself. No external source is needed. It is for this reason that active homing missiles are often called "fire-and-forget" because the launch aircraft does not need to continue illuminating the target after the missile is launched.

Active homing missiles typically use radar seekers to track their target. These seekers are also sometimes called monostatic because, unlike semi-active guidance, the transmitted and reflected waves are at the same angle with respect to the line of sight between the missile and target. Examples of active homing missiles include the AMRAAM air-to-air and Exocet anti-ship missiles.

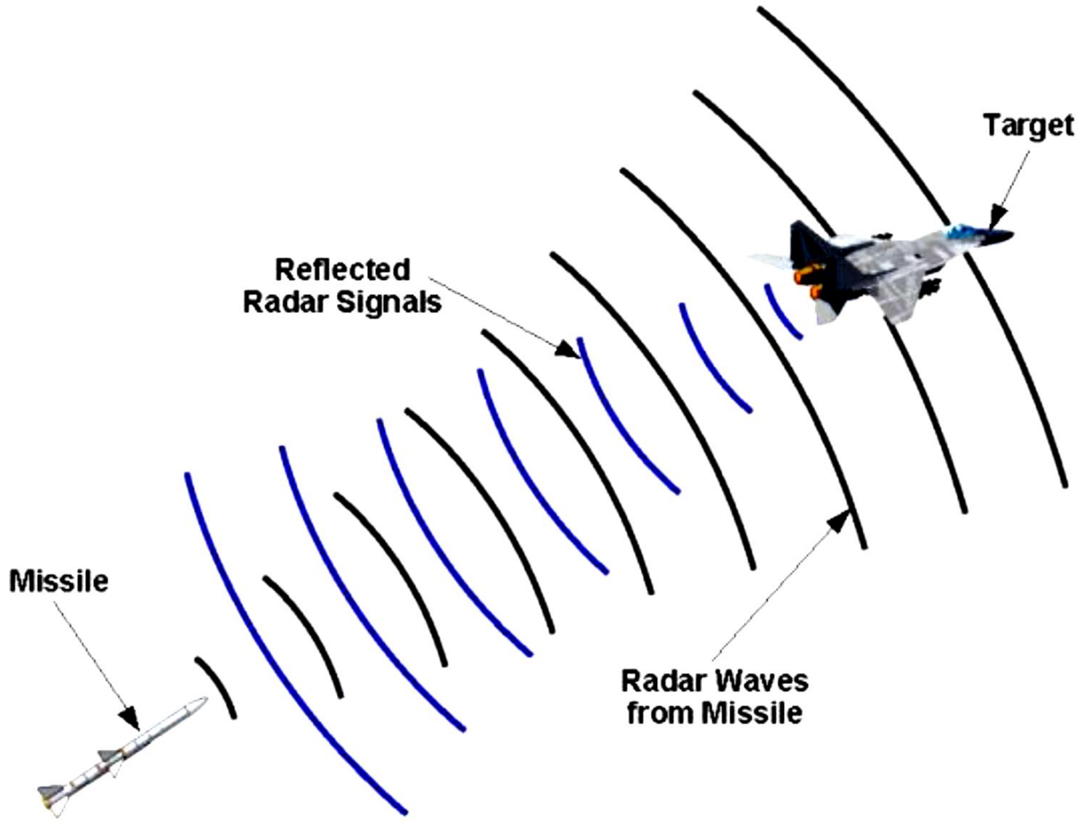


**Figure 1-27 Active homing guidance**

### Passive Homing Guidance

A passive homing system is like active in that the missile is independent of any external guidance system and like semi-active in that it only receives signals and cannot transmit. Passive missiles instead rely on some form of energy that is transmitted by the target and can be tracked by the missile seeker.

This energy could take many forms. For example, infrared seekers like those used on Sidewinder home in on the heat signature generated by a target. Anti-radiation missiles like HARM track the radio frequency energy transmitted by ground-based radar stations. Passive torpedoes use sonar, or sound waves, created by the engines of ships to attack their targets. Electro-optic sensors like those used on Maverick rely on visual images to guide towards a target.

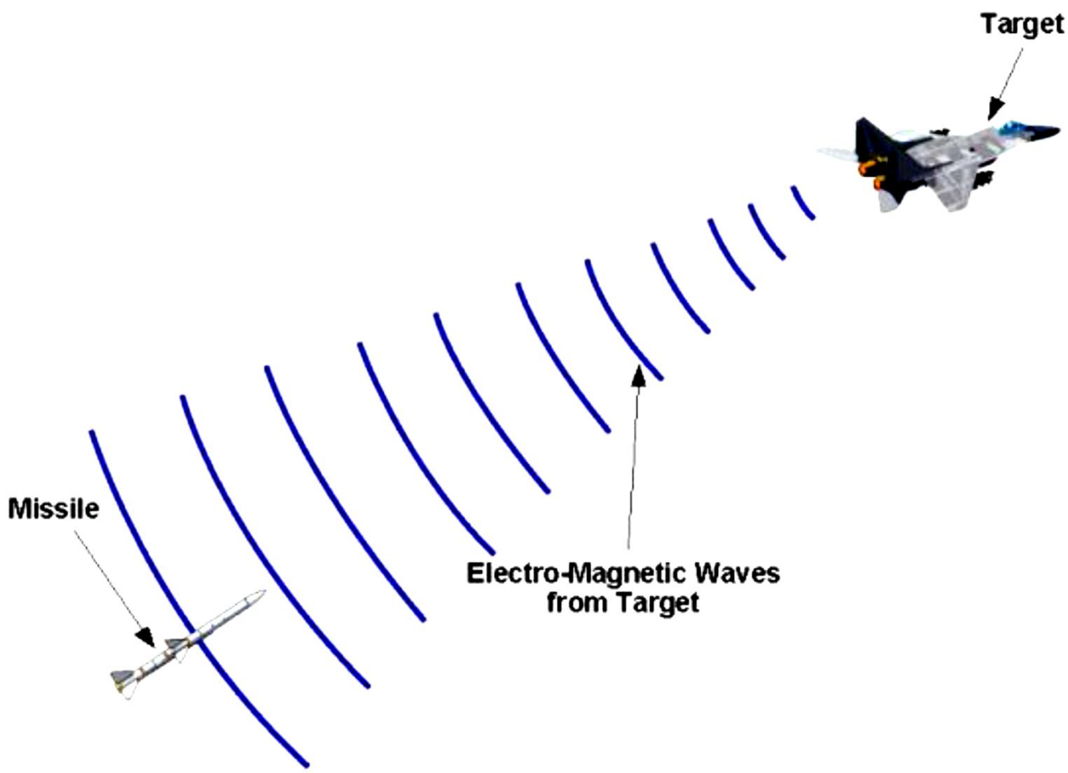


**Figure 1-28 Passive homing guidance**

## NAVIGATION GUIDANCE

Like homing guidance, navigation guidance includes several subcategories. In this section, it will be describe inertial, ranging, and geophysical navigation techniques.

### Inertial Navigation Guidance

> In inertial guidance, the vehicle's position, velocity and attitude are derived from gyroscope and accelerometer readings.

Inertial navigation relies on devices onboard the missile that senses its motion and acceleration in different directions. These devices are called gyroscopes and accelerometers.

The purpose of a gyroscope is to measure angular rotation, and a number of different methods to do so have been devised. A classic mechanical gyroscope senses the stability of a mass rotating on gimbals. More recent ring laser gyros and fiber optic gyros are based on the interference between laser beams. Current advances in Micro-Electro-Mechanical Systems (MEMS) offer the potential to develop gyroscopes that are very small and inexpensive.

While gyroscopes measure angular motion, accelerometers measure linear motion. The accelerations from these devices are translated into electrical signals for processing by the missile computer autopilot. When a gyroscope and an accelerometer are combined into a single device along with a control mechanism, it is called an inertial measurement unit (IMU) or inertial navigation system (INS).

The INS uses these two devices to sense motion relative to a point of origin. Inertial navigation works by telling the missile where it is at the time of launch and how it should move in terms of both distance and rotation over the course of its flight. The missile computer uses signals from the INS to measure these motions and insure that the missile travels along its proper programmed path. Inertial navigation systems are widely used on all kinds of aerospace vehicles, including weapons, military aircraft, commercial airliners, and spacecraft. Many missiles use inertial methods for midcourse guidance, including AMRAAM, Storm Shadow, Meteor, and Tomahawk.

### Ranging Navigation Guidance

Unlike inertial navigation, which is contained entirely onboard the vehicle; ranging navigation depends on external signals for guidance. The earliest form of such navigation was the use of radio beacons developed primarily for commercial air service. These beacons transmit radio signals received by an aircraft in flight. Based on the direction and strength of the signals, the plane can calculate its location relative to the beacons and navigate its way through the signals.

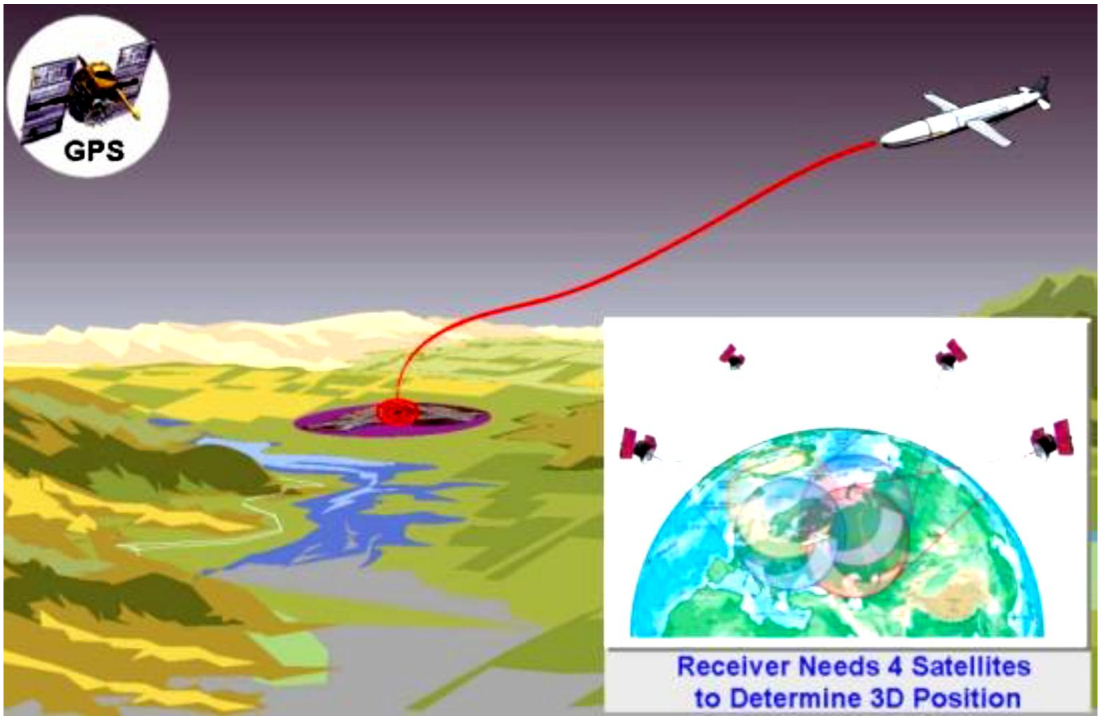


**Figure 1-29 GPS used in ranging navigation guidance**

The advent of the global positions system (GPS) has largely replaced radio beacons in both military and civilian use. GPS consists of a constellation of 24 satellites in geosynchronous orbit around the Earth. If a GPS receiver on the surface of the Earth can receive signals from at least four of these satellites, it can calculate an exact three-dimensional position with great accuracy. Missiles like JSOW and the JDAM series of guided bombs make use of GPS signals to determine where they are with respect to the locations of their targets. Over the course of its flight, the weapon uses this information to send commands to control surfaces and adjust its trajectory.

### Geophysical Navigation Guidance

A technique mostly used is called digital scene matching. In concept, digital scene matching is little different than looking out the window of your car and using landmarks to navigate your way to a specific location. Missiles make use of this technique by comparing the image seen below the weapon to satellite or aerial photos stored in the missile computer. If the scenes do not match, the computer sends commands to control surfaces to adjust the missile's course until the images agree. Digital scene matching is used on the Tomahawk cruise missile.

## 1.4 TRAJECTORIES PERFORMANCES

Guided trajectory types:

1. straight-line trajectory (not common but it can be part of cruise missile trajectory)


**Figure 1-30 Straight-line trajectory**

2. proportional navigation (PN) & augmented proportional navigation (APN) (homing guidance)

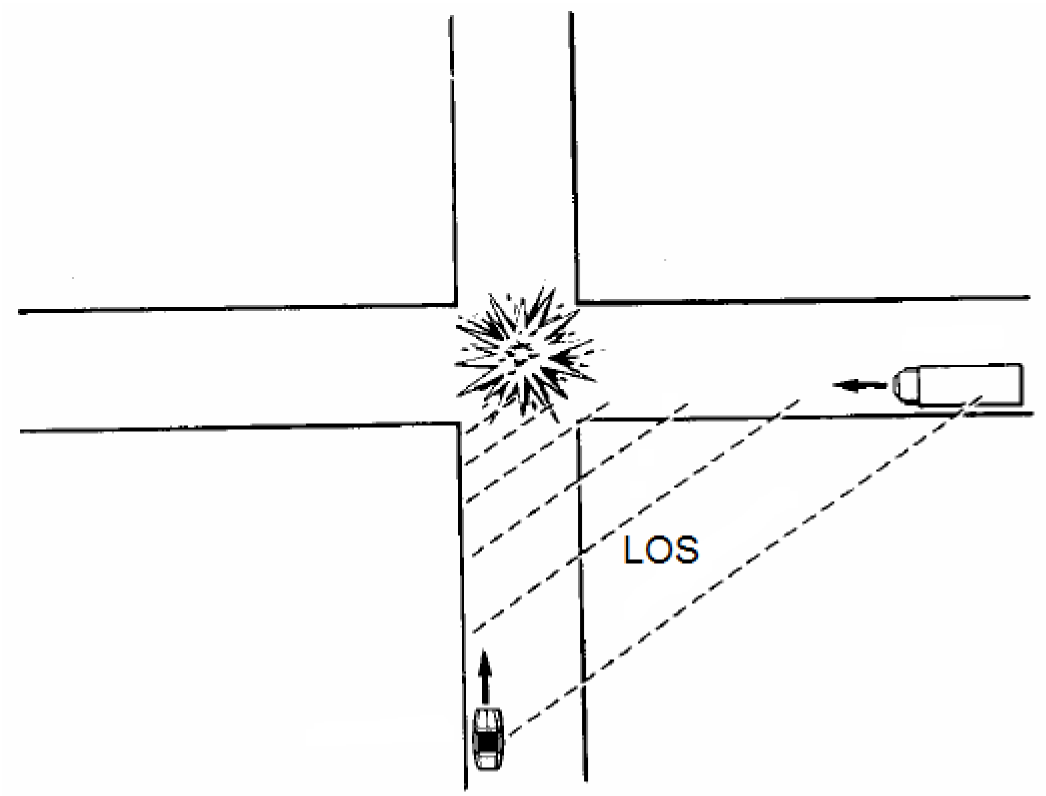


**Figure 1-31 PN principle**

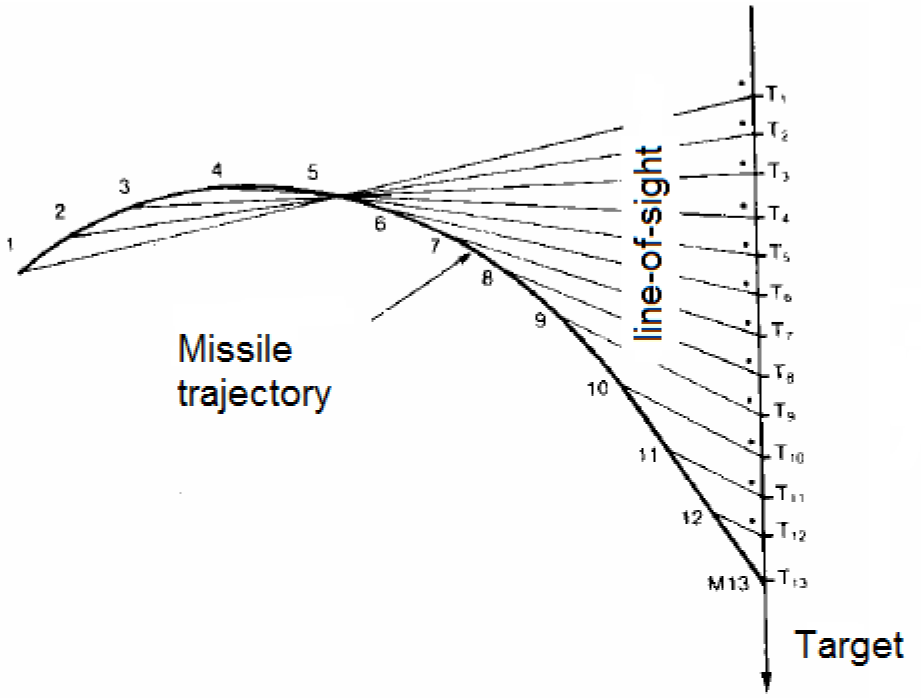


**Figure 1-32 PN kinematic**

*Proportional navigation guidance law:*

$$f = K\dot{\phi}$$

*$f$ – Normal acceleration*

*$\dot{\phi}$ – LOS rate*

*$K$ – PN constant*

3. three point guidance, pursuit guidance (LOS & homing guidance)

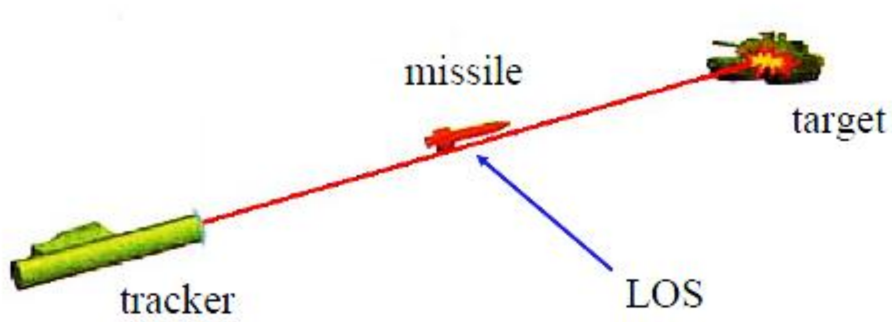


**Figure 1-33 Three point guidance**

*Flight path of the missile which flies directly toward the target is described as PURSUIT GUIDANCE*

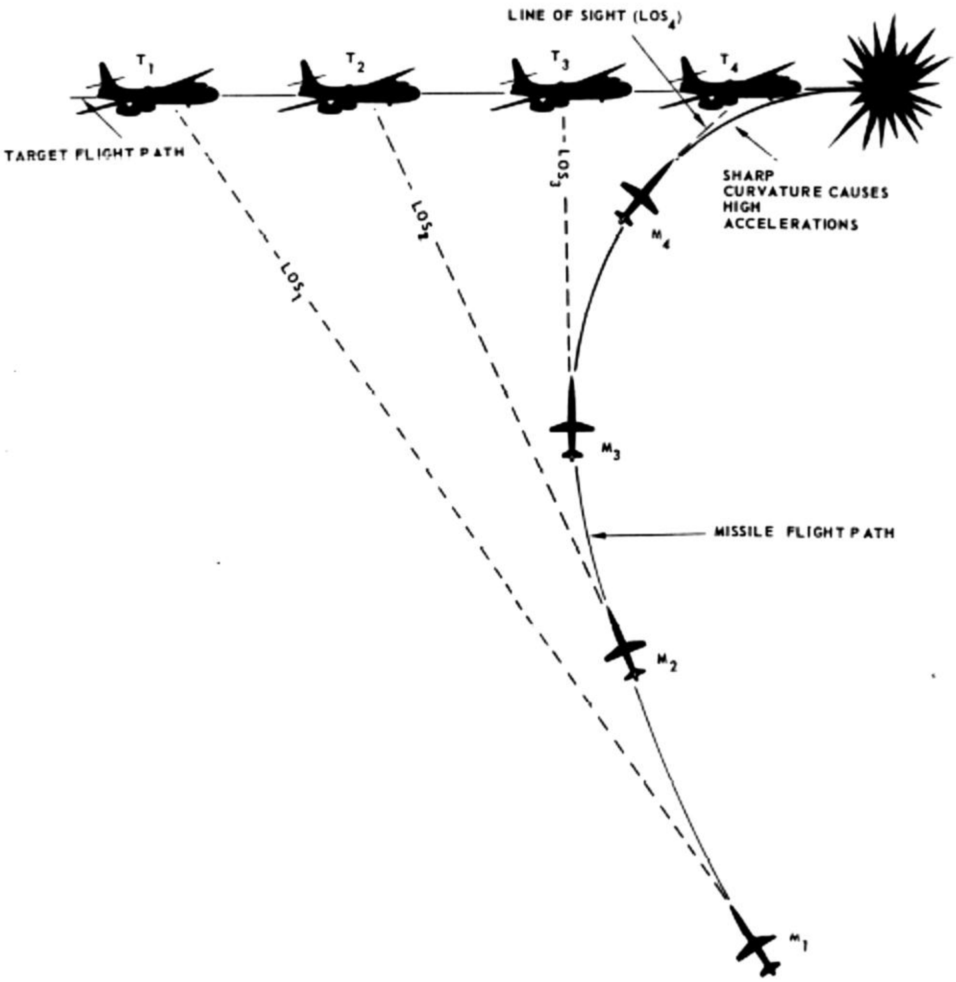


**Figure 1-34 Pursuit guidance kinematics**

4. programed guidance with corrections in control points (cruise missiles)

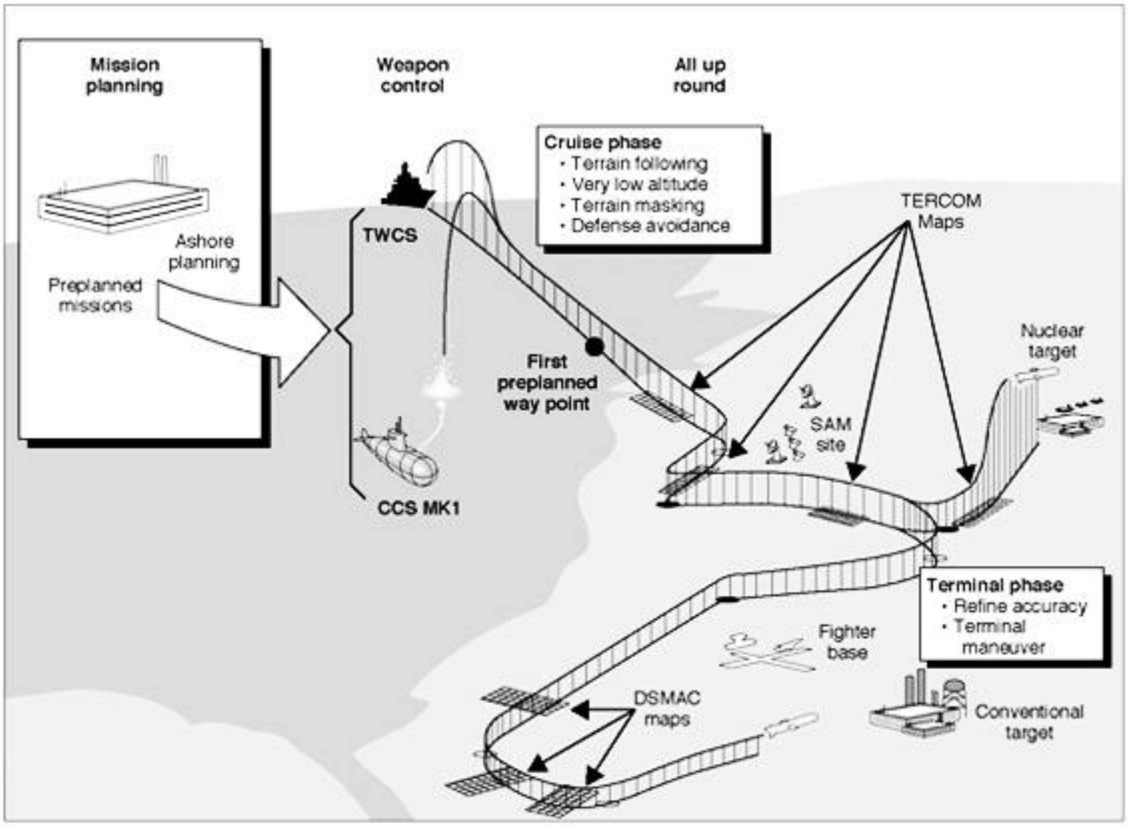


Figure 1-35 Cruise missile trajectory

5. ballistic trajectory (inertial guidance)

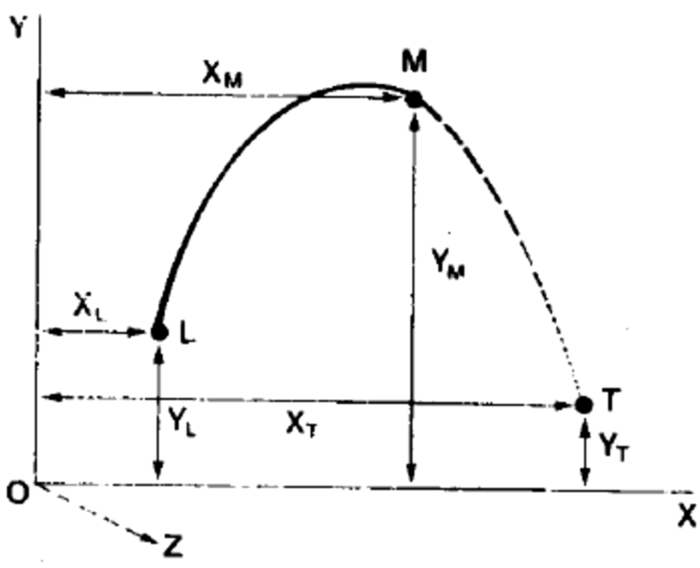


Figure 1-36 Ballistic trajectory

# 2 MISSILE CONTROL METHODS

*Main task of missile control system:*

*Achieved desired trajectory using deflections of control surfaces*

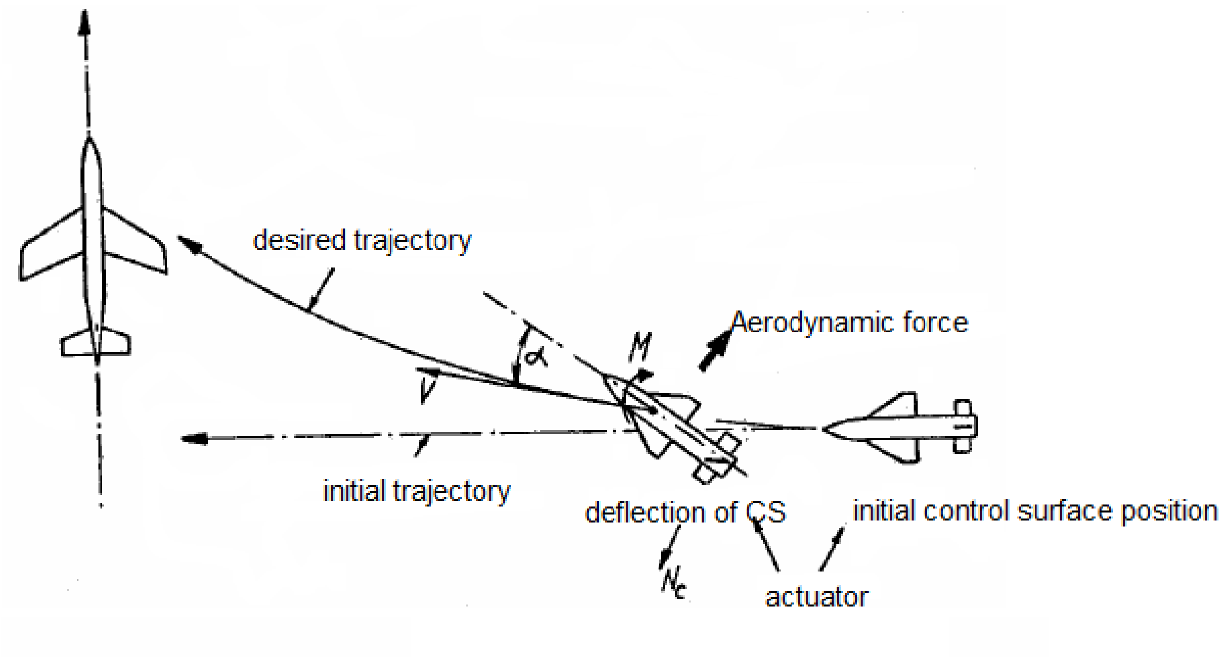


**Figure 2-1 Control method principle**

To achieve desired aerodynamic force, it is necessary to:

- Make deflections of control surfaces which will produce aerodynamic force and aerodynamic torque
- Pitch missile at the angle of attack that will produce aerodynamic force

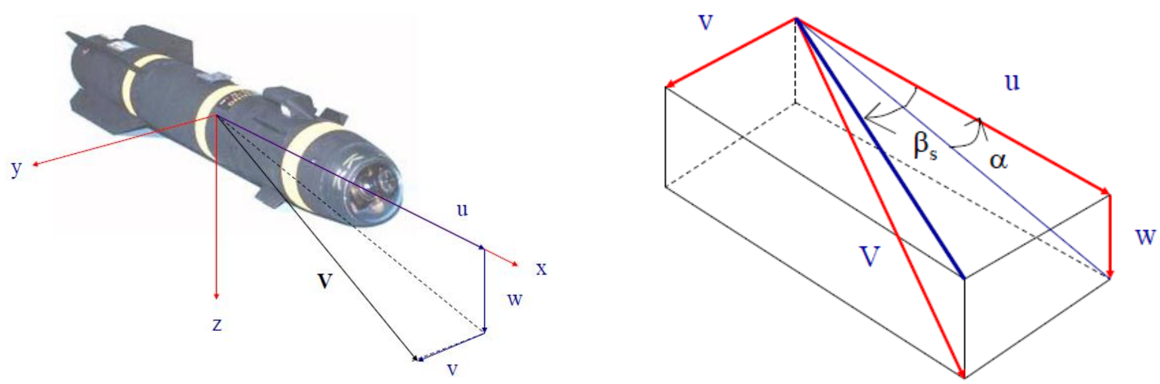


**Figure 2-2 Velocity components, attack angle and sideslip angle**

System for stabilization and control consists of:

- Autopilot (gyros, accelerometers…)
- Actuator
- Control surfaces

### *REQUIREMENTS FOR GUIDANCE, CONTROL AND NAVIGATION SYSTEMS*

- Achieve desired command by intensity and direction
- Achieve desired roll angel or roll rate

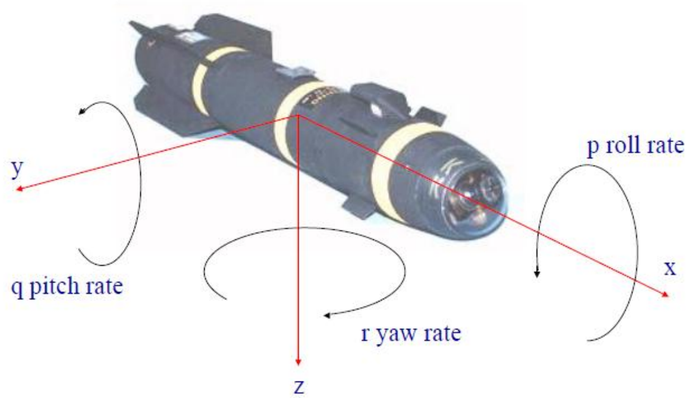


**Figure 2-3 Angular velocity components**

- Compensate disturbance forces
- Autopilot input signal need to be time consistent with guidance signals
- Achieve linear characteristic: control deflection – force
- Minimize hinge moment

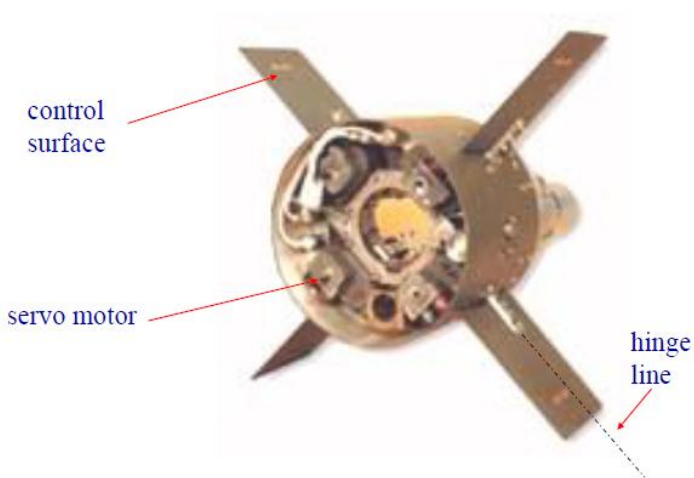


**Figure 2-4 Missile aerodynamic control surfaces**

- Minimize drag force due to deflection

## 2.1 CARTESIAN AND POLAR CONTROL METHODS, CONTROL METHODS CLASIFICATION & DEFINITIONS

### *CARTESIAN CONTROL*

- Separate sets of control surfaces for pitch (up/down) and yaw (left/right)
- Guidance generates required latax for pitch and yaw planes
- Pitch and yaw controls can act simultaneously

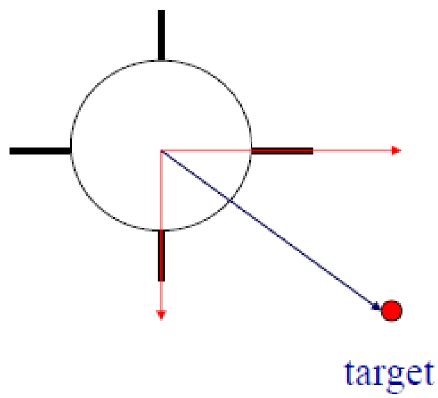


**Figure 2-5 Cartesian control**

### *POLAR CONTROL*

- Guidance generates roll command and required pitch latax (twist & steer)
- Needs roll angle reference

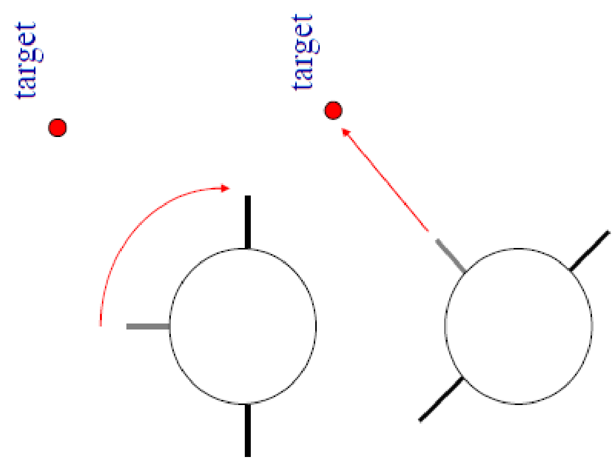


**Figure 2-6 Polar control**

### *CONTROL METHODS CLASSIFICATION*

1. Cartesian control
    a. Aerodynamic control
        i. Roll control
            1. Gyro-stabilized missile
            2. Roll rate stabilization
            3. Roll angel stabilization
        ii. Pitch/Yaw control
            1. Tail control
            2. Canard control
            3. Wing control
    b. Thrust vector control
        i. Gimbaled engine(s) or nozzle(s)
        ii. Reactive fluid injection
        iii. Auxiliary engines (fixed or movable)
        iv. Exhaust vanes
2. Polar control – Aerodynamic
    a. Tail control
    b. Canard control
    c. Wing control

### *CONTROL METHOD DEFINITIONS*

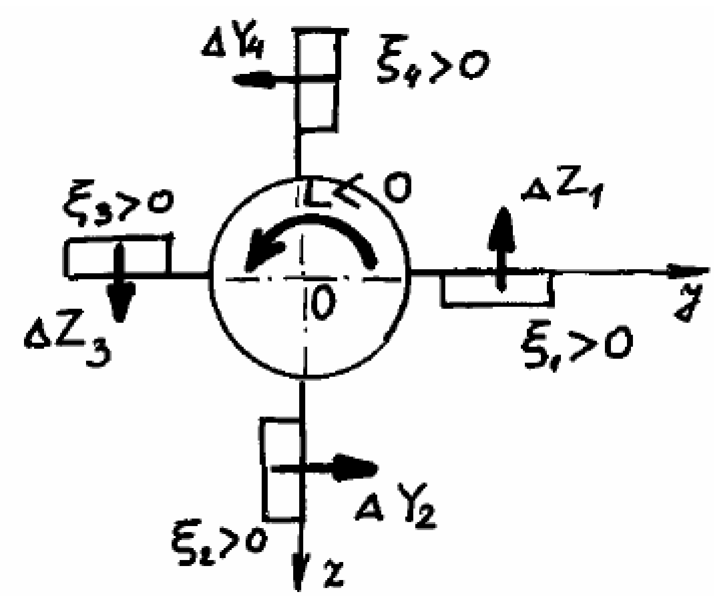


**Figure 2-7 Definition of positive deflections (view in flight direction)**

*Roll command*

$$\xi = \frac{1}{4}(\xi_1 + \xi_2 + \xi_3 + \xi_4) \;\; \xi > 0 \Longrightarrow L < 0$$

*Pitch command*

$$\eta = \frac{1}{2}(\xi_1 - \xi_3) \;\; \eta > 0 \Longrightarrow Z < 0$$

*Yaw command*

$$\zeta = \frac{1}{2}(\xi_2 - \xi_4) \;\; \zeta > 0 \Longrightarrow Y < 0$$

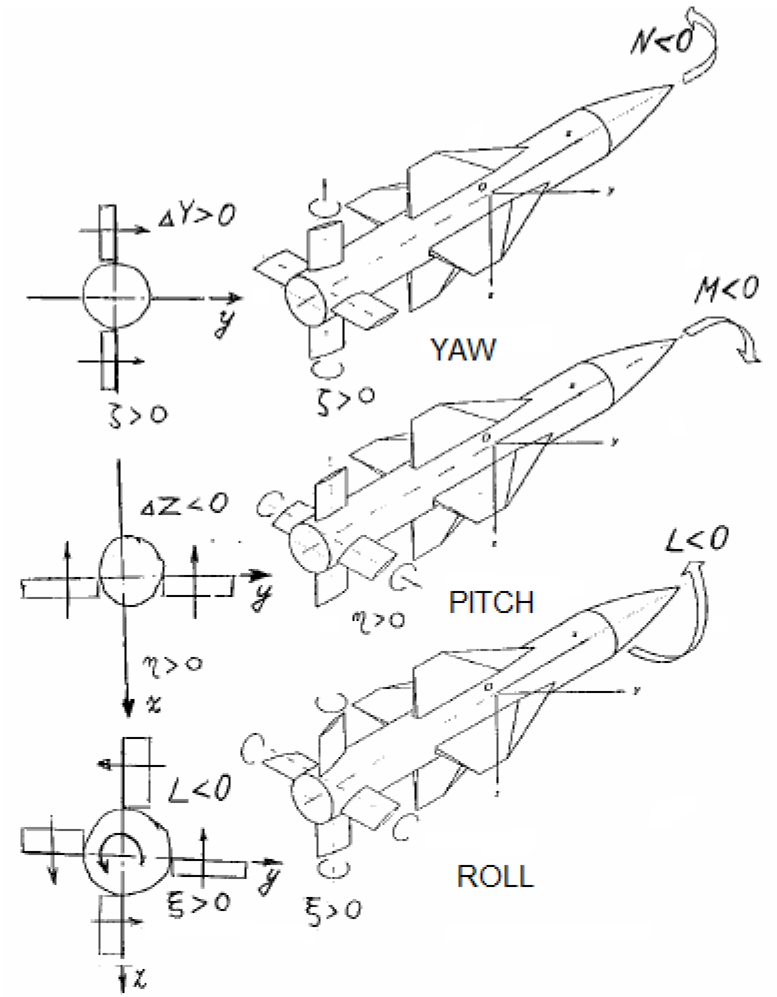


**Figure 2-8 Aerodynamic control torques**

## 2.1.1 TAIL CONTROL

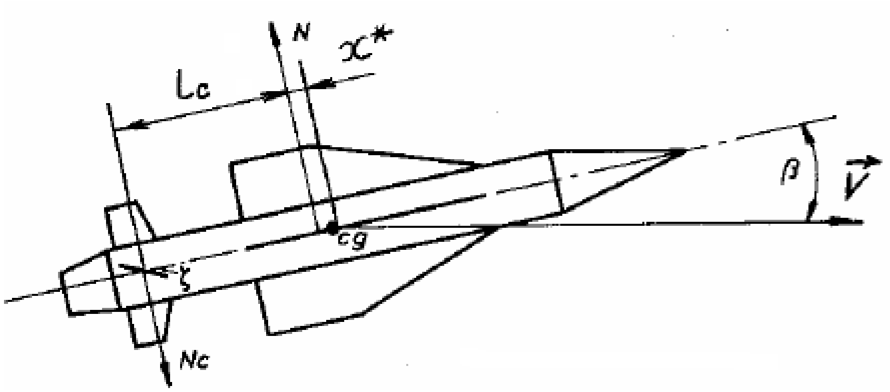


**Figure 2-9 Tail control configuration**

- Allows uniform component mass distribution
- Unfavorable installation
- Most stable in supersonic flight (minimal roll problems from downwash from fixed surfaces)
- Lowest induced drag from control surfaces
- Highest angle of attack without control surfaces stalling
- Slower response due to reduced lift from tails

$$N_c l_c = N x^*$$

$x^*$ - Static stability margin

$l_c$ – The distance from control force point to CG

EXAMPLE:

$$\frac{l_c}{x^*} = 10 \implies N = 10N_c, \qquad N_{tot} = N - N_c = 9N_c \implies N_{tot} < N$$

## 2.1.2 CANARD CONTROL

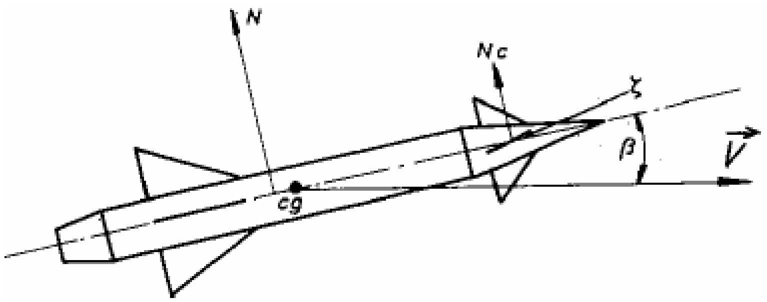


**Figure 2-10 Canard control configuration**

- Produces largest moment for given surface area
- Highest efficiency
- Canards can stall in relatively low angle of attack
- Downwash from canards complicates roll control
- Favorable installation

EXAMPLE:

$$\frac{l_c}{x^*} = 10 \implies N = 10N_c, \qquad N_{tot} = N + N_c = 11N_c \implies N_{tot} > N$$

## 2.1.3 WING CONTROL

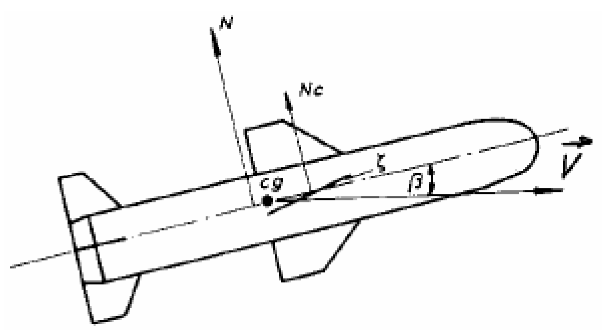


**Figure 2-11 Wing control configuration**

- Minimizes angle of attack, reducing body drag and relaxing seeker FOV requirements
- Direct control of angle of attack
- Wings must be large, increasing drag, actuator size, power consumption, missile mass and volume requirements

## 2.2 ACTUATOR SYSTEM DESIGN

### *ACTUATOR REQUIREMENTS*

Anything other than the warhead and fuze of a missile is dead weight. Minimum weight and volume are almost over-riding requirements for any missile servo. Good shelf life, low cost and reliability are obvious requirements, and a facility to test is an advantage although it can be argued that if reliability is excellent then a requirement to test, which involves time, training and test equipment, should be dispensed with. Now a tactical guided missile can weigh anything from 10 kg to 1000 kg with flight times of 5 to 100 seconds or more. Some systems are liable to experience very noisy guidance signals while others may have a comparatively quiet ride. It is this diversity in size, duration of flight and type of duty cycle that has resulted in a number of successful designs of servos to operate control surfaces or thrust vector elements. Now all missile servos incorporate mechanical stops to limit the angular travel to a safe limit, typically ± 15°-25°. Hence, maximum fin rate will always be defined. The load inertia will be quoted of course in addition to the maximum aerodynamic hinge moment due to the control surface center of pressure not always coinciding with the axis of rotation. Some thrust vector methods involve considerable coulomb friction due to seals.

### 2.2.1 PNEUMATIC ACTUATORS

### *STORED COLD GAS SERVOS*

The advantage of using a fluid operated motor compared with an electric device to produce a force or torque is the much higher pressures that can be utilized with high pressure gas or hydraulic fluid. In a stored gas system working pressures are typically about  and this is much higher than the effective pressure obtainable in any electric motor; this makes for a small final control element or "actuator".

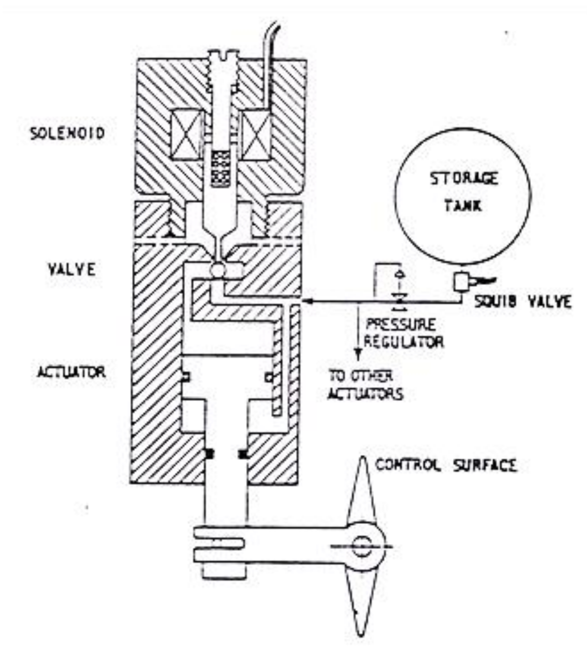


**Figure 2-12 Stored cold gas servos**

Figure shows the main features of a pneumatic fin servo; the error signal is obtained by summing the input demand voltage with the negative feedback from a potentiometer. Compensation and amplification of the error signal are also performed electrically. Actuation power is obtained from stored gas which is released just before firing by means of a solenoid-operated start valve. A reducing valve regulates the pressure to a lower value such that the servo supply pressure remains approximately constant until the bottle pressure is less than the reduced pressure.

The system illustrated represents a gas operated servo in its simplest form. The double acting actuator is the "area-half-area" type and operates in the on-off mode. Conversely if the compensated error signal is converted into a pulse-width modulated signal the servo operates as a linear system with a superimposed high frequency carrier equal to the modulating frequency. If the compensated error signal is simply used to operate the solenoid when this signal is positive and not to operate it when it is negative then a limit cycle will result.

A small high frequency limit cycle is not a bad thing in a servo where the coulomb friction is appreciable as it tends to improve the accuracy. Helium or nitrogen is stored at , and an electro-explosive cutter which ruptures a sealing membrane is a standard technique. Gas containers are subjected to a weight check and mass spectrometer leakage test. Such a system is attractive from a long term storage point of view and the reliability is excellent because the fluid is clean. The electronics and valve design are simple and cheap to manufacture.

If however the load inertia is considerable, the bandwidth obtainable with a pneumatic servo is limited; the dynamic lag from valve position to load speed is a second order one due to the compressibility of the gas and the pneumatic natural frequency decreases with an increase in load inertia. The pneumatic open loop frequency can be increased by increasing the gas supply pressure or by increasing the diameter of the actuator. An important design factor in any high performance pneumatic servo is an amplifier specifically designed to reduce the electrical time constant associated with the inductive solenoid winding.

Pneumatic servos are usually associated with small missiles and fairly short times of flight; the weight of fuel plus bottle tends to be prohibitive if the total energy demand is large.

### *HOT GAS SERVOS*

The advantage of burning cordite or some other mono-fuel such as isopropyl nitrate and using the hot gas to drive the actuator is the reduction of the size and hence weight of the fuel container. The cordite can bum at the supply pressure and need not be stored at say ten times the supply pressure. Since however cordite burns at a greater rate at a higher pressure a relief valve must be used to allow generated gas to escape to atmosphere at times when the demand is low. Figure shows a typical arrangement using an equal area double-acting piston. An area-half area design is often favored for hot gas servos, in which case only one valve is required per actuator; it is necessarily bigger however since the volume of gas to be supplied is larger. However, reliability is increased and manufacturing costs are reduced. Although a direct bleed from the main propulsion motor plus a heat sink has been used on a system with a flight time about 12 seconds, a separate charge of much cooler burning cordite is generally used. Since cordite produces very dirty hot gas it will be appreciated that a long flight time creates some severe design problems, and most hot gas systems used so far have been restricted to missiles with short times of flight.

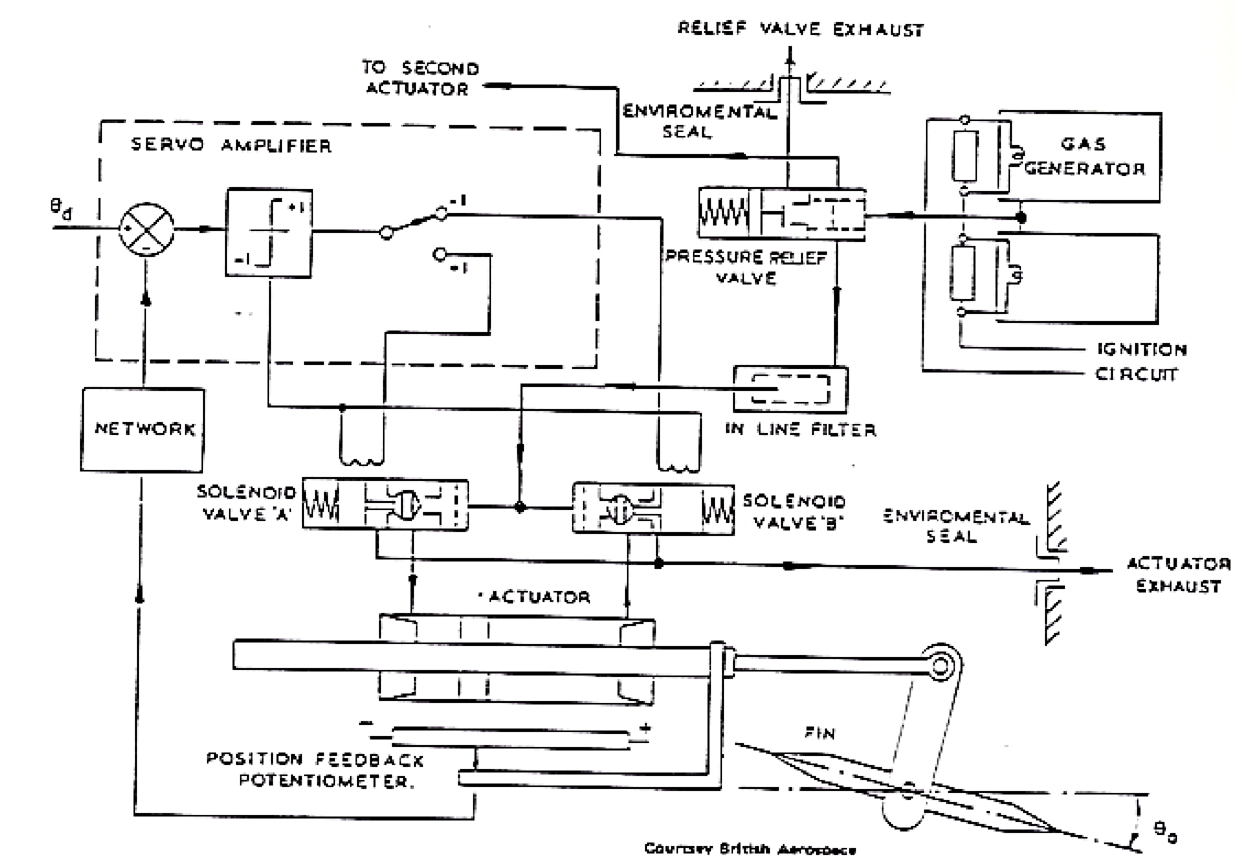


**Figure 2-13 Hot gas servo actuator**

## 2.2.2 HYDRAULIC ACTUATORS

> *HYDRAULIC ACTUATOR = control valve + [cylinder or rotary motor] + hydraulic power source*

**Figure 2-14 Actuator: control valve + cylinder**

**Figure 2-15 Actuator: control valve + rotary motor**

Selection of the optimum actuator's configuration for a specific application requires consideration of relevant factors: available energy source, load capacity, stroke, speed of response, leakage limitations, environmental conditions, storage life and conditions, size, weight and cost.

Control valve: flow control valve (FCV); pressure control valve (PCV); directional control valve (DCV). Flight applications usually require engaging of directional control valves. Directional control valve perform direct distribution of flow to various path of the system.

The categories that describe DCVs: type ((spool, poppet and rotary), numbers of positions, number of ways, number of lands, center configuration and valve driver class.

DCV's elements: body, spool sleeve subassembly (spool moves inside sleeve) and activation device.

Pumps have to provide the system with hydraulic energy (flow under desired pressure). As the power is product of pressure and flow - three strategies exists in controlling of the delivered power:

- control the pressure with a constant flow,
- control the flow with a constant pressure and
- combination pressure and flow control

In flight applications, standard choice for hydraulic actuator is control the pressure with constant flow; also other two strategies are possible.

Accumulator is standard component of hydraulic circuit.

Hydraulic accumulator has to:

- Provide an energy to the system
- Maintain relatively constant pressure levels in the system
- Perform leakage compensation
- Compensate pressure peak breaking down
- Perform shock and vibration cushioning.

Selection of accumulator type

Accumulators come in a variety of sizes and ratings and generally can be one of two types: piston and bladder. For both types the energy storage takes place in the compression of gas, usually an ideal inert gas such as nitrogen. Piston accumulators could achieve pressure range that is bigger than with bladder type.

Operation:

The piston accumulator is device of light steel construction (alternative is aluminum or same kind of composite), designed as a gas (nitrogen) pressurized vessel. The piston separates nitrogen and oil inside the accumulator. Accumulator operation is based on small compressibility of oil and on the opposite side, high compressibility of nitrogen.

Hydraulic accumulator is source of hydraulic potential (flow and pressure) that could be used during burning time of rocket engine (TVC action time) for actuating of the LRE.At the same time, the hydraulic accumulator is the shock absorber - reduces pressure surges caused by servo valves.

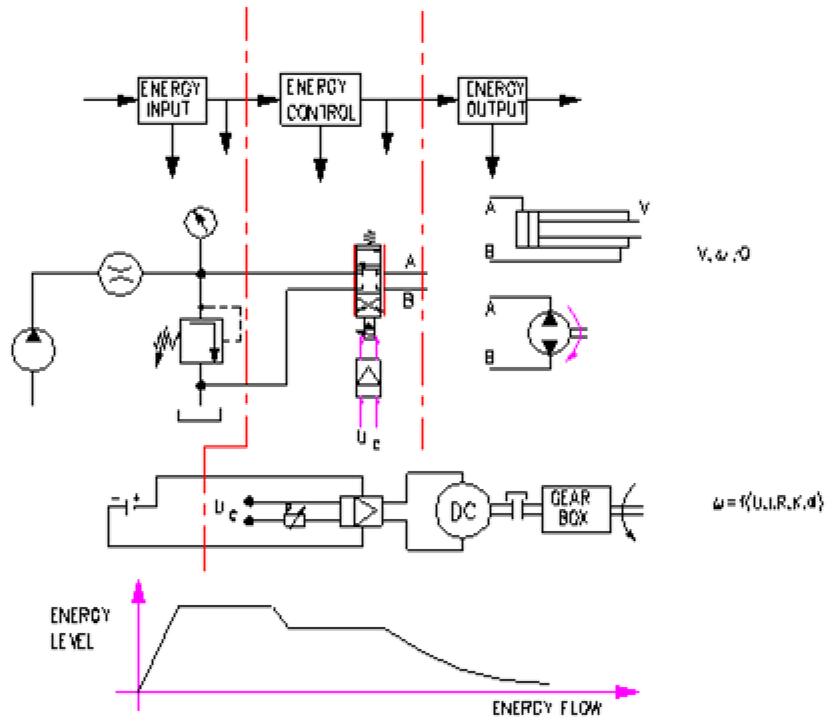


**Figure 2-16 System powered with pump**

## 2.2.3 ELECTROMECHANICAL ACTUATORS (EMA)

The use of EMA system is becoming increasingly popular in the aerospace community for variety of reasons that include maintainability, reduced number of parts, rapid production and reduced weight and costs. Recent studies have shown that hydraulic actuation systems cost the space program many valuable hours for tests, maintenance and repairs. Conventional centralized hydraulic systems used in today's launch vehicles perform well but they are an operational nightmare.

Advantages of EMA:

- increased reliability
- improved safety through the elimination of high pressure and hazardous fluids
- reduction in check-out time
- increased ability to launch on demand

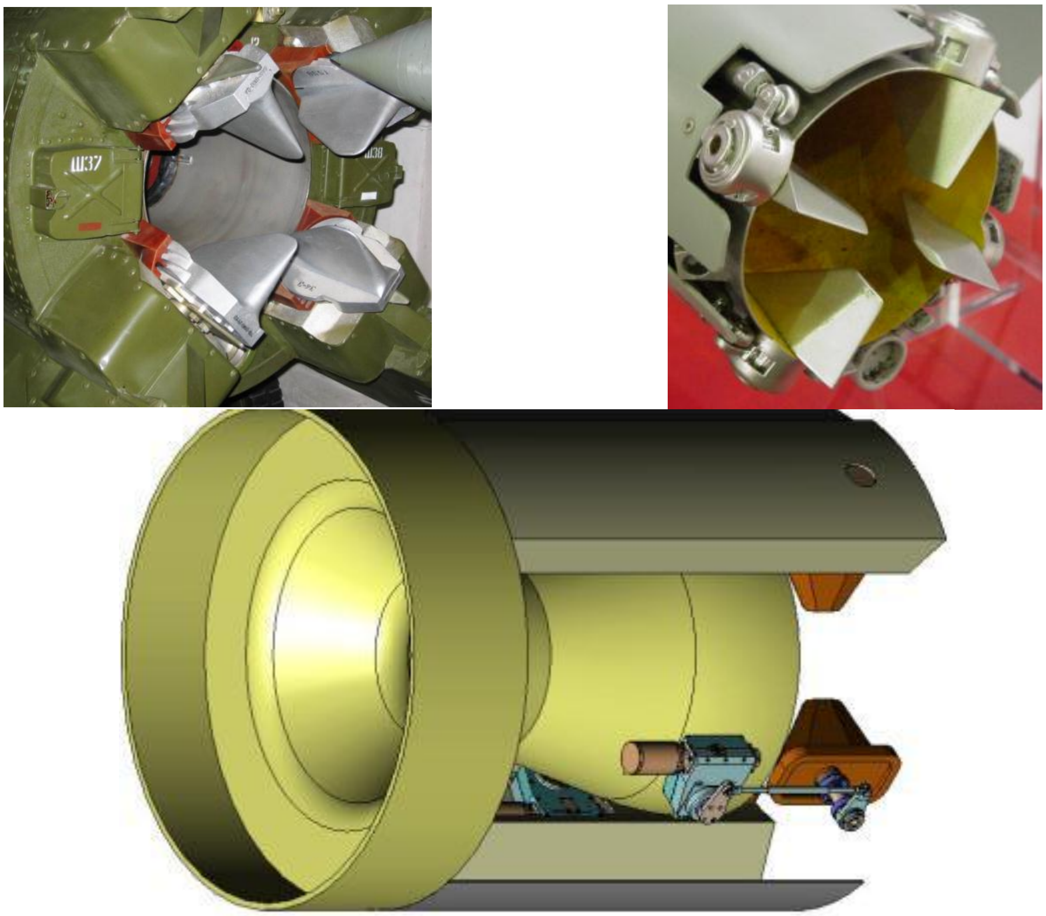


**Figure 2-17 EMA used for TVC**

The primary mechanical scheme of EMA is relatively simple even the component design could be complex.

The design consists of the following major components:

- Control surface
- Electromotor with (or without) gearhead & encoder

*Gearhead makes possible control of large load inertia with a comparatively small motor inertia. Without the gearhead, acceleration or velocity control of the load would require that the motor torque, and thus current, would have to be as many times greater as the reduction ratio which is used; DC electric motors produce large output speed with relatively small torque. In actuating system, however, opposite situation is required, so we need relative small speed (angular velocity) and large torque. Because of that, reducing system between electric motor and vane (or fin) may be introduced if it is necessary*

- Transmission subassembly

*Converts rotational input into linear output*

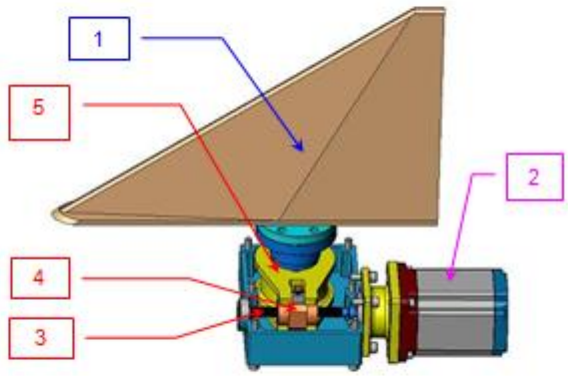


**Figure 2-18 EMA components**

*Basic principle of work*

*The electromotor and gearhead (2) in the combination with transmission subassembly - transforms rotation motion (of electromotor) via roller screw (3) in translation motion of nut (4). Translational motion of nut is then via fork lever (. 4) trans-formed into rotational motion of vane's (1) shaft.*

### *TRANSMISSION SYSTEM*

*Direct fitting of electric motor (torque-motor) to vane*

Used for applications with large torque and low speed.

**Figure 2-19 Direct fitting of electric motor (torque-motor) to vane**

*Connection of electromotor and vane using rolled threads ball screws along with some lever mechanism*

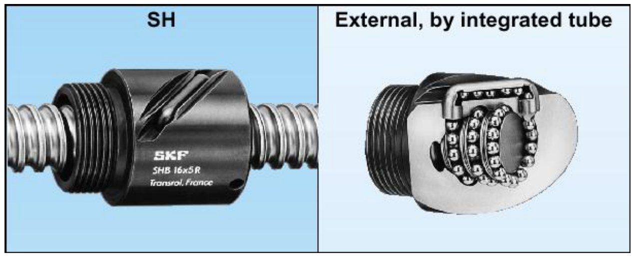


**Figure 2-20 Roller thread mall screws**

This type of screws transforms rotating motion of electric motor to linear motion of nut. However, instead of classic screw-nut conjunction, hardened steel balls rotating between them are used. In that way sliding friction is replaced by rolling friction thus increasing efficiency of the screw from 30% to around 90%.

*Connection of electromotor and vane using "Planetary roller screws"*

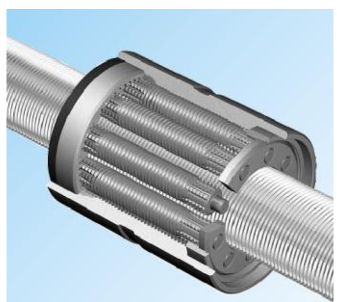

**Figure 2-21 Planetary roller screws**

Connection of electromotor and vane using "planetary roller screws" with lever mechanism or two coaxial cylinders in which screw and nut are placed, making compact construction (resembles to hydraulic cylinder). They work in a similar manner as ones with ball screws, but instead of balls, they have threaded rollers. The helix angle of the roller thread is exactly the same as the nut thread, so the roller does not move axially relative to the nut as it rolls. It has high efficiency (0.80-0.90) because there is no sliding. The absence of recalculation means that the nut is robust and capable of high rotation speed with smooth running. They also have small gap that could be eliminated by preloading. They are less sensitive to impact loads and they can carry higher loads, comparing to those with ball screws.

### *GEAR SYSTEM*

To satisfy linear velocity requirements, a speed reduction is usually needed to the output shaft. The two pass gear system utilizes a design such that backlash is nearly eliminated. Spur gears transmit high torques necessary to drive the system in either direction. A two piece idler shaft gear allows for on assembly adjustment to aid in minimizing rotational play.

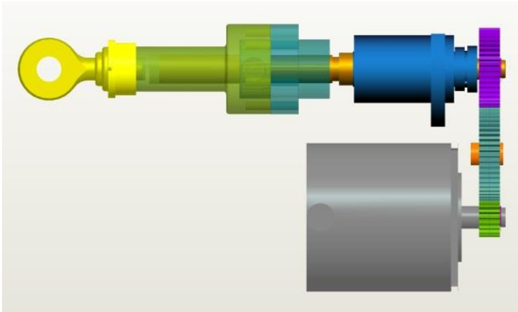

**Figure 2-22 Motor with gearbox**

*Linear Screw*

Rotational motion is converted to linear motion using a roller screw. Roller screws are high efficiency linear devices which provide a robust means of transmitting very high loads with considerable accuracy. They consist of a threaded screw shaft and a nut which houses contacting rolling elements. Triangular threads, with an included angle of 90 degrees, are machined onto the main screw shaft. Thread pitch may range from 0.015 inch to as much as 1.250 inches with 4, 5, or 6 starts. The rollers housed in the nut are machined with a single start triangular thread. Contact is made between the nut and shaft by the rollers. A barreled thread form provides a large contact radius for high load carrying capacity and rigidity.

Two critical areas of highest concern are:

- Dynamic load capacity for the given geometric envelope
- Shock load capability

The Dynamic load rating in an application depends on the type and magnitude of the load applied and the life of the screw in millions of revolutions. In a TVC system, the maximum dynamic load is only experienced at very short intervals during a flight. This characteristic duty cycle aids in compacting the actuators' geometry which, of course, is much to the advantage of the overall system design. Extreme shock loads and adverse environments may also be encountered on a mission. Transient shocks much larger than those loads experienced under normal continuous operation may be experienced by a TVC actuator at engine start up.

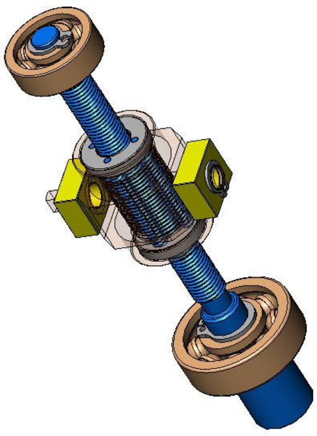

**Figure 2-23 Linear screw**

## *SELECTION OF ELECTROMOTOR*

Regardless of how simple or complex the application, there are some common requirements to consider for the selection of the proper motor (and/or controller). Some common considerations are:

*Required output torque*

Motor torque is a combination of the **internal torque loses** (a function of motor design) and **external torque load**. External torque load is a function of load inertia and load acceleration.

*Required speed range*

How fast should motor run when loaded and unloaded.

*Available space for motor mounting*

Motor length and its maximum diameter must be taken into consideration. Also, motor dimensions may be dictated by performance requirements.

*Source of power for the motor*

Source of power could be AC or DC. Current limits and voltage range could be limiting factors.

*Special shaft and/or mounting requirements*

What length and diameter of shaft are needed, and is a rear shaft extension required (for encoders, brakes, etc.) are the questions that should be kept on mind.

*Environmental considerations*

Temperature / Humidity / Shock and vibration / Altitude / Presence of chemicals, contaminants, vapor, etc.

*"Heat sinking" of motor*

A motor can be heat sinked by mounting it on a mass of thermally conductive material. The material conducts heat away from the motor. Heat sinking has a dramatic effect on motor performance. Effective heat sinking increases the continuous output torque capability of the motor.

*Expected velocity profile*

A velocity profile is a graph that shoes how quickly the motor accelerates to rated speed, the time the motor runs at rated speed, and how quickly the motor decelerates to zero speed.

### *TYPICAL MOTOR TECHNICAL SPECIFICATION PARAMETERS*

| Electrical Parameter | Typical Symbol | Unit | Definition |
|---|---|---|---|
| Reference Voltage | V | Volts | This is the rated terminal voltage. |
| Rated Current | Ir | Amps | Current drawn by the motor when it delivers the rated torque. |
| Peak Current (stall) | Ipk | Amps | This is the maximum current allowed to be drawn by the motor. |
| No Load Current | INL | Amps | Current drawn by the motor when there is no load on the motor shaft. |
| Back EMF Constant | KE | V/RPM or V/rad/s | Using this parameter, back EMF can be estimated for a given speed. |
| Resistance | R | Ohms | Resistance of each stator winding. |
| Inductance | L | mH | Winding inductance. This, along with resistance, can be used to determine the total impedance of the winding to calculate the electrical time constant of the motor. |
| Motor Constant | KM | Oz-in/√W or NM//√W | This gives the ratio of torque to the power. |
| Electrical Time Constant | τE | ms | Calculated based on the R and L of the windings. |

Motor torque, $T$, is related to the armature current, $i$ , and a constant factor $K_t$. The back EMF $e$, is related to the rotational velocity:

$$T = K_t\, i$$

$$e = K_e \frac{d\theta}{dt}$$

In SI units (which we will use), $K_t$ (armature constant) is equal to $K_e$ (motor constant).

| Mechanical Parameter | Typical Symbol | Unit | Definition |
|---|---|---|---|
| Speed | N | RPM or rad/s | Rated speed of the motor. |
| Continuous Torque | Tc | Oz-in or N-M | This is the torque available on the shaft for the given speed range. |
| Peak Torque or Stall Torque | Tpk | Oz-in or N-M | This is the maximum torque that motor can deliver for a short duration of time. This torque may not be available for all the speed ranges. |
| Torque Constant | Kt | Oz-in/A or N-M/A | This is the torque produced for every ampere of current drawn by the motor. Since the torque varies linear with current, this parameter can be used to interpolate the torque delivered for a given current and vice versa. |
| Friction Torque | TF | Oz-in or N-M | This is the torque loss due to friction which includes mainly the bearing friction. |
| Rotor Inertia | JM | Oz-in-s$^2$/N-M-s$^2$ | Rotor moment of inertia. This is useful to determine the acceleration and deceleration rates, the dynamic response of the system and to calculate the mechanical time constant of the rotor. |
| Viscous Damping | D | Oz-in/RPM or N-M-s | |
| Damping Constant | KD | Oz-in/RPM or N-M-s | |
| Temperature | T | °F or °C | Operating ambient temperature. |
| Maximum Winding Temperature | θmax | °F or °C | Maximum allowed winding temperature. If the winding temperature exceeds this limit, winding leakage current may increase or there are chances of winding breakdown. |
| Thermal Impedance | RTH | °F/W or °C/W | This is the thermal impedance posed by the motor to the ambient. |
| Thermal Time Constant | τTH | min | Time constant based on the thermal impedance. A motor with a heat sink will have a higher time constant than a motor without a heat sink. |

## *MOTOR TORQUE – SPEED CHARACTERISTICS*

Usually, for the Permanent Magnet DC motors, it is given their Speed to Torque characteristics.

$$K \cdot i = T + b \cdot \omega \rightarrow i = \frac{T}{K} + \frac{b}{K}\omega$$

$$R \cdot i + K \cdot \omega = V = T\frac{R}{K} + (\frac{b}{K}R + K) \cdot \omega = T\frac{R}{K} + \frac{b \cdot R + K^2}{K} \cdot \omega = T\frac{R}{K} + \frac{R}{K}(b \cdot + \frac{K^2}{R}) \cdot \omega$$

$$T = -(b \cdot + \frac{K^2}{R}) \cdot \omega + K \cdot \frac{V}{R}$$

$$\omega = 0, T = \tau_s = \frac{V}{R} * K$$

$$T = 0, \omega = \omega_n = \frac{K^2}{b \cdot R + K^2} \cdot \frac{V}{K} = \frac{1}{\frac{b \cdot R}{K^2} + 1} \cdot \frac{V}{K}$$

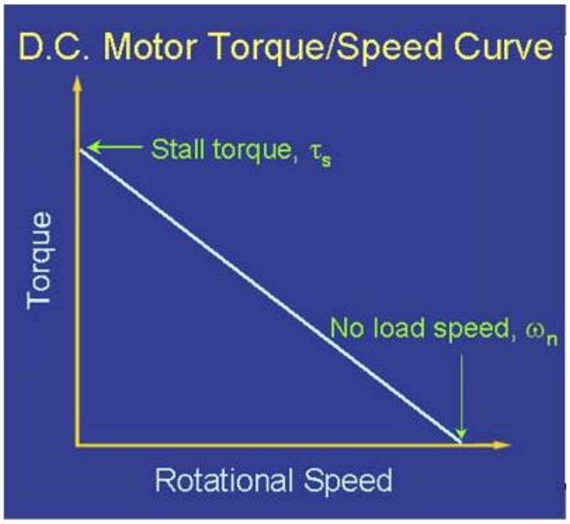


**Figure 2-24 DC motor Torque/Speed curve**

The graph above shows a torque/speed curve of a typical DC motor. Note that torque is inversely proportional to the speed of the output shaft. In other words, there is a tradeoff between how much torque a motor delivers, and how fast the output shaft spins.

Motor characteristics are frequently given as two points on this graph:

- The stall torque, $\tau_s$ , represents the point on the graph at which the torque is a maximum, but the shaft is not rotating.
- The no load speed, $\omega_n$ , is the maximum output speed of the motor (when no torque is applied to the output shaft).

The curve is then approximated by connecting these two points with a line, whose equation can be written in terms of torque or angular velocity as equations:

$$T = \tau_s - \frac{\tau_s}{\omega_n} \cdot \omega$$

$$\omega = \omega_n - \frac{\omega_n}{\tau_s} \cdot T$$

We defined power as the product of torque and angular velocity $P = T\omega$. This corresponds to the area of a rectangle under the torque/speed curve with one corner at the origin and another corner at a point on the curve (see figures below). Due to the linear inverse relationship between torque and speed, the maximum power occurs at the point where:

$$P_{max} = T_{mp}\omega_{mp}$$

$$\omega_{mp} = \frac{\omega_n}{2}$$

$$T_{mp} = \frac{\tau_s}{2}$$

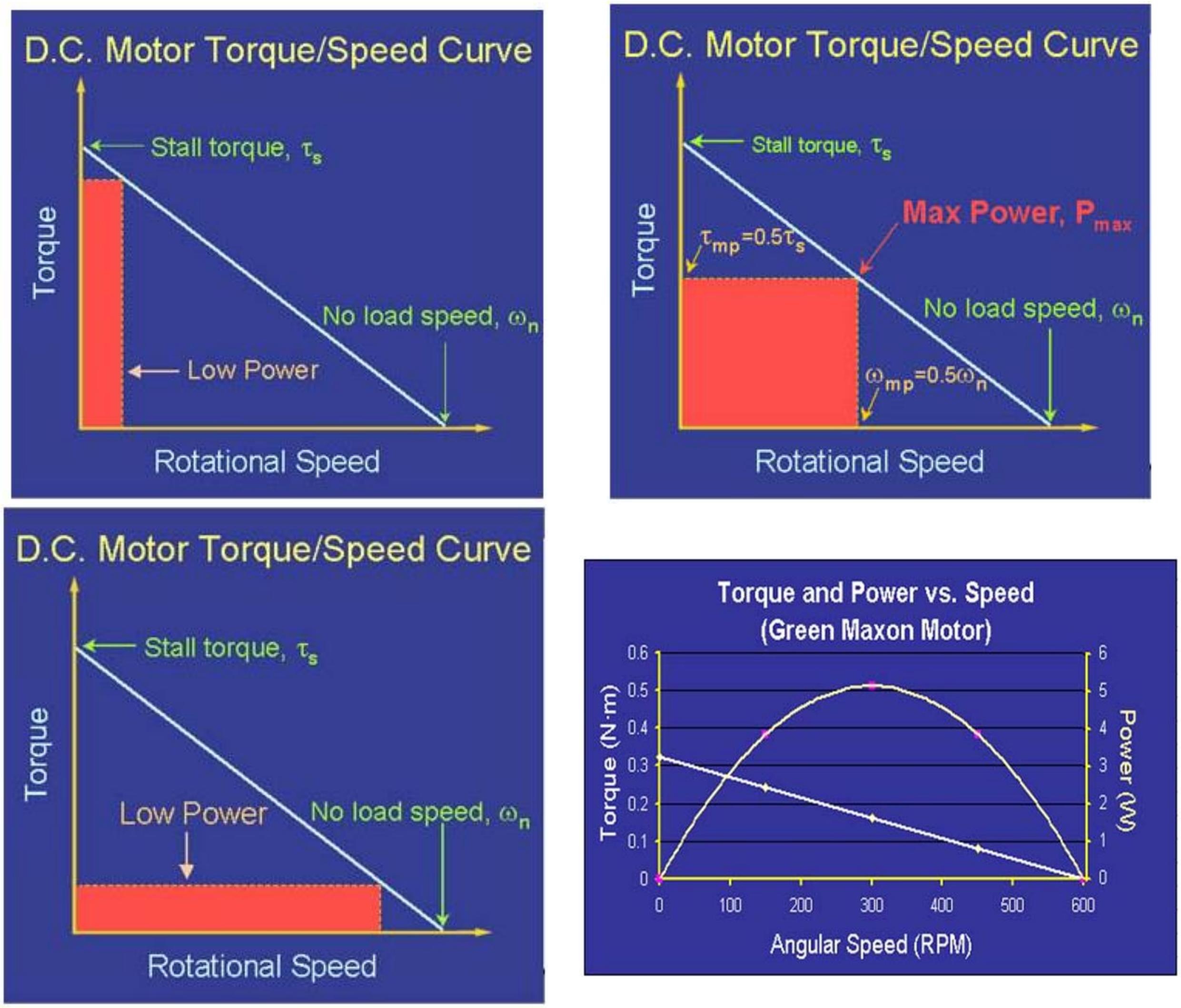


**Figure 2-25 DC motor Torque/Speed maximization**

## 2.3 MISSILE AUTOPILOT DESIGN - ROLL, PITCH AND YAW AUTOPILOTS, DESIGN & SIMULATIONS

The function of the autopilot is to **stabilize** and **guide** the missile by requesting fin deflections, which cause the missile body to rotate and hence translate. The fin servos respond to the commands ordered by the autopilot, and the actual fin deflection is computed by the balance between servo torque and aerodynamic hinge moment. These fin deflections then act to force the airframe dynamic model. The autopilot requirements and limitations are closely related to the overall design of the guidance subsystem. The aerodynamic characteristics of the missile airframe are an integral part of the autopilot design and operation. Therefore, the autopilot refers to the missile airframe dynamics and associated stability augmentation system, which is designed so that the missile lateral acceleration follows the autopilot acceleration commands as closely as possible. The design of an autopilot must be tailored to each individual missile airframe configuration and its associated **aerodynamic characteristics**, which are nonlinear functions of missile **velocity**, **angle of attack, control surface deflection, and altitude**. Therefore, a properly designed autopilot provides a nearly linear response characteristic if changes in these parameters about their nominal design values are small. It should be pointed out, however, that there are some missile designs that do not require an autopilot. **It is desirable to avoid a large angle of attack, since the associated drag results in a rapid loss of missile velocity. Furthermore, the airframe structural limit must not be exceeded.** It is common practice in missile design to limit the commanded lateral acceleration in order to prevent both angle-of-attack saturation and structural failure. Therefore, autopilot command limiting is assumed to be the dominant nonlinear effect, and all other nonlinear characteristics, such as actuator angle and angle rate limiting, aerodynamic nonlinearities, and instrumentation nonlinearities, are assumed to be secondary or equivalently represented as acceleration-limiting, or as changes in autopilot dynamics.

It is standard practice in the design of missile autopilots to utilize a linearized second-order airframe model. The airframe acceleration command must be limited in an actual missile in order to prevent structural failure or an excessively large angle of attack, which causes increased missile drag and loss of lateral (note that in missiles, **lateral movement usually means up–down or left–right**) acceleration capability, often referred to as airframe acceleration saturation. Therefore, we can define the function of the autopilot subsystem as follows:

- **Provide the required missile lateral acceleration response characteristics,**
- **Stabilize or damp the bare airframe, and**
- **Reduce the missile performance sensitivity to disturbance inputs over the missile's flight envelope**

**Autopilots are commonly classified as either controlling the motion in the pitch/yaw planes, in which case they are called lateral autopilots, or controlling the motion about the fore-and-aft axis, in which case they are called roll autopilots (or longitudinal autopilots).** Note that in aircraft design, the autopilot nomenclature is somewhat different from that of missile autopilots. Specifically, in aircraft nomenclature, autopilots designed to control the motion in the pitch plane are called longitudinal autopilots, while those designed to control motion in the yaw plane are called lateral autopilots.

Strictly speaking, a typical interceptor missile has three separate autopilots for control of roll, pitch, and yaw. The pitch and yaw autopilots control the lateral acceleration of the missile in accordance with some guidance law, such as the proportional navigation guidance law. Although the roll autopilot is not used directly in homing, nevertheless it is designed to enable maximum homing performance in the other two axes.

A realistic autopilot can be designed that requires knowledge of very few specific aerodynamic parameters, yet its response characteristics are easily related to the important missile aerodynamic properties. Figure illustrates a block diagram of a generic autopilot, which uses accelerometer feedback in order to control the lateral acceleration of the missile.

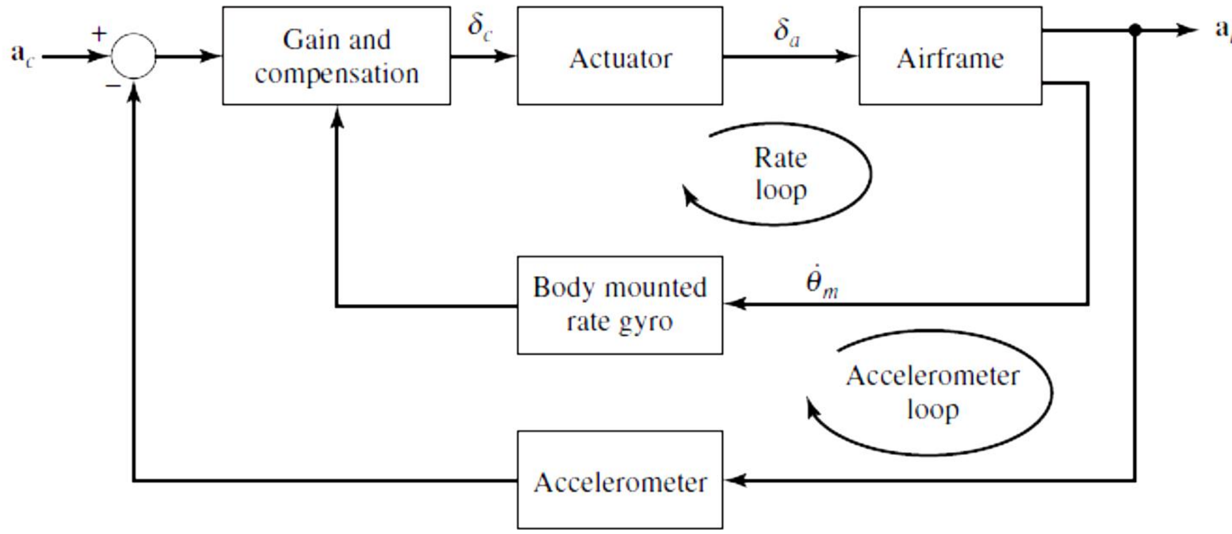


**Figure 2-26 Autopilot block diagram**

Using a linearized airframe model, the closed-loop transfer function for the general autopilot configuration of Figure can be developed for specific gains and compensation. Commonly, lateral acceleration control is used in accordance with the proportional navigation guidance law, which requires a missile lateral acceleration proportional to the measured missile-to-target line-of-sight (*LOS*) rotation rate *(dλ/dt)*. Furthermore, the body-mounted rate gyroscope senses the body-attitude rate, *dθm/dt,* which is used by the autopilot to increase the effective damping ratio of the airframe's short-period poles. The missile motion in space is completely defined by the acceleration normal to the velocity vector and the rate of change of the velocity magnitude. The commanded normal acceleration is the input to a combination of limiters and transfer functions that simulate the autopilot, control system, and aerodynamics. Specifically, the commanded acceleration is passed to the autopilot in a body frame sense.

Another effect of importance to a real missile arises if the missile is rolling and the pitch/yaw autopilots fail to compensate for the roll. This effect, which manifests itself as roll cross-coupling, causes the lateral acceleration calculated in one plane to be executed, due to system lags, in another plane. For this reason, missiles are often fitted with roll-attitude hold autopilots. The autopilot also assumes that the missile roll rate is either zero or known and compensated for.

In Figure, *ωn* is the system natural frequency, *ζ* is the system-damping ratio, and *s* is the Laplace operator. Before passing into the autopilot, **the commanded accelerations are checked to ensure that they do not exceed structural or aerodynamic limits.** That is, the inputs to the autopilot block transfer function are restricted to some maximum value if limits are exceeded. **The autopilot block transfer function can be represented either as a first- or second-order lag with inputs of commanded acceleration and outputs of realized output acceleration**.

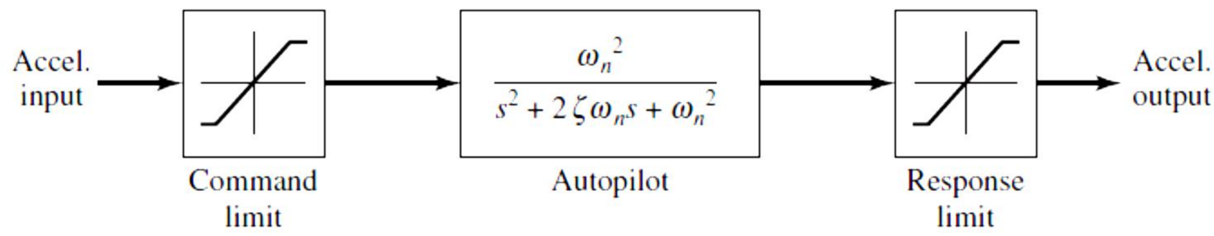


**Figure 2-27 Autopilot transfer function**

### 2.3.1 ROLL AUTOPILOT

*The basic function of the roll autopilot is to roll-rate stabilize the missile, that is, to provide missile stabilization of roll attitude about the longitudinal axis.*

This is accomplished by sensing roll rate, and using the signal to deflect the fins (or wings) by an amount sufficient to counteract roll disturbances. Moreover, the response of the system must be sufficiently fast to prevent the accumulation of significant roll angles. When mounted on an aircraft, the missiles may be mounted at some angle other than their correct flight orientation. In order to align the polarization of the illuminator and the missile front antenna, the missile must be rolled to its umbilical up position (with respect to the attitude of the launching aircraft) after launch. To produce this required roll, a fixed *dc* voltage is supplied to the missile. At shorter ranges, the roll command is not necessary and may be removed to improve the effectiveness of the missile at shorter ranges.

***Roll in a missile can be caused by:***

- Asymmetric loading of the lifting and control surfaces in supersonic flight, which occurs when pitch and yaw incidences (i.e., angles) occur simultaneously and are not equal,
- Atmospheric disturbances, especially if the missile is flying close to the ground.

Some missiles are deliberately designed to have a high roll rate, with appropriately timed periodic lateral acceleration so as to null the *LOS* rotation rate. However, high roll rates can cause cross-coupling between the symmetric pitch and yaw autopilot channels, thereby tending to destabilize the system.

***In still other missile designs, the roll autopilot is designed to hold the roll attitude of the missile nearly constant for two major reasons:***

- Because of the lags in the guidance system, rolling at moderate or high frequencies may cause a lateral corrective acceleration to occur out of the proper plane, thereby causing an increase in the miss distance;
- Severe continuous rolling may cause loss of tracking the target or loss of aerodynamic control.

One common type of roll autopilot utilizes a spring-restrained rate gyroscope for measurement of roll rate, in conjunction with proportional-plus-integral (*PI*) compensation in the autopilot amplifier, in order to give the approximate equivalent of roll-rate plus roll-angle feedback. Other roll autopilot designs utilize a free vertical gyroscope as an attitude reference. That is, in order to maintain a desired roll angle, an attitude reference must be used.

***Variation of dynamic pressure with flight conditions alters the autopilot characteristics from one of fast response with minimum stability at high dynamic pressures to one of relatively slow response with maximum stability at low dynamic pressures.***

In addition, the roll autopilot has velocity compensation to further increase the roll autopilot effectiveness over the operational envelope. Another function of the roll autopilot, say in air-to-air engagements, is to roll the missile in response to command signals initiated by the launching aircraft. In other words, and as stated above, a commanded rotation of the missile is necessary to achieve proper umbilical-up missile orientation when the configuration of the launching aircraft makes it impractical to launch the missile with this orientation.

***Maximum fin deflection is limited by the missile velocity and the altitude band in which the missile is flying.***

Rolling moments are obtained from the following four sources and converted into acceleration about the missile longitudinal axis:

- *Induced Roll:* The four fins on the missile produce a rolling moment when the wind direction is not symmetric.
- *Fin Blanking:* When the fins are displaced, asymmetric air flow causes differential lift on either side of the body. The rolling moment induced will depend on the angle of attack and Mach number; therefore, to modify these effects, a modifying function is commonly used.
- *Aileron Moment:* The effective aileron deflection $\delta a$, obtained by differential fin commands, is used to calculate a rolling moment (assumed to vary linearly with $\delta a$, but with a slope varying with Mach number.
- *Roll Damping:* The roll damping moment is assumed proportional to roll rate, and the coefficient *Cl* is looked up as a function Mach number alone.

### 2.3.2 PITCH/YAW AUTOPILOT

> *Basically, the pitch/yaw autopilots (also known as lateral autopilots) each consist of a major accelerometer feedback loop that provides the desired conversion of commanded acceleration to missile acceleration, and a minor rate feedback loop that provides the necessary damping of missile pitch or yaw rates.*

Therefore, because the pitch and yaw autopilots must control the lateral (lateral movement means up–down or left–right) acceleration of the missile in accordance with the proportional navigation guidance law, each autopilot must have feedback from an accelerometer. Additionally, one or usually two inner loops with feedback from a spring-restrained rate gyro are required for compensating the poles of the airframe response. (These two loops could also be mechanized with an integrating gyro, but at a higher cost than the improvement in drift performance would warrant.) For a symmetric cruciform missile, the pitch and yaw autopilot channels are identical. Therefore, only one will be discussed.

> *Variation of dynamic pressure with flight conditions also alters the pitch/yaw autopilot characteristics, as in the roll autopilot, from the one extreme of fast response with minimum stability at high dynamic pressures to the other extreme of relatively slow response with maximum stability at low dynamic pressures.*

This effect can be minimized by providing altitude gain switching, which permits a prelaunch selection of the proper launch logic as a function of launch altitude and target altitude. This launch logic is used to determine the proper in-flight switching, which occurs as the missile goes from midcourse to terminal phase. In addition, an in-flight course correction command called English bias is processed by the pitch/yaw autopilot to correct for a missile launch at other than the desired lead angle. Because missile acceleration and slowdown during the boost and glide phases of flight affect the missile lead angle for proper intercept, axial compensation provides lateral commands to the pitch/yaw autopilot in order to adjust the lead angle. From the time the flight control pressure (e.g., hydraulic) is up, pitch or yaw stabilization is obtained by sensing pitch or yaw rates with the pitch or yaw rate gyros, respectively.

# 3 DESIGN OF LINE-OF-SIGHT SYSTEMS, SEEKERS, AND DESIGN AND SIMULATION OF HOMING SYSTEMS

## 3.1 LINE-OF-SIGHT SYSTEMS (LOS)

> *In LOS guidance the missile follows the line of sight (LOS) from an external tracker to the target*

There are three points of interests

1. the tracker,
2. the missile
3. the target.

Hence LOS guidance is also referred to as "**three point guidance**"

These guidance systems usually need the use of radars and a radio or wired link between the control point and the missile; in other words, the trajectory is controlled with the information transmitted via radio or wire (see Wire-guided missile). These systems include:

- Command guidance - The missile tracker is on the launching platform. These missiles are totally controlled by the launching platform that sends all control orders to the missile. The 2 variants are
  - Command to Line-Of-Sight (CLOS)
  - Command Off Line-Of-Sight (COLOS)
- Line-Of-Sight Beam Riding Guidance (LOSBR) - The target tracker is on board the missile. The missile already has some orientation capability meant for flying inside the beam that the launching platform is using to illuminate the target. It can be manual or automatic.

### *COMMAND TO LINE-OF-SIGHT (CLOS)*

The CLOS system uses only the angular coordinates between the missile and the target to ensure the collision. The missile is made to be in the line of sight between the launcher and the target (LOS), and any deviation of the missile from this line is corrected. Since so many types of missile use this guidance system, they are usually subdivided into four groups: A particular type of command guidance and navigation where the missile is always commanded to lie on the line of sight (LOS) between the tracking unit and the aircraft is known as command to line of sight (CLOS) or three-point guidance. That is, the missile is controlled to stay as close as possible on the LOS to the target after missile capture.is used to transmit guidance signals from a ground controller to the missile. More specifically, if the beam acceleration is taken into account and added to the nominal acceleration generated by the beam-rider equations, then CLOS guidance results. Thus, the beam rider acceleration command is modified to include an extra term. The beam-riding performance described above can thus be significantly improved by taking the beam motion into account. CLOS guidance is used mostly in short-range air defense and antitank systems.

### *MANUAL COMMAND TO LINE-OF-SIGHT (MCLOS)*

Both target tracking and missile tracking and control are performed manually. The operator watches the missile flight, and uses a signaling system to command the missile back into the straight line between operator and target (the "line of sight"). This is typically useful only for slower targets, where significant "lead" is not required. MCLOS is a subtype of command guided systems.

### *SEMI-MANUAL COMMAND TO LINE-OF-SIGHT (SMCLOS)*

Target tracking is automatic, while missile tracking and control is manual.

### *SEMI-AUTOMATIC COMMAND TO LINE-OF-SIGHT (SACLOS)*

Target tracking is manual, but missile tracking and control is automatic. Is similar to MCLOS but some automatic system positions the missile in the line of sight while the operator simply tracks the target. *SACLOS has the advantage of allowing the missile to start in a position invisible to the user, as well as generally being considerably easier to operate. SACLOS is the most common form of guidance against ground targets such as tanks and bunkers.

### *AUTOMATIC COMMAND TO LINE-OF-SIGHT (ACLOS)*

Target tracking, missile tracking and control are automatic.

### *COMMAND OFF LINE-OF-SIGHT (COLOS)*

This guidance system was one of the first to be used and still is in service, mainly in anti-aircraft missiles. In this system, the target tracker and the missile tracker can be oriented in different directions. The guidance system ensures the interception of the target by the missile by locating both in space. This means that they will not rely on the angular coordinates like in CLOS systems. They will need another coordinate which is distance. To make it possible, both target and missile trackers have to be active. They are always automatic and the radar has been used as the only sensor in these systems.

### *LINE-OF-SIGHT BEAM RIDING GUIDANCE (LOSBR)*

LOSBR uses a "beam" of some sort, typically radio, radar or laser, which is pointed at the target and detectors on the rear of the missile keep it centered in the beam. Beam riding systems are often SACLOS, but do not have to be; in other systems the beam is part of an automated radar tracking system.

LOSBR suffers from the inherent weakness of inaccuracy with increasing range as the beam spreads out. Laser beam riders are more accurate in this regards, but are all short-range, and even the laser can be degraded by bad weather. On the other hand, SARH becomes more accurate with decreasing distance to the target, so the two systems are complementary.

## 3.1.1 LOS GUIDANCE, GEOMETRY OF ATTACK & ACCELERATION REQUIREMENTS

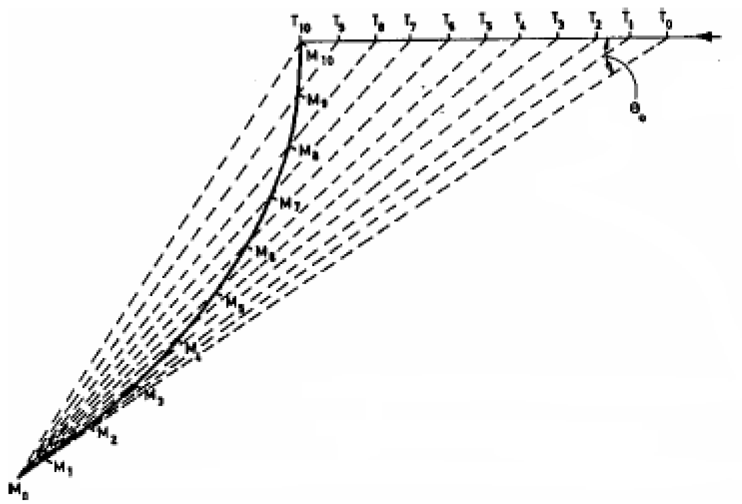


**Figure 3-1 LOS trajectory**

In an ideal implementation of LOS guidance, missile will lay on the line of sight. Missile trajectory will be curved, and trajectory curvature will increase as the missile approaches the target. Velocity of the missile will not lay on LOS. In the terminal phase, angle between velocity vector and LOS may have a very high value. This angel is called **the collision course**. For trajectory realization it is very important to know the angle between missile longitudinal axis and LOS.

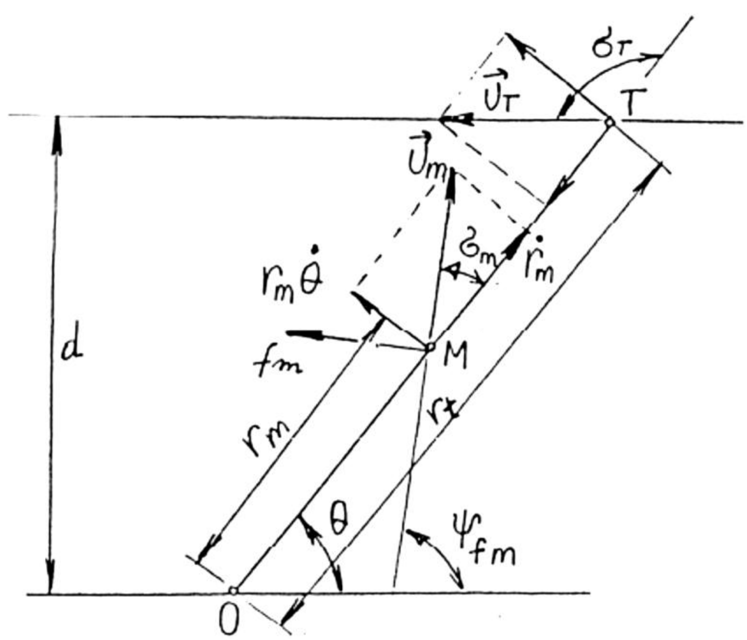


**Figure 3-2 LOS guidance planar geometry**

$\sigma_m$ – collision course (angle between missile velocity and LOS)

$\sigma_b$ – "peleng" (angle between longitudinal missile axis and LOS)

$f_m$ –lateral missile acceleration

Based on equations of motion and LOS planar geometry we can write kinematic equations:

$$\dot{r_m} = U_m \cos \sigma_m$$

$$r_m \dot{\theta} = U_m \sin \sigma_m$$

Having in mind that:

$$\sigma_m = \psi_{f_m} - \theta$$

Equation will take form:

$$\dot{r_m} = U_m \cos(\psi_{f_m} - \theta)$$

$$r_m \dot{\theta} = U_m \sin(\psi_{f_m} - \theta)$$

After differentiation of the second equation and including in to the first equation is obtained

$$\dot{r_m}\dot{\theta} + r_m\ddot{\theta} = U_m \cos\left(\psi_{f_m} - \theta\right)\left(\dot{\psi_{f_m}} - \dot{\theta}\right) + \dot{U_m} \sin\left(\psi_{f_m} - \theta\right)$$

$$U_m \dot{\psi}_{f_m} \cos\left(\psi_{f_m} - \theta\right) = 2U_m \dot{\theta} \cos\left(\psi_{f_m} - \theta\right) + r_m \ddot{\theta} - \dot{U}_m \sin\left(\psi_{f_m} - \theta\right)$$

From previous equation it can be seen that lateral acceleration should be:

$$f_m = U_m \dot{\psi}_{f_m}$$

$$f_m = 2U_m \dot{\theta} + \frac{r_m \ddot{\theta}}{\cos \sigma_m} - \dot{U}_m \tan \sigma_m$$

Lateral acceleration which is necessary to realize LOS guidance should compensate three components:

- $2U_m\dot{\theta}$ Coriolis acceleration
- $\frac{r_m\ddot{\theta}}{\cos \sigma_m}$ relative acceleration
- $\dot{U}_m \tan \sigma_m$ tangent missile acceleration

## 3.2 HOMING GUIDANCE

### *PROPORTIONAL NAVIGATION*

Proportional navigation (also known as PN or Pro-Nav) is a guidance law (analogous to proportional control) used in some form or another by most homing air target missiles. It is based on the fact that two objects are on a collision course when the direction of their direct Line-of-Sight does not change. PN dictates that the missile velocity vector should rotate at a rate proportional to the rotation rate of the line of sight (Line-Of-Sight rate or LOS-rate), and in the same direction.

### *ACTIVE HOMING*

Active homing uses a radar system on the missile to provide a guidance signal. Typically electronics in the missile keep the radar pointed directly at the target, and the missile then looks at this "angle" of its own centerline to guide itself. Radar resolution is based on the size of the antenna, so in a smaller missile these systems are useful for attacking only large targets, ships or large bombers for instance. Active radar systems remain in widespread use in anti-shipping missiles, and in "fire-and-forget" air-to-air missile systems such as AIM-120 AMRAAM and R-77

### *SEMI-ACTIVE HOMING*

Semi-active homing systems combine a passive radar receiver on the missile with separate targeting radar that "illuminates" the target. Since the missile is typically being launched after the target was detected using a powerful radar system, it makes sense to use that same radar system to track the target, thereby avoiding problems with resolution or power, and reducing the weight of the missile. Semi-active radar homing (SARH) is by far the most common "all weather" guidance solution for anti-aircraft systems, both ground- and air-launched.

It has the disadvantage for air-launched systems that the launch aircraft must keep moving towards the target in order to maintain radar and guidance lock. This has the potential to bring the aircraft within range of shorter-ranged IR-guided (infrared-guided) missile systems. It is an important consideration now that "all aspect" IR missiles are capable of "kills" from head on, something which did not prevail in the early days of guided missiles. For ships and mobile or fixed ground-based systems, this is irrelevant as the speed (and often size) of the launch platform precludes "running away" from the target or opening the range so as to make the enemy attack fail.

SALH is similar to SARH but uses a laser as a signal. Another difference is that most laser-guided weapons employ a turret-mounted laser designator which increases the launching aircraft's ability to maneuver after launch. How much maneuvering can be done by the guiding aircraft will depend on the turret field of view and the system's ability to maintain a lock-on while maneuvering. As most air-launched, laser-guided munitions are employed against surface targets the designator providing the guidance to the missile need not be the launching aircraft; designation can be provided by another aircraft or by a completely separate source (frequently troops on the ground equipped with the appropriate laser designator).

### *PASSIVE HOMING*

Infrared homing is a passive system that homes in on the heat generated by the target. Typically used in the anti-aircraft role to track the heat of jet engines, it has also been used in the anti-vehicle role with some success. This means of guidance is sometimes also referred to as "heat seeking"

Contrast seekers use a television camera, typically black and white, to image a field of view in front of the missile, which is presented to the operator. When launched, the electronics in the missile look for the spot on the image where the contrast changes the fastest, both vertically and horizontally, and then attempts to keep that spot at a constant location in its view. Contrast seekers have been used for air-to-ground missiles, including the AGM-65 Maverick, because most ground targets can be distinguished only by visual means. However they rely on there being strong contrast changes to track, and even traditional camouflage can render them unable to "lock on".

#### *RETRANSMISSION HOMING*

Retransmission homing, also called Track Via Missile or TVM, is a hybrid between command guidance, semi-active radar homing and active radar homing. The missile picks up radiation broadcast by the tracking radar which bounces off the target and relays it to the tracking station, which relays commands back to the missile.

### 3.2.1 GEOMETRY OF HOMING GUIDANCE

Homing missiles are sometimes called 'guidance method of two dots". Target tracking system is located in the rocket. In the case of this guidance it is used angular velocity of LOS instead of angle of LOS. Therefore, we have just one integration into the closed loop system. In this type of guidance there is no problem with stability of the system at large distances form the target. But on the other side in case of near distances open loop gain grow to infinity and system becomes unstable. This characteristic is known as homing kinematic instability.

Homing system is implemented so that the rate of velocity direction should be k times higher than rate of LOS.

$$\dot{\psi}_f = k\dot{\theta}$$

Where $k$ is proportional navigation constant.

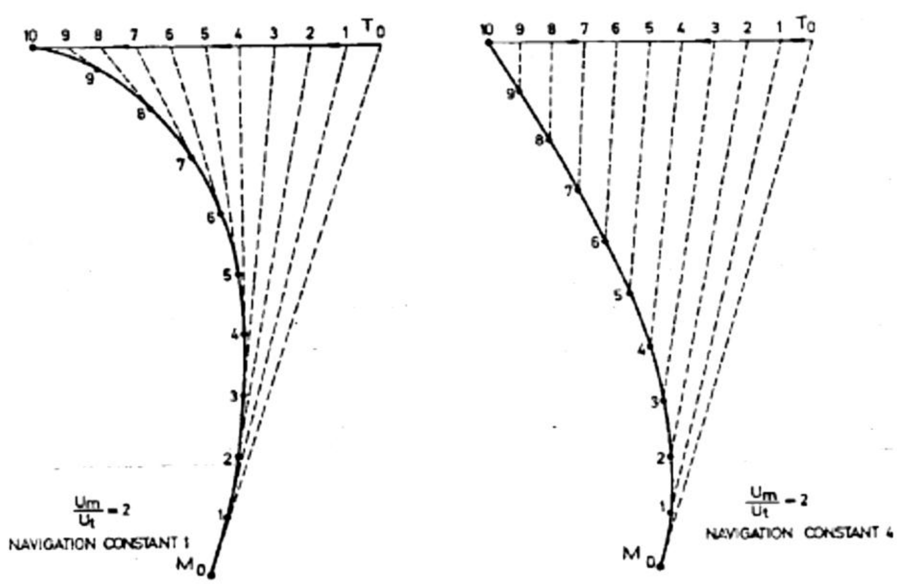


Figure 3-3 PN constant influence on trajectory

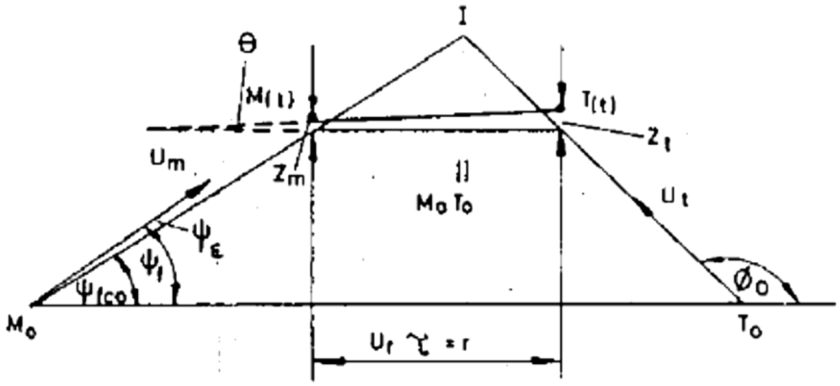

**Figure 3-4 PN guidance geometry**

Assumptions:

- Linear homing head and autopilot model
- Linearity of PN guidance geometry kinematic
- Collision course $\psi_{fco}$

$$U_m \sin\psi_{fco} = U_t \sin\phi_0$$

Let's get the imaginary line missile-target, which is parallel to the initial LOS. Values $z_m, z_t$ denote deviations of missiles and target perpendicularly to initial LOS. The actual LOS is determined by angle $\theta$:

$$\tan\theta = \frac{z_t - z_m}{r} = \frac{z_t - z_m}{U_r\tau} \approx \theta$$

Relative radial velocity of missile is:

$$U_r = U_m cos\psi_f - U_t cos\phi_0$$

$\tau$ – Time to collision

In case of small disturbances missile miss in flight plan can be noted as:

$$M = z_t - z_m \quad for\, r = 0$$

Actual collision course in case of aiming error is:

$$\psi_f = \psi_{fco} \pm \psi_e$$

In case of small aiming errors:

$$cos\psi_f = \cos(\psi_{fco} \pm \psi_e) = cos\psi_{fco} cos\psi_e \mp sin\psi_{fco} \sin\psi_e = cos\psi_{fco} \mp \psi_e sin\psi_{fco}$$

If we neglect aiming error relative velocity will be:

$$U_r = U_m cos\psi_{fco} - U_t cos\phi_0$$

Total gain of homing system is:

$$K = K_{HH}K_{PN}K_{autopilot}$$

All homing systems with the same dynamic delay will be identical if they have the same value for the following size:

$$N = \frac{Kcos\psi_{fco}}{U_r}$$

Constant N is non-dimensional at it is called kinematic gain. This gain should be in the range:

$$N = 3.5 - 4$$

PN constant can be expressed in function of kinematic gain and relative velocity:

$$k = \frac{NU_r}{U_m cos\psi_{fco}}$$

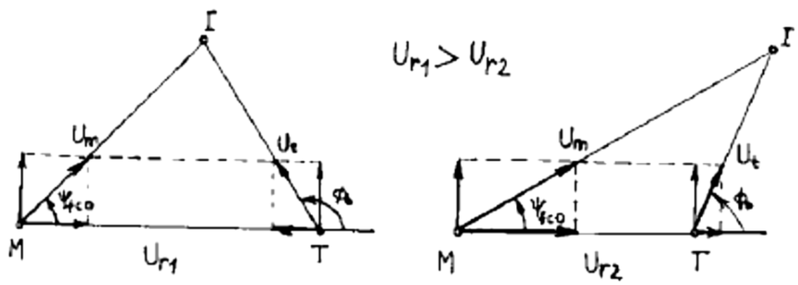


**Figure 3-5 Shooting regime and relative velocity**

## 3.3 MISSILE SEEKER SYSTEMS

Missile seekers make the measurements for target detection and homing by sensing the radio frequency (RF), infrared (IR), and/or visible energy that targets emit or reflect.

***TYPES OF SEEKERS:***

- Passive IR seekers (e.g. THAAD. EKV) detect IR energy emitted from targets using a focal plane array (FPA), scanning detector, or a single detector with a spinning reticle. Some IR seekers also include a visible sensor, which measures reflected visible light. Anti-radiation missiles (e.g, HARM) use passive RF seekers that home directly on the radar transmissions from ground- or sea-based anti-aircraft radar.

- Active seekers (e.g. PAC-3, Standard Missile) track targets with on-board radar.
- Semi-active seekers (e.g. Patriot) detect radar energy reflected from targets tracked and illuminated by ground- or ship-based radar

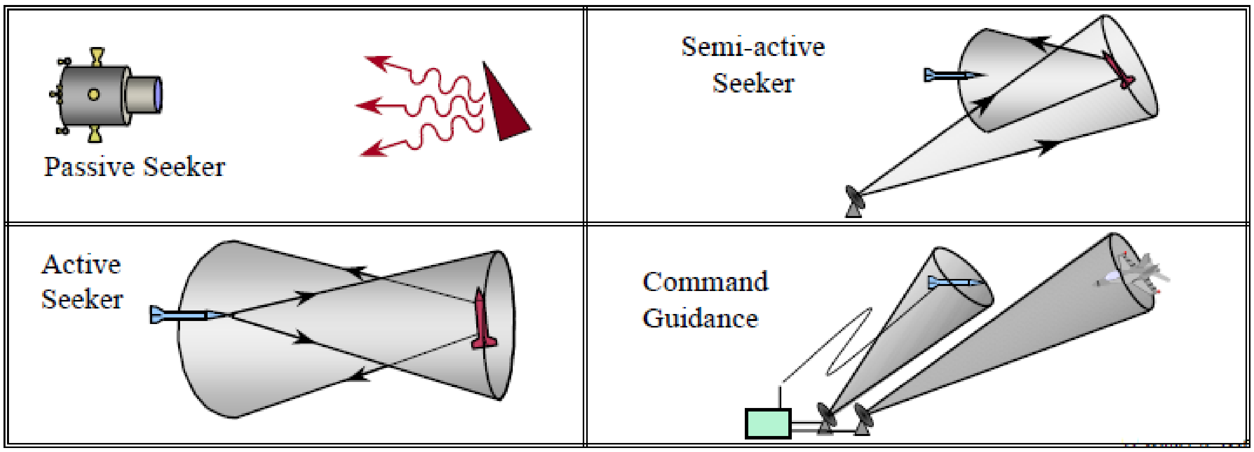


**Figure 3-6 Seekers types**

### 3.3.1 RADAR SEEKERS

Most missile radars transmit many different waveforms. The most common is a train of narrow pulses, each containing many cycles of the transmitted energy as shown at the figure. The number of cycles radar transmits each second is called its frequency, which is measured in Hertz (Hz).

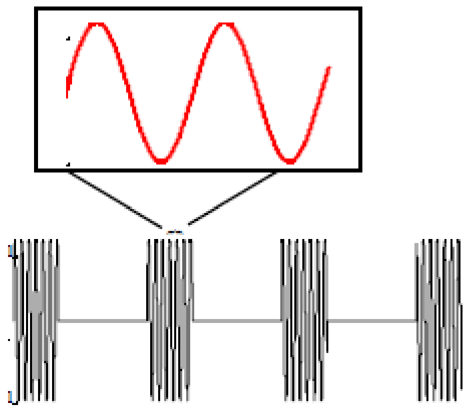

**Figure 3-7 Radar waveforms**

The period of the radar frequency is

$$T = \frac{1}{f}$$

Radar's wavelength $\lambda$ (meters) is the distance between successive peaks of its transmitted energy in space.

Radar waves travel at the speed of light, $= 3 \cdot 10^8 \frac{m}{s}$, so frequency and wavelength are related by

$$\lambda f = c$$

The formulas above apply to all portions of the electromagnetic spectrum (IR, visible, etc), not just for radar.

Radars transmit electromagnetic waves and track the LOS angle $\sigma$ via the energy reflected back from targets using a parabolic dish antenna or a phased array. Their angular accuracy is approximately their beam width

$$\theta = \frac{\lambda}{D}$$

Where D is the diameter of the dish or array

LOS rate $\dot{\sigma}$ used for homing algorithms can be found by filtering (usually Kalman filtering).

Radar can also measure a target's range r

$$r = \frac{c\Delta t}{2}$$

### *ACTIVE RADAR SEEKER*

The active radar seeker, from a radar engineer's view, may be defined as an application-specific compact missile-borne monopulse tracking radar whose antenna is mounted on a gyro-stabilized platform such that the antenna is isolated/decoupled from the body movement of the missile. The above basic idea stems from the requirement of generating highly accurate target information necessary for precise homing guidance of the missile. A typical gimbaled antenna active radar seeker is shown in Figure below.

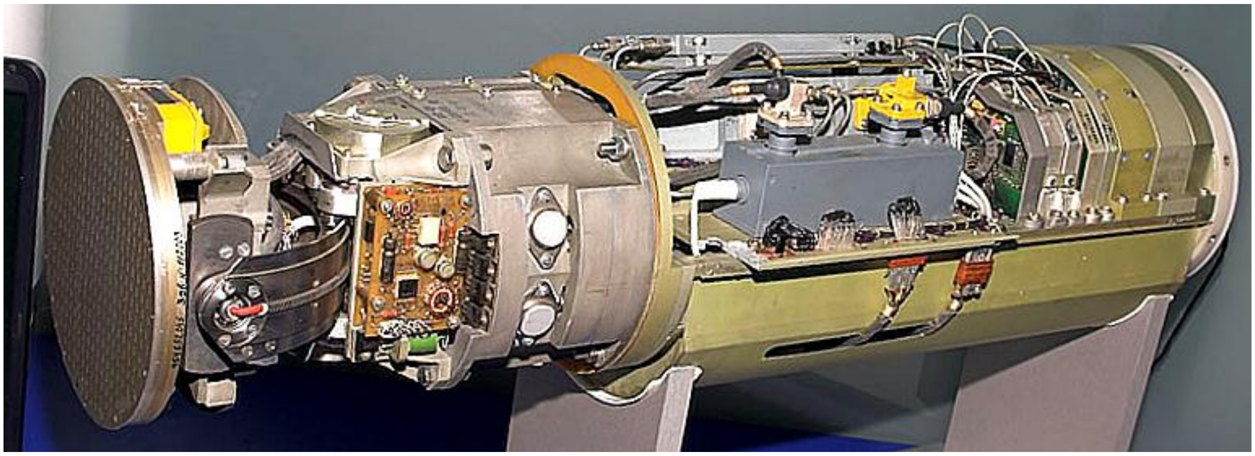

**Figure 3-8 Phazotron PSM-E series Ku-band MMW active radar seeker developed for the Kh-25MAE**

The active radar seeker essentially consists of blocks, shown in Figure 4-11, configured as a coherent three-channel monopulse master oscillator power amplifier (MOPA) system capable of tracking the target in terms of angle as well as relative velocity (in terms of Doppler frequency shift, ***A)*** and range, R (optional).

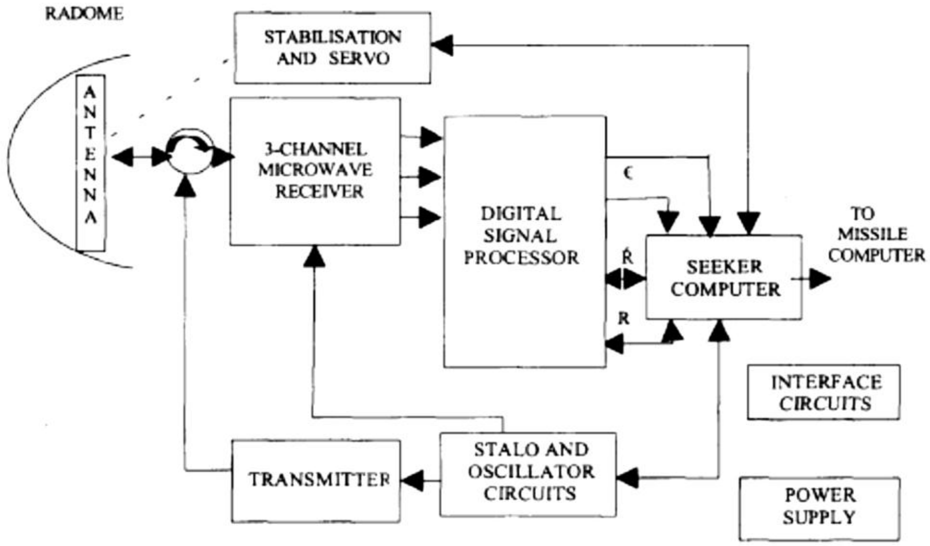


**Figure 3-9 Active radar seeker - basic blocks**

## 3.3.2 PASSIVE SEEKERS

Passive missile seekers usually operate in the infrared (IR) spectrum, which is subdivided into short wave IR (SWIR) ($1\ \mu m$ $to$ $3 \mu m$), medium wave IR (MWIR) ($3\ \mu m$ $to$ $5 \mu m$), long wave IR (LWIR) ($8\ \mu m$ $to$ $12 \mu m$) and very long wave IR (beyond $12 \mu m$).

In addition to a FPA (usually in dewar for cooling), passive seekers include optical components (lenses, mirrors, stops, and baffles).

Many different optical configurations are possible for IR seekers. Their purpose is to magnify incoming IR energy, and direct it onto the detector by mirrors and/or lenses. One common configuration, called a Cassisrain system, is shown below.

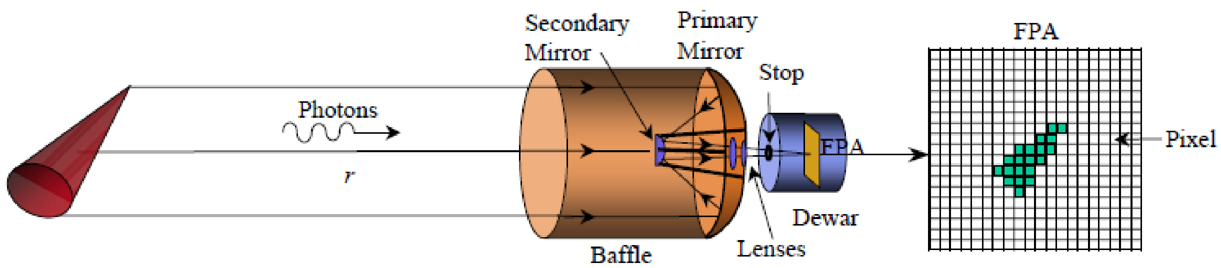


**Figure 3-10 Passive seeker system**

### *IR SEEKER*

Infrared seekers are passive devices, which, unlike radar, provide no indication that they are tracking a target. This makes them suitable for sneak attacks during visual encounters, or over longer ranges when used with a forward looking infrared system or similar cuing system. They are, however, subject to a number of simple countermeasures, most notably dropping flares behind the target to provide false heat sources. This only works if the pilot is aware of the missile, and modern seekers have rendered these increasingly ineffective even in that case.

The three main materials used in the infrared sensor are lead(II) sulfide (PbS), indium antimonide (InSb) and mercury cadmium telluride (HgCdTe). Older sensors tend to use PbS, newer sensors tend to use InSb or HgCdTe. All perform better when cooled, as they are both more sensitive and able to detect cooler objects.

Early infrared seekers were most effective in detecting infrared radiation with shorter wavelengths, such as the 4.2 micrometre emissions of the carbon dioxide efflux of a jet engine. This made them useful primarily in tail-chase scenarios, where the exhaust was visible and the missile's approach toward it was carrying to toward the aircraft as well. In combat these proved extremely ineffective as pilots attempted to make shots as soon as the seeker saw the target, launching at angles where the target's engines were quickly obscured or flew out of the missile's field of view. Such seekers, which are most sensitive to the 3 to 5 micrometre range, are now called single-color seekers. This led to new seekers sensitive to both the exhaust as well as the longer 8 to 13 micrometer wavelength range, which is less absorbed by the atmosphere and thus allows dimmer sources like the fuselage itself to be detected. Such designs are known as "all-aspect" missiles. Modern seekers combine several detectors and are called two-color systems.

All-aspect seekers also tend to require cooling to give them the high degree of sensitivity required to lock onto the lower level signals coming from the front and sides of an aircraft. Background heat from inside the sensor, or the aerodynamically heated sensor window, can overpower the weak signal entering the sensor from the target. (CCDs in cameras have similar problems; they have much more "noise" at higher temperatures.) Modern all-aspect missiles like the AIM-9M Sidewinder and Stinger use compressed gas like argon to cool their sensors in order to lock onto the target at longer ranges and all aspects. (Some such as the AIM-9J and early-model R-60 used a peltier thermoelectric cooler).

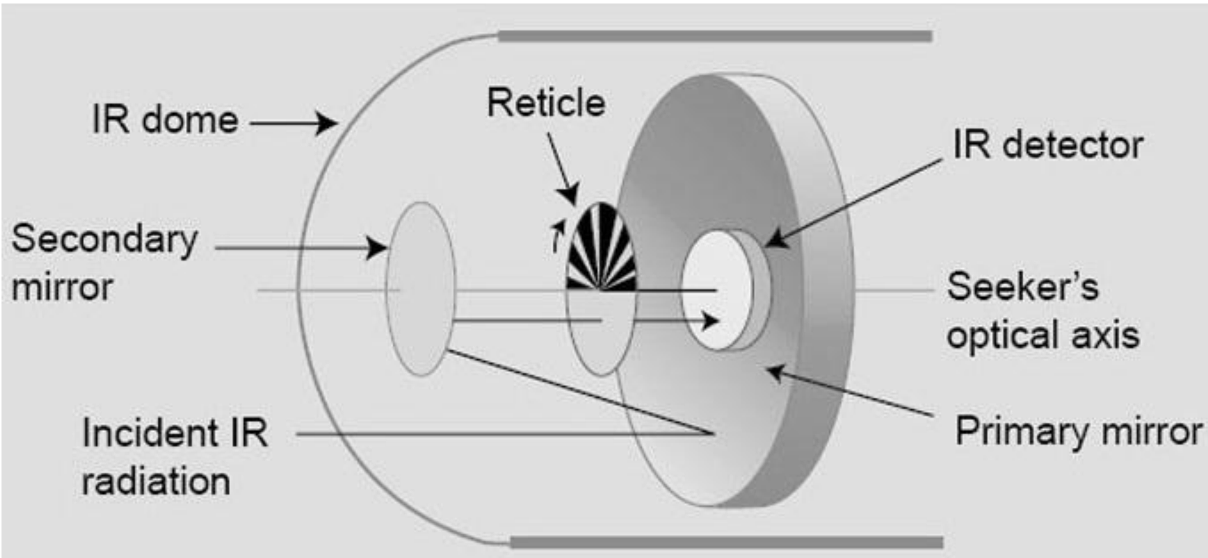


**Figure 3-11 IR homing head design**

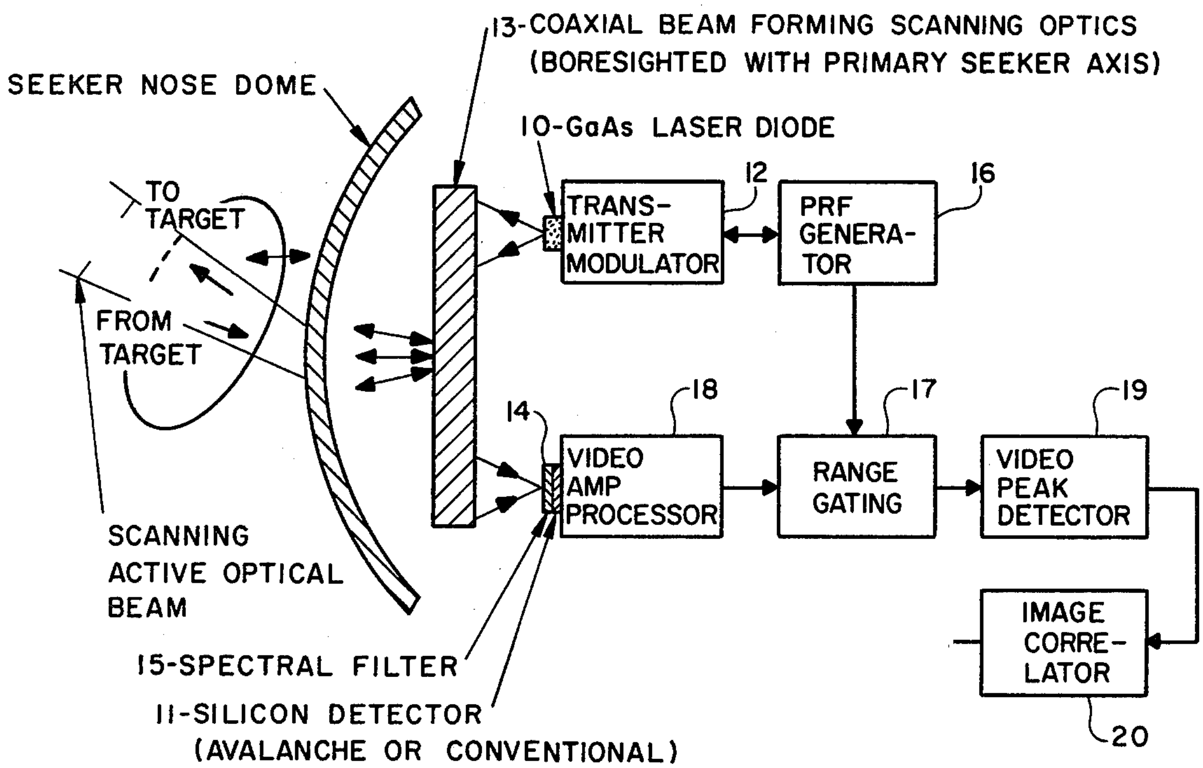


Figure 3-12 IR seeker block diagram

### *TV HOMING HEAD – CONTRAST SEEKER*

Optical contrast seekers, or simply contrast seekers, are a type of missile guidance system using a television camera as its primary input. The camera is initially pointed at a target and then locked on, allowing the missile to fly to its target by keeping the image stable within the camera's field of view. The first production missile to use a contrast seeker was the AGM-65 Maverick

Analog television cameras scan an image as a series of horizontal lines that are stacked vertically to form a grid or "frame". The camera's progression through the frame is carefully controlled by electronic timers. As the camera scans the image, the brightness of the location currently being scanned is represented as a voltage. The series of varying voltages forms an amplitude modulated (AM) signal that encodes the brightness variations along any given line, and spikes of the signal indicate when the line or frame changes.

The contrast seeker is a simple device that can be implemented using very basic analog electronics. It first uses some form of automatic gain control to adjust the image brightness until it contains some areas with high-contrast spots. This produces a bias voltage signal to represent the background brightness level, making brighter objects stand out. Any rapid change in contrast along a given scan line causes the voltage from the camera to suddenly change. These rapid changes trigger circuitry that sends the voltage of the television's horizontal and vertical deflection magnet drivers into capacitors. Thus the capacitors store a value representing the Y and X locations of any high contrast spot within the image.

The missile is initially brought onto the target manually, normally using a small cueing input on the pilot's control stick or by the weapons officer in two-seat aircraft. When the trigger is pressed to pickle the target, the timers trigger on a bright spot near the center of the image, and that time is written to an analog memory. Normally the recorded spot is indicated on the screen in the cockpit and the pilot can select other high-contrast spots within the image, in an attempt to select one that is either the target or very close to it. From then on the contrast trigger circuits fire only for signals close to that X and Y location, filtering out other objects.

Once a suitable target image has been selected, the seeker enters tracking mode. As the television scanning process continues, the contrast triggers continually produce new X and Y locations that are compared to the ones in the memory. The difference between the voltages gives an X and Y error, which is used by the seeker's gimbal mounting to turn the camera so it re-aligns with the original location. The guidance system then compares the angle of the camera to the angle of the missile body, and sends commands to the aerodynamic controls to bring it back onto a collision course. To address the need to track moving targets, a proportional navigation system is normally used, which naturally produces the required "lead"

Contrast seekers are, obviously, subject to problems when the contrast spots changes. This can occur quite easily if the target changes angle, causing the absolute brightness of the object to change, or if it moves, which can change the contrast relative to the surroundings. For instance, a tank on a roadway might provide a very high contrast tracking spot, only to have that disappear when it drives off the road into low bush. It can also be fooled by artificial lighting changes and similar effects. It is for this reason that the timers are gated, to limit the area in which the changes can take place without breaking lock.

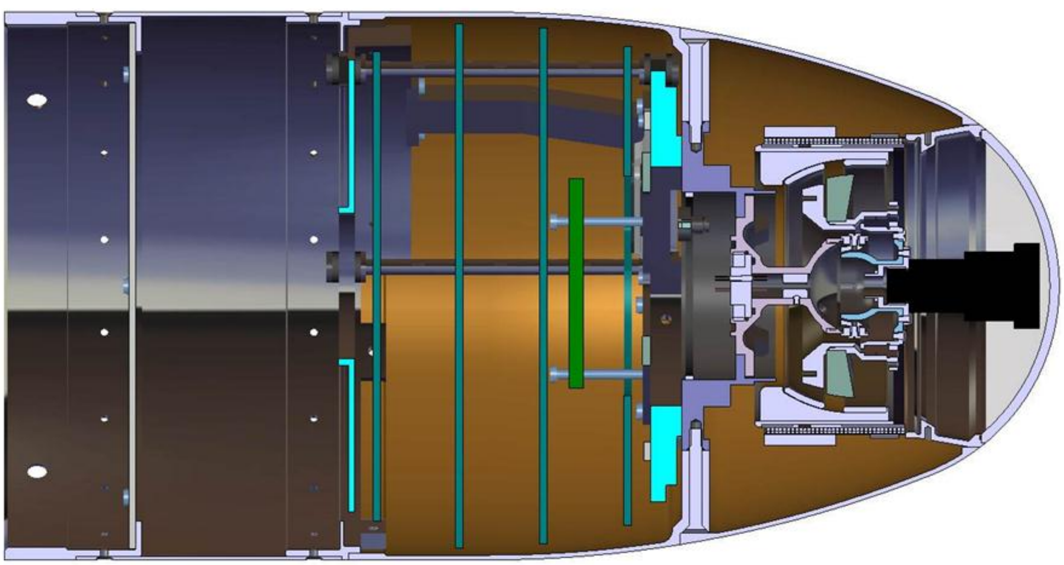

**Figure 3-13 TV homing head**

### *LASER SEEKER*

Laser guided weapons, such as the Lockheed Martin Hellfire, and Lahat and Nimrod, developed by IAI/MBT offer many advantages for heliborne and airborne use. The SAL seeker is relatively low cost, offering high precision operational flexibility, despite its adverse weather limitations.

Semi-Active Laser (SAL) guidance techniques combine very high precision, with "man-in-the-loop" capability. Weapon's sensors, utilizing SAL to detect a coded laser spot which is created by a target designator illuminating the target. The laser spot clearly marks the target to an aerial attacker or guided weapon.

Laser guided weapons can also be used effectively in urban terrain. However, constant lines of sight between the target, laser designator and the weapon must be maintained. Ground designation and UAV designation provide effective support for laser targeting. Ground designation is effective when vertical targets (walls, doors and windows in buildings) are engaged while UAVs or other aerial platforms can take advantage of the unobstructed view of the scene. UAVs and airborne designators become very efficient when engaging moving cars, which can be hidden from ground observers behind other vehicles, buildings or trees. Airborne designators can also designate combatants hidden behind walled patios or in orchards, etc. However, due to their high aspect angle, they are limited in the targeting of vertical surfaces, such as windows or doors. Except for line of sight verification and allocation of codes prior to the mission, the use of SAL does not impose further limitations or complexities when operating in good visibility conditions, (day or night) and therefore it is suitable for ad-hoc engagements of targets of opportunity and close air support. Modern laser guided weapons are integrating both GPS and laser guidance capability, offering high precision, all-weather attack capability.

Laser guided weapons are adversely affected by visibility conditions (clouds, smoke etc) and obscurants. There are also operational limitations on the flight envelopes which have to be flown, exposing aircraft and helicopters to anti-aircraft weapons. When deployed or designated from standoff range, laser guided weapons usually reach the target at flat angles, which do not have sufficient vertical velocity for deep penetration of flat structures (such as underground bunkers). Therefore, such weapons are preferably dropped from shorter distances and high altitude, or through a "loft" maneuver to maintain steep attack angles and high penetration speed.

**Figure 3-14 SAL seeker**

# 4 MISSILE NAVIGATION METHODS

Missile navigation algorithm is necessary in any type of trajectory guided missile. It is widely used in inertial navigation guidance (ballistic trajectory) and ranging navigation guidance (cruise missiles).

## 4.1 INERTIAL NAVIGATION SYSTEMS

> *Inertial navigation algorithm is based on integration of signal from inertial sensors: gyros and accelerometers.*

Many types of inertial instruments have been invented in the past, are currently being invented, and will continue to be invented as the market for guidance, navigation, and control continues to expand. Based on the technologies used, which decide the size, cost, and performance, some of the inertial instruments have found a niche in current applications, while some did not progress much beyond the laboratory/prototype stage.

### 4.1.1 ACCELEROMETERS

The accelerometer is the basic sensor of inertial navigation. This section will examine its theoretical operation, its limitations, and its practical implementation.

***GRAVITY***

> *Gravity is defined as the force per unit mass required to keep a test mass in the same position relative to the Earth.*

It as shown that an object fixed to the surface of the earth experiences a centrifugal force as viewed from a reference frame spinning with the earth.

The natural trajectory for an object positioned at a radius of R, given some initial tangential velocity, is in an orbit around the planet. Its real trajectory is not a circle centered at the center of the earth (an orbit) but a circle of constant latitude.

It is the practice in geodesy to combine both gravitational attraction and centrifugal force together into a single force, divided by the mass, and call the result **gravity.** Local density changes in the earth give rise to local changes in the gravitational attraction, and it also varies obviously with elevation. The centrifugal force varies with latitude as shown in Figure below.

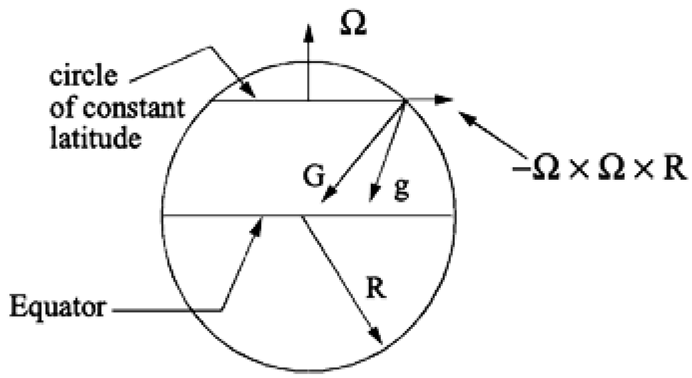


**Figure 4-1 Gravity acceleration**

## *SPECIFIC FORCE*

Imagine an accelerometer constructed from a test mass on a spring. It is intuitively clear that such a device must measure the influence of gravitation because it will deflect with the weight of the mass. It is not enough to simply calibrate this constant out of the device, because, when mounted on a vehicle climbing through the atmosphere at constant speed, the mass would slowly rise, since gravitation varies with distance, compressing the spring and indicating an erroneous acceleration upward. This is yet another example of the principle of relativity, for an observer in such a vehicle cannot tell he is in a gravitational field. This is the fundamental reason why an INS can only operate in a known (or known to be insignificant) gravitational field, and why space vehicles using an INS must know the gravitational field of any planets they encounter to avoid missing the real accelerations that they cause.

Provided a vehicle always operates near the surface of the earth, the gravitational constant acceleration can in principle, be calibrated out. However, as a practical matter, real vehicles rotate so the direction of the gravitational attraction vector is not constant in the vehicle frame of reference. Some systems however, completely avoid the gravitation problem by maintaining the accelerometer sensitive axis normal to the local gravitation vector. The real issue is that the deflection of the mass is directly influenced by the total force on the mass and not by its acceleration. When stationary in a gravitational field, a mass can experience force even when there is no acceleration. When free-falling, a mass can register no force even when there is acceleration.

> *A quantity called the specific force is defined as the total force divided by the mass.*

## *EFFECT OF THE EARTH'S GRAVITATION*

It was shown that a nonrotating geocentric frame of reference is inertial. This allows the expression of Newton's laws in such a frame. However, it is now clear that choosing an inertial frame is only half the problem. It is also necessary to know the value of the gravitation vector field in which the device operates. In general, this field is determined by the superposition of all influences of all bodies in the universe. This can be approximated easily by considering only those bodies which exert significant gravitational force on the test mass.

Gravitational force is proportional to the mass on which it is exerted, so it is completely equivalent to speak of acceleration due to gravitation. It is clear that when operating near the earth, its gravitation is significant, as is the gravitation of the sun or moon or planets when operating close to them. However, on the earth, the sun and moon give rise to the tides, so they are not completely insignificant forces.

Luckily, the acceleration due to the sun is about 0.0006 g and that of the moon is similarly small, so they can be neglected as engineering approximations when operating near the earth.

### *BASIC EQUATION OF INERTIAL NAVIGATION*

The frame of reference problem is finally solved in principle. The basic equation of inertial navigation is the specific force equation which relates the accelerometer readout to the gravitational force on the vehicle and its inertial acceleration. Consider a spring and mass accelerometer mounted to a vehicle chassis. The spring experiences forces T required to counteract gravity when the vehicle is stationary, and further forces as the vehicle executes accelerated motion dragging the mass along with it Using a subscript of i to indicate a geocentric inertial frame, T to represent total force registered by the accelerometer, G to represent gravitational attraction, and m for the accelerometer test mass, Newton's second law gives:

*Total force:*

$$T - G = ma_i$$

*Specific force:*

$$\frac{T}{m} = a_i + \frac{G}{m}$$


### *PRACTICAL ACCELEROMETERS*

All accelerometers operate similarly, though details of construction differ. A practical accelerometer is a precision instrument which couples a mass to the instrument case through an elastic, viscous, or electromagnetic restraint. Typically, the mass is allowed a single degree of freedom which may be linear or rotary. Rotary deflecting devices, called pendulous devices, are often used. Calibration of the restraint, whatever its form, provides a measure of the specific force that is tending to cause a deflection along the free degree of freedom.

Most devices operate in a closed loop, called a rebalance loop, which prevents the mass from actually moving. In this case, the effort required to retard motion is a measure of the specific force experienced. These devices are generally superior to those permitting error inducing motion. Most devices incorporate fluids to damp vibrations or to support critical components through buoyancy.

The design of inertial grade accelerometers is concerned with achieving:

- null repeatability, $5 \cdot 10^{-4} g$ is typical
- low threshold (sensitivity), $10^{-4} g$ is typical
- linearity, 0.10% over $10g$ is typical
- large dynamic range
- time constant, few milliseconds and less is typical
- low sensitivity to vibration

Accelerometers may be designed to be mounted on a stable platform. These incorporate their own small gyroscopes and perform integration as part of their function. The string accelerometer maintains mass between two strings under tension. Beat frequencies between the vibrating strings provide the output. There are many exotic types including, nuclear, electrostatic, cryogenic, gaseous, particle stream, solid state, and vibrating reed instruments.

The majority of electromechanical accelerometers are the restrained mass or force rebalance types, in which a proof mass is supported in a plane perpendicular to the input (sense) axis by a flexure, torsion bar, or pivot and jewel. The motion of this proof mass under changes of acceleration is detected by a pickoff. A rebalance force may be generated through a servo feedback loop to restore the proof mass to its null position. The force rebalance type of accelerometer has been successful not only because it is relatively small, simple, very rugged, and reliable, but also because it can be designed to meet different performance and application requirements by careful selection of the flexure and mass configuration, electromagnetic pickoffs and forces, servo electronics, fluid and damping, and materials. Force rebalance accelerometers can operate in strapdown or gimbaled modes. The output needs to be digitized.

The highest performance accelerometer available is the **Pendulous Integrating Gyro Accelerometer** (PIGA), which is used for strategic missile guidance The PIG part of the PIGA is identical to the floated single-degree-freedom, integrating gyro with the addition of a pendulous mass located on the spin axis. The PIGA is a very stable, linear device, with very high resolution over a wide dynamic range. PIGAs are relatively complex and perceived to have high life-cycle costs due to the three rotating mechanisms (gas bearing, servo-driven member (SDM), and slip ring)

Another type of accelerometer is **the resonator** or open-loop type such as the vibrating string accelerometer. This device has low shock tolerance.

**Angular accelerometers** were initially used in the 50s for dynamic compensation of AC (alternating current) servomechanisms. The basic configuration is a fluid-filled ring with a vane extending into it. Under rotational motion of the ring, the vane is restrained by a torquer, whose current indicates the angular displacement. Such devices are used in applications requiring high bandwidth (2000Hz), small magnitude stabilization, or jitter compensation. However, angular displacement sensors are not as accurate as floated gyros or DTGs below about 20Hz, but the high cost of these gyros restricts their use.

In less than 20 years, **MEMS (micro electro-mechanical systems)** technology has gone from an interesting academic exercise to an integral part of many common products. Silicon micromechanical instruments can be made by bulk micromachining (chemical etching) single crystal silicon or by surface micromachining layers of polysilicon. Many manufactures are developing gyros and accelerometers using this technology. Their extremely small size combined with the strength of silicon makes them ideal for very high acceleration applications. Silicon sensors provide many advantages over other materials, such as quartz or metal, for micro sized rate sensor development. These advantages include excellent scale factor matching and stability, long life, bias stability, virtually no degradation, and the ability to handle larger stress levels.


**Figure 4-2 Piezoelectric accelerometer**

**Figure 4-3 MEMS accelerometer**

### 4.1.2 MECHANICAL GYROSCOPES

> *While the accelerometer measures translational motion, the **gyroscope or gyro** is a device used to measure or maintain orientation.*

Cartesian coordinate systems are defined by the orientation of their reference axes, so the gyroscope can be used to solve the coordinate system problem. The properties of the gyroscope are based on the spinning rotor which was first investigated by Euler and Faucault in the eighteenth century. The two primary properties of interest to inertial navigation are the **rigidity** of orientation in inertial space and **precession.**

#### *SIMPLE GYROSCOPE*

A stable angular reference might be constructed based on the rigidity property of the gyro as follows. Use a symmetric body (disk or ring), spin it about its axis, and isolate it from external torques. The method used to isolate the spun mass from external torques is to use a system of hierarchical, independent, ideally frictionless pin joints called gimbals as illustrated in Figure below.

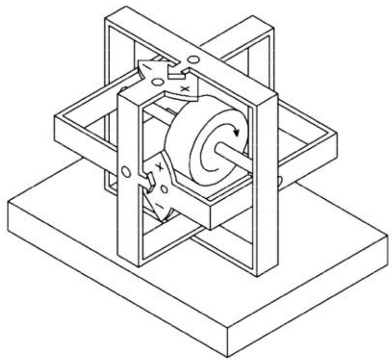

**Figure 4-4 Gimbal platform**

As the platform is rotated, the gyro will tend to remain stably pointed in inertial space. Sensors mounted on the gimbals, as indicated schematically in the figure, can measure the vehicle rotation in this way.

## *SINGLE AXIS GYROSCOPE*

Real gyros are built based on the second property of the spinning rotor, precession, and employ servomechanisms to implement a stable table on which the accelerometers are mounted.

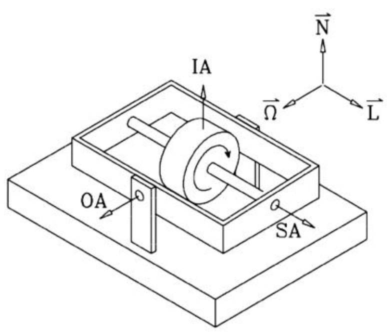


**Figure 4-5 Single axis gyroscope**

Any attempt to yaw the platform on which the gyro is mounted about the input axis will give rise to a nodding of the device about the output axis. A rotational transducer such as a resolver placed on the output axis will register this rotation which is proportional to the applied torque. Using this technique, gyros can be employed as null indicating devices in the feedback loops of servomechanisms.

Single axis gyros such as this one are called restrained gyros since they are prevented from motion about all axes but one, and, in practice, the output angles are never allowed to become large. There are also two degree of freedom gyros in common use. These are called free gyros, incorporate an extra set of gimbals and are used to control angular velocity around two orthogonal axes. Gimbals may be used both in the gyro itself and in the stable table, so sometimes it is necessary to distinguish these.

### *PRACTICAL GYROSCOPES*

Gyros can be used as **rate** gyros or as **integrating gyros.** The former are characterized by spring restraint of the output axis and low damping and derive their name from the fact that the output angle is proportional to the input rate. These are useful as turn indicators, but are not used in precision inertial systems. The latter are characterized by no spring restraint and much higher damping of the output axis, and derive their name from the fact that the output angle is proportional to the integral of the input rate.

Typical requirements for a gyro are the measurement of one thousandth of a degree per hour of angular motion. The ability to manufacture precision gyroscopic instruments which are sufficiently free of unpredictable drift is the ultimate limiting factor in the performance of inertial navigation technology. Gyro drift is the overriding performance concern and it is caused by parasitic precessional torques which are outside the system mechanization loops.

Gyroscopes develop imperfect performance for all of the following reasons:

- mass imbalance of the rotor
- thermal gradients
- anisoelasticity
- buoyancy imbalance (in floated devices)

Current technology relies on the following techniques to improve performance:

- notation in fluid to buoy components and damp oscillations
- high incidence of cylindrical symmetry in components
- precise manufacturing tolerances
- matched coefficients of thermal expansion
- temperature control
- preloaded bearings

Some practical designs of gyroscopes include the floated gyro, originally developed at MIT, air bearing gyros, electrostatic gyros, optical gyros (laser gyros, fire optical gyros FOG), direct torque gyros DTG and MEMS gyros.


**Figure 4-6 FOG**

Figure 4-7 Ring laser gyro


Figure 4-8 MEMS gyro

Figure 4-9 DTG gyro

## 4.2 SATELLITE NAVIGATION SYSTEMS

Satellite navigation was developed to answer the military need to precisely determine the position of air, sea, and land vehicles. It is rapidly becoming a definitive navigation tool since it provides continuous, high accuracy positioning anywhere on the surface of the planet and the near space region, 24 hours a day, under all weather conditions. The receiver systems are small, lightweight, and easy to use, and have made handheld global positioning systems a reality.

The military has used the TRANSIT constellation of satellites since 1964 and the resource industries have used the Starfix system since 1986. Both of these systems are now obsoleted by two more recent systems, the Global Positioning System developed by the United States Department of Defense, and its Soviet counterpart, called Glonass.

All of these satellite navigation systems are based on identical principles, and GPS and Glonass are virtually identical from the perspective of equations which we will define.

Since GPS provides a positioning fix, it can complement inertial navigation technique. Integration of both techniques achieves performance which is an improvement over what would be possible using either system alone.

There are many navigation signals available from the satellites and many different ways of processing them. These processing techniques give rise to trade-offs between the accuracy of the fix, how quickly it can be updated, and the range over which the accuracy can be maintained. Accuracies vary from tenths of a kilometer to tenths of a centimeter depending on the signals used and how they are processed. The systems can be used to measure position, velocity, attitude, and very precise time.

### 4.2.1 PRINCIPLES OF OPERATION

A constellation of satellites is maintained in earth orbit and radio receivers at or near the surface of the planet are used to decode the transmissions of the satellites, to compute from them the motion and position of the receiver. GPS was developed primarily as a military navigational aid, so it was designed as a one way broadcast system. Receivers do not have to transmit any signals back to the satellites that might give away their position.

This aspect of the design has a number of other implications. Satellites receive no signals, so there is no satellite capacity limit on the number of users. Receivers transmit no signals, so anyone with knowledge of the communications protocol can use the system without being detected. Civilian use of GPS has been provided for in the design and there is no charge associated with satellite signal reception. The eventual number of civilian users is estimated to be in the millions.

#### *POSITION MEASUREMENT*

Positioning is based on the principle of range triangulation. The receiver needs to know the range to the satellites that are being used and the positions of these satellites. It can determine its own position from only this information. As in all triangulation situations, the number of independent observations required depends on how many dimensions of position are required (2D or 3D) and how much information is already known.

The two dimensional case is indicated in Figure. Two satellites are required in general. If the positions of the satellites are known, and the ranges from the receiver to the satellites can be measured, the receiver position is constrained to be one of the two points of intersection of the circular lines of position. After the two simultaneous equations are solved, a rough initial estimate, or a third satellite can be used to resolve which of the two solutions applies.

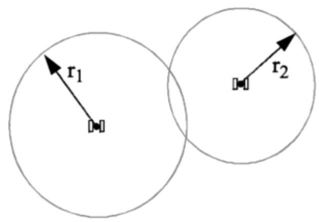


**Figure 4-10 2D range triangulation**

The receiver position is given by the simultaneous solution to the following two equations

$$r_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2}$$

$$r_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2}$$

Where the satellite position $(x_1, y_1), (x_2, y_2)$ and ranges $r_1, r_2$ are considered known quantities.

The three dimensional case is similar except that three satellites are used and the lines of position are spheres centered at the satellites. Two spheres intersect in a circle and another sphere intersects this circle in at most two places. The ambiguity can be resolved with another satellite or an initial estimate. For vehicles whose altitude is accurately known, only two satellites are required.

The satellite positions are known as the **ephemeris data.** This information is broadcast to the receivers from orbit. The second required piece of information, the ranges, are determined by measuring the time required for the propagation of the signal from each satellite to the receiver, and multiplying by the radio wave speed. To make this possible, each satellite also broadcasts the time along with its position. Then the receiver can subtract the time that the message was sent from the time that it was received.

The receiver performs the very difficult task of measuring a range of about 50,000 kilometers to the nearest meter or so. Therefore, an error of only one part in fifty million in either the propagation time or the speed of light will cause an error of one meter in the observed range. Prediction of the wave speed and tremendously precise measurement of time are both key design features of satellite navigation.

### *TIME MEASUREMENT*

In order for a receiver to compute the propagation time directly, it must have its clock synchronized with the clocks used on the satellites, and all of the satellites must be similarly synchronized among themselves. This is because every nanosecond of time error causes one foot of range measurement error. It seems at first glance that all satellites and all receivers everywhere must all agree on the time to the nearest nanosecond.

For the satellites, this synchronization is actually achieved in practice through the use of very precise atomic clocks. However, it is impractical for a receiver to use an atomic time standard. Cheap crystal oscillators are used instead. These clocks are precise enough but not accurate enough, so they may be offset in time from the satellites. The solution to the problem is to consider time a fourth unknown in the ranging process, and to use a fourth satellite to provide an extra constraint equation. If altitude is known accurately (as it is on the sea) one less satellite is needed.

The time offset of the receiver is called the **user clock bias.** The effect of this bias is shown in Figure. The satellites are synchronized, so the user clock bias gives rise to ranges to all of them which are in error by the same amount. These erroneous ranges are called **pseudoranges.**

If the user clock is slow, considering the two dimensional case, the three ranges will create a roughly triangular region as shown. Offsetting the sides of the triangle by fixed amounts on all sides, the receiver can determine the clock bias as an extra unknown.

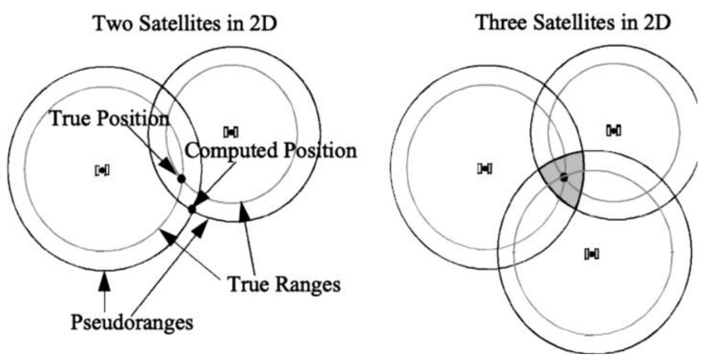


**Figure 4-11 Pseudoranges**

In three dimensions, the receiver position is given by solution of the following system of equations

$$r_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2 + (z - z_1)^2} + c\Delta t$$

$$r_2 = \sqrt{(x - x_2)^2 + (y - y_2)^2 + (z - z_2)^2} + c\Delta t$$

$$r_3 = \sqrt{(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2} + c\Delta t$$

$$r_4 = \sqrt{(x - x_4)^2 + (y - y_4)^2 + (z - z_4)^2} + c\Delta t$$

Although the equations are written in terms of geocentric cartesian coordinates, the answer can be converted into latitude, longitude, and altitude on any reference ellipsoid model of the earth, or into grid coordinates on a map.

### *WAVE SPEED MEASUREMENT*

The measurement of wave speed, or equivalency atmospheric delay, is accomplished in receivers in one of two different ways. Some receivers use a mathematical model of the atmosphere to compute the delay. The coefficients in this model are broadcast regularly by the satellites themselves.

Since the effect depends in a known way on frequency, some receivers are able to actually measure the difference between each of the two navigation signal carriers transmitted by each satellite. This technique is far more accurate than die use of a mathematical model. Atmospheric delay is one of die most significant sources of error.

#### *VELOCITY MEASUREMENT*

The principle used to measure velocity is the Doppler frequency shift. The Doppler shift for the frequency of each satellite is a direct measure of the relative velocity of receiver and satellite along the line between them. Each satellite has a very high velocity relative to a stationary receiver because of both the orbital motion of the satellite and the rotation of the receiver with the earth. The velocity solution is obtained by differentiating the four dimensional navigation solution. These are the equations solved in the receiver to determine velocity.

#### *ORIENTATION MEASUREMENT*

The principle used to measure orientation of a rigid body is to measure the differential three dimensional positions of different points on the body. In practice, measurements of the radio carrier phase are used to achieve the positioning accuracies required for acceptable orientation measurement. Very accurate real-time orientation measurement for rigid vehicles has been achieved.

The accuracy of angular measurement increases with the length or **baseline** distance between the points positioned. Using four satellites, the user clock error and the 3D position of each point can be calculated. If baselines are known accurately enough, less satellites are required. At least three points must be measured to determine all three angles of rotation.

### 4.2.2 IMPLEMENTATION

#### *SIGNAL CHARACTERISTICS*

GPS was intended to meet both a military and a civilian demand for global all weather navigation. Satellite radio sources are particularly suited for this purpose since they can radiate their signals to every point on the surface of the planet. It is intended to be a multiple simultaneous access positioning utility, available to as many users as there are receivers. Further, receivers are required to be small enough to carry in one hand. The signals are intended to provide highly accurate measurements of both position and velocity while at the same time being easy to acquire quickly.

#### *CARRIERS*

GPS signals are, like most radio signals, a modulated carrier signal. The carrier signals, being predictable sinusoids, cannot carry information themselves. Information is carried by the modulation signals which are mixed with the carriers.

It is known that the radio wave delay through the atmosphere varies with time and position by as much as 100 meters in equivalent range. Mathematical models do not remove most of this error, so the design includes two separate L band carriers, denoted LI and L2. Atmospheric delay varies with the square of the carrier frequency. With two carriers, the differential delay of both can be used to measure the absolute delay.

The choice of L band carriers represents a trade-off between the limited available bandwidth in UHF and the excessive atmospheric absorption, called space **loss, of** C band. The L band frequency is high enough to accommodate the 20 MHz bandwidth required and the centimeter wavelength enables precision velocity measurement.

### *MODULATION SIGNALS*

The requirement for small receivers translates into a requirement for small antennae on the receiver, which in turn means that they must work at very small signal levels. In order to achieve this, both modulation signals are based on a particular kind of digital pseudorandom noise code called a Gold code. These **PRN codes** allow receivers with antenna only a few inches across to extract very low power signals from the noise by correlating them with expectations.

PRN codes have two other advantages. They are particularly suited to the multiple access requirement. Each satellite transmits a different code, so the receiver can distinguish between the signals of all satellites which arrive at its antenna simultaneously. Also, PRN codes are particularly impervious to deliberate or unintentional jamming, so they support the need for secure military access.

There are two modulation signals used. The civilian signal, called C/A for clear access or coarse acquisition, has a 300 meter pulse wavelength. The military signal is designated P for precise, and is called Y when it is encrypted. It has a pulse wavelength of 30 meters. Pseudorange resolution is directly related to the wavelength of the code, so the civilian code has a lower precision. The civilian C/A code is mixed with the LI carrier while the P code is mixed with both carriers and broadcast redundantly.

The use of two codes, one of which is encrypted, gives the military access to a completely separate, secure, positioning system. A second reason for using two codes is that, since the P code is difficult to acquire, the C/A was designed for easy acquisition.

### *NAVIGATION MESSAGE*

In addition to the PRN codes, a second block of information is digitally broadcast on both carriers. This **navigation** message is transmitted at the very slow rate of 50 bits per second, and is updated and repeated every 12.5 minutes. The navigation message contains a large amount of information provided to assist receivers in their operation.

The system time of the week, known as the **handover word,** is provided to assist in acquiring the P code after the C/A code is acquired. The transmitting satellite provides precise ephemeris data on its own orbit for use in the navigation solution. Also, less accurate ephemeris data for all other satellites, known as the **almanac,** is transmitted to permit receivers to predict when new satellites will rise above the horizon and become available. The ephemeris data is described by the classical ellipse parameters discovered and defined by Kepler.

Atmospheric delay model coefficients are provided since the magnitude of the delay varies significantly throughout the day. Satellites also provide information on their own health and the precision of range measurements that can be expected if their signals are used. In this way, receivers can pick an optimal set of satellites in view for their computations. The complete set of GPS modulation signals is summarized in Figure:

| | C/A | P | NAV |
|---|---|---|---|
| L1 | ✔ | ✔ | ✔ |
| L2 | | ✔ | ✔ |

**Figure 4-12 GPS modulation signals**

### *USER SEGMENT*

The GPS user segment consists of the receivers in use on the surface and in the near space region. The basic functions of the receiver are to perform code correlation in order to measure propagation tune and clock bias, to perform compensation for known errors, and to implement the solution to the navigation equations. Receivers internally generate exactly the same codes as the satellites. Time of arrival measurement is accomplished by shifting the internal codes in time and computing the correlation with the externally received codes as shown in Figure.

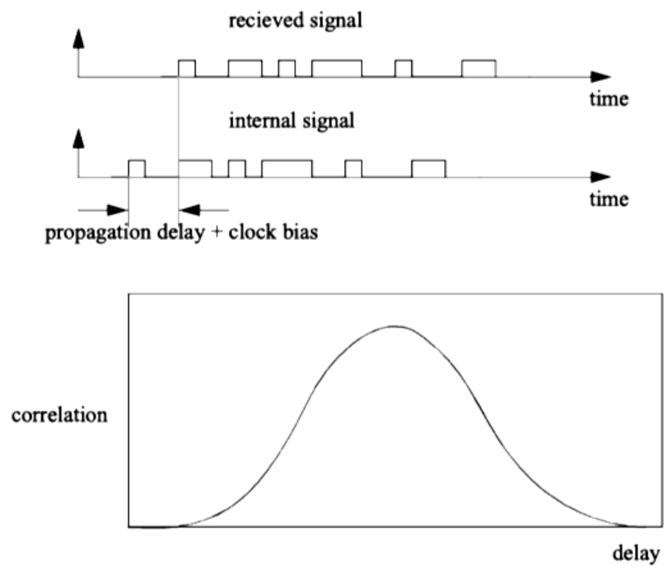


**Figure 4-13 GPS signal correlation**

The C/A code repeats once every millisecond while the P code period is one week long. The C/A code is short, so the amount of shifting required is limited, and a match can be found quickly. A match occurs when the correlation function peaks. Once a match is computed, the code is said to be acquired. Once acquired, the time of signal propagation is given by the time shift that was necessary. This time includes the clock bias which is removed later. The next step is to derive the wave speed.

Receivers can compensate for atmospheric delay to compute an accurate wave speed by using either mathematical model coefficients transmitted by the satellites, or by using two carriers. Multiplying this by the propagation time gives the pseudorange to the satellite. When this process is repeated for four different satellites, the receiver can solve the navigation equations and determine its position.

### *SPACE SEGMENT*

The GPS space segment consists of a constellation of 18 earth satellites and three spares that circle the earth in each of six orbits as shown in Figure. The six orbital planes are each inclined at 55° to the equatorial plane, and separated by 60° in longitude. Orbits are nearly circular, 11,000 miles in altitude and repeat exactly twice per **sidereal** day. This implies that the satellites are in exactly the same place four minutes later every day.

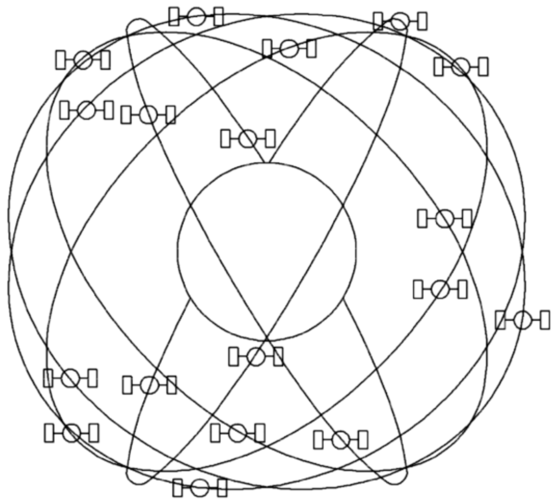

**Figure 4-14 Satellites orbits**

The constellation is designed to ensure that at least four satellites are visible at any time. Satellites are considered visible above the nominal **mask angle** of 5°-10°. When the elevation angle above the horizon exceeds the mask angle, the signals are strong enough for good reception. However, foliage or other attenuation may increase the mask angle significantly.

Use of GPS requires a clear line of sight, and the signals cannot penetrate water, soil, or walls very well. Hence, GPS cannot be used underwater, in thick forest, or in mines and tunnels. Signals can be obstructed by tall buildings or by parts of a rotating vehicle if the receiver antenna has not been placed carefully.

### *GROUND SEGMENT*

The GPS ground segment consists of a series of five ground stations spaced in longitude around the globe. One of the stations is designated the Master Control Station (MCS). The function of GPS Master Control is to track the positions of all satellites very precisely and to maintain the overall system time standard.

Space vehicle positions are very predictable due to the negligible effect of atmosphere at orbital altitude. Nevertheless small perturbing forces exist which are caused by the oblateness of the earth, the presence of the moon, and even the pressure of solar radiation accumulated over time. The effect of these forces is to perturb the satellite orbits from their ideal Keplerian elliptical shape.

Inverting the ranging process, considering MCS to be the user and the ground stations to be the satellite, makes it possible to determine the position of each satellite very accurately. MCS updates the satellites on their orbital parameters at least once per day so that precise information can be later transmitted to the receivers.

Satellites must be synchronized in time to a precision of nanoseconds to avoid the range errors that arise from timing errors. To do this, MCS maintains an overall system time standard known as **GPS system time** through the use of highly accurate cesium atomic clocks. Also, each satellite incorporates two cesium and two rubidium atomic clocks for redundancy.

Atomic clocks are used since the more common crystal oscillators are not stable enough for this application. MCS regularly transmits to each satellite the offset of its clock from GPS system time. Satellite clock offsets are then retransmitted to receivers so that the precise GPS system time of signal emission is known. Overall constellation control is indicated in Figure.

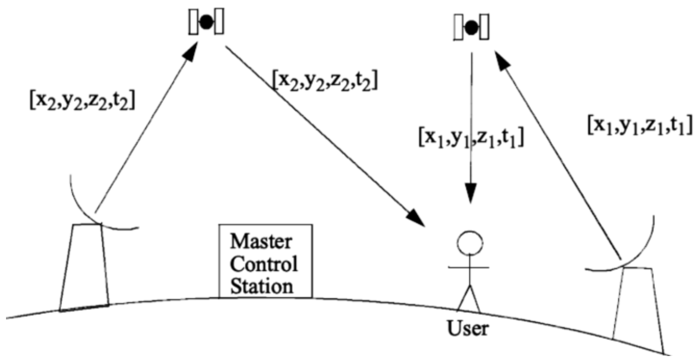


**Figure 4-15 GPS constellation control**

## *PERFORMANCE*

The performance of GPS can be characterized along three independent dimensions: the precision and accuracy of the fix, the rate at which it is updated, and the range over which the accuracy can be achieved.

If GPS is used to determine vehicle position in real time for navigation purposes, high accuracy must be traded against the update rates necessary for real time control. Further, if the vehicle's range of motion is restricted by the application, differential techniques can be used to improve accuracy considerably, at the cost of a second receiver, a second communications link, and the limited range.

The accuracy achievable at any update rate is also a function of the signal combinations used and the processing performed on them. The inherent resolutions of the C/A, P, and carrier signals vary from meters, to decimeters, to millimeters. It is not the intention here to provide a quantitative assessment of GPS accuracy in any of its modes of use. Rather, this section will equip the reader to understand the factors that affect accuracy to a degree that will make informed design decisions

## *RANGE RESOLUTION*

The fundamental range resolution available in any mode of operation is related to the signals used. Specifically, achievable range resolution is a fraction of the signal wavelength. The signal wavelengths are summarized below:

| Signal | Frequency | Wavelength |
|---|---|---|
| L1 | 1575.42 MHz | 20 cm |
| L2 | 1227.60 MHz | 24 cm |
| C/A code | 1 MHz | 300 m |
| P code | 10 MHz | 30 m |

**Table 4-1 Signal wavelength**

### *MEASURES OF ACCURACY*

There are two measures of accuracy that are commonly used to describe GPS performance. The military commonly uses the Circular Error Probable (CEP) in two dimensions, and the Spherical Error Probable (SEP) in three dimensions. These measures are defined to be the median value of accuracy. Hence 50% of measurements can be expected to fall within the SEP of the true value and 50% will fall outside.

The scientific community prefers the standard deviation as a measure of accuracy. In GPS circles, the 2drms value or twice the standard deviation is often used. Between 95% and 98% of measurements can be expected to fall within the 2drms value of the true position, depending on the ellipticity of the distribution. Both measures of accuracy are related by the approximate formula:

$$2 \times drms \cong 2.5 \times CEP$$

### *SELECTIVE AVAILABILITY*

The nominal precision of GPS service is precisely regulated and deliberately controlled by the Department of Defense. Two levels of positioning service are distinguished for GPS. The standard positioning service (SPS) is designated for civilian use and has a nominal precision of 120 meters 2drms in 3 dimensions. SPS is based on reception of the C/A PRN code. Velocity measurement is intended to be accurate to 0.3 meters/second, and time is accurate to 300-400 nanoseconds.

The precise positioning service (PPS) is designated for US and NATO military use and has a nominal precision of 10 meters 2drms in three dimensions. PPS is based on the simultaneous reception of the encrypted P code on both carriers. Only those users who have knowledge of the encryption will have access to the higher PPS precision.

Through a technique known as selective availability (S/A) the physically achievable 40 meter precision of SPS is deliberately degraded to the 120 meter nominal value. This is achieved by artificially introducing errors into the broadcast ephemeris data, and by dithering the satellite clock. As a result, S/A errors have both a slowly varying and a rapidly varying component. For a fixed receiver, measured pseudorange will vary apparently randomly over time as shown in Figure.

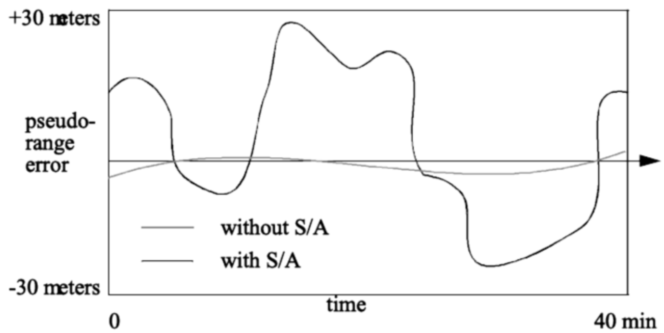


Figure 4-16 Pseudorange in function with time

### 4.2.3 SOURCES OF GPS ERROR

In order to assess the possible accuracies available in an application, an understanding of the sources of error and when they apply is necessary. The total error in the positioning fix can be separated into two components: the errors in the satellite pseudoranges, and the projection of these errors onto the navigation coordinate system. The first component involves physics; the second is purely a mathematical transformation.

Pseudorange error sources are additive. The geometric dilution of precision or GDOP is multiplicative. GDOP can be as high as 4 or much more under certain circumstances, so it is important to know whether precision estimates include GDOP or not, and if they assume a particular value. It is also important to ascertain whether a particular precision estimate is horizontal or 3D.

#### *SOURCES OF PSEUDORANGE ERROR*

The significant sources of pseudorange errors and their nominal values are summarized in Table

| Error Source | Nominal Value (rms) |
|---|---|
| **Selective Availability** | 8m |
| **Atmospheric Delays** | 4m |
| **Satellite Clock & Ephemeris** | 3m |
| **Multipath** | 1-3m |
| **Receiver Electronics & Vehicle Dynamics** | 1.5m |
| **TOTAL** | 10m |

Table 4-2 Pseudorange errors

It should be clear from this table that the GPS design reserves the ability to remove the two most important errors for exclusive military use. With the exception of S/A, the deliberate signal degradation discussed earlier, atmospheric delay is the most important pseudorange error source. It is also called **group** delay. It varies with time and place. It can result in as much as 30 meters equivalent range error when the satellite is at **zenith.**

Atmospheric delay consists of two components: ionospheric and tropospheric. The ionospheric delay is caused by diffraction due to charged particles. It can vary by a factor of 5 from day to night, being highest during the day, and by a factor of 3 due to the elevation angle of the satellite. This means that each satellite has a different, time varying pseudorange error. It depends on solar magnetic activity and geomagnetic latitude, being particularly affected during magnetic storms and it is usually greatest at the poles, and at the equator.

Tropospheric delay is caused by water and other atmospheric constituents giving rise to local changes in the index of refraction. It can be modeled very well by a simple mathematical formula. It also varies with time and place, giving about 2.3 meters delay vertically and ten times this at the horizon.

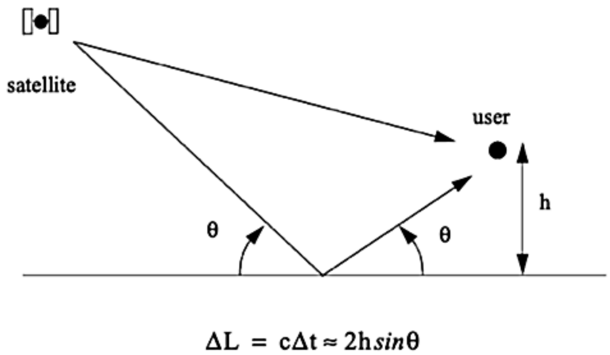


**Table 4-3 Multipath error**

Multipath error arises when some of the radio signal reaches the receiver through reflection in addition to the portion which arrived along the direct line of sight. The result of both signals arriving at the antenna is destructive interference. This error source can be substantial above water, since water is a perfect radio reflector under certain conditions, but it is usually less than 1 meter. As indicated in Figure, the time delay for ground reflection depends on altitude. Multipath error is more pronounced when the receiver is close to the surface. Therefore, it is usually advisable to mount GPS antennae as high as possible.

Other less important sources of error include the unintentional component of satellite ephemeris and clock errors, imperfections in receiver circuitry, computational approximation and truncation, and vehicle dynamics. The effect of vehicle motion can be substantial if receiver specifications are exceeded.

### *GEOMETRIC DILUTION OF PRECISION*

The final GPS positioning accuracy is strongly dependent on the geometry of the satellite lines of position used. The transformation from four pseudorange errors into positioning errors is accomplished by multiplying by a geometry factor known as the geometric dilution of precision. This GDOP occurs in all triangulation problems and is defined as:

$$DOP = \frac{FixError}{RangeError}$$

where the errors are represented as standard deviations. The DOP scales the range error into the fix error, so small values indicate favorable geometry.

The concept is easy to visualize in two dimensions. Consider Figure below. Range errors, shown as shaded regions, combine to form a roughly rectangular region of uncertainty in the fix. The greater the angle between the satellites used, the smaller is the region of uncertainty. It can be formally shown that there is a direct relationship between uncertainty and the volume of this region.

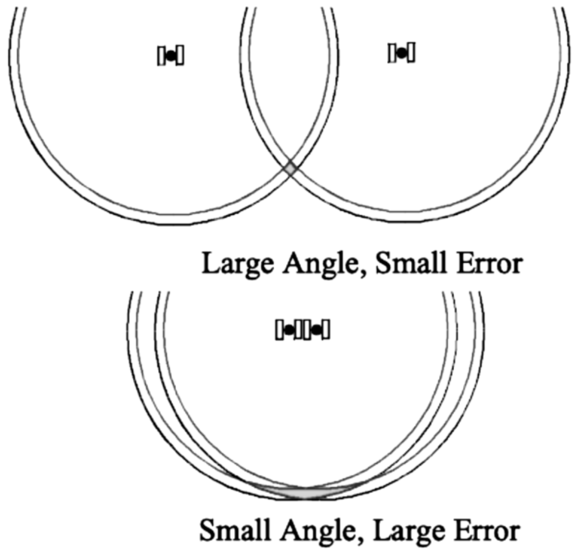


**Table 4-4 DOP errors**

It the concept of GPS, five terms are defined

- TDOP – time dilution of precision
- PDOP – position dilution of precision (3D)
- HDOP – horizontal dilution of precision
- VDOP – vertical dilution of precision
- GDOP – geometric dilution of precision

$$GDOP = \sqrt{PDOP^2 + TDOP^2}$$

$$PDOP = \sqrt{HDOP^2 + VDOP^2}$$

## 4.3 INTEGRATIONS OF DIFFERENT NAVIGATION ALGORITHMS

| INS | DGPS |
|---|---|
| • High position and velocity accuracy over **short term**<br>• **Accurate attitude information**<br>• Accuracy decreasing with time<br>• **High measurement output rate**<br>• **Autonomous**<br>• **No signal outages**<br>• Affected by gravity | • High position and velocity accuracy over the **long term**<br>• Noisy attitude information (multiple antenna arrays)<br>• **Uniform accuracy,** independent of time<br>• **Low measurement output rate**<br>• **Non-autonomous**<br>• Cycle slip and loss of lock<br>• Not sensitive to gravity |
| INS/DGPS<br>• High position and velocity accuracy<br>• Precise attitude determination<br>• High data rate<br>• Navigational output during GPS signal outages<br>• Cycle slip detection and correction<br>• Gravity vector determination | |

### *KALMAN FILTERING BASICS*

Within the significant toolbox of mathematical tools that can be used for stochastic estimation from noisy sensor measurements, one of the most well known and often-used tools is what is known as the *Kalman Filter*. The Kalman filter is named after Rudolph E. Kalman, who in 1960 published his famous paper describing a recursive solution to the discrete-data linear filtering problem. The Kalman Filter is essentially a set of mathematical equations that implement a predictor-corrector type estimator that is *optimal* in the sense that it minimizes the estimated *error* covariance—when some presumed conditions are met. Since the time of its introduction, the *Kalman Filter* has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation.

Basically, Kalman Filter is a special case of sequential least square estimation, where initial measurement is equal to the first measurement and design matrix of the second position is an identity matrix. Kalman Filter also takes the velocity as an unknown parameter.

# 5 GNC TESTING AND EVALUATION

## 5.1 LABORATORY TESTING

### 5.1.1 CALIBRATION

Sensor calibration is a method of improving sensor performance by removing structural errors in the sensor outputs. Structural errors are differences between a sensors expected output and its measured output, which show up consistently every time a new measurement is taken. Any of these errors that are repeatable can be calculated during calibration, so that during actual end-use the measurements made by the sensor can be compensated in real-time to digitally remove any errors. Calibration provides a means of providing enhanced performance by improving the overall accuracy of the underlying sensors.

The calibration process consists of placing the DUT (device under test) into configurations where the inertial input stimuli for the sensor is known, thus allowing us to determine the actual error in each measurement. The sensors examined are the sensors that are typically found on inertial sensor packages, consisting of accelerometers, gyroscopes, magnetometers...

#### CALIBRATION OF IMU

> *Calibration procedure should produce transformation matrix compensating for IMU axis misalignment, DC offset, gravity dependent drift and earth rotation.*

#### CALIBRATION OF GYRO STABILIZED HOMMING HEAD

> *The main purpose of Homing Head Calibration procedure is to synthesize calibration polynomials that convert data measured from sensor to Homing Head angular rates. It is critical that the Homing Head is mounted in manner that it later will be installed in rocket.*

### 5.1.2 SOFTWARE IN THE LOOP (SIL) TESTING

> *The main purpose of SIL testing is to verify complete software of GNC unit.*

In software-in-the-loop (SIL) phase, the actual Production Software Code is incorporated into the mathematical simulation that contains the models of the Physical System. This is done to permit inclusion of software functionality for which no model(s) exists, or to enable faster simulation runs.

Purpose:

- enable the inclusion of control algorithm functionality for which no model exists
- increase simulation speed by including compiled code in place of interpretive models
- verify that code generated from a model will function identically to the model
- guarantee that an algorithm in the modeling environment will function identically to that same algorithm executing in a production controller

### 5.1.3 ACTUATORS TESTING

> *The main purpose of ACT test is to verify work of actuator system. During actuator testing dynamic o actuator will be verified and dead zone, backslash and other mechanical errors can be revealed.*

### 5.1.4 HARDWARE IN THE LOOP TESTING (HIL)

> *The main purpose of HIL test is to verify roll, pitch and yaw autopilots algorithm, navigation & guidance algorithm with actuator system involved and rates from actual IMU sensor, in conditions that simulate real flight conditions.*

Hardware-in-the-loop (HIL) simulation, or HWIL, is a technique that is used in the development and test of complex real-time embedded systems. HIL simulation provides an effective platform by adding the complexity of the plant under control to the test platform. The complexity of the plant under control is included in test and development by adding a mathematical representation of all related dynamic systems. These mathematical representations are referred to as the “plant simulation”. The embedded system to be tested interacts with this plant simulation.

## 5.2 MISSILE DESIGN EVALUATION & FLIGHT TESTING

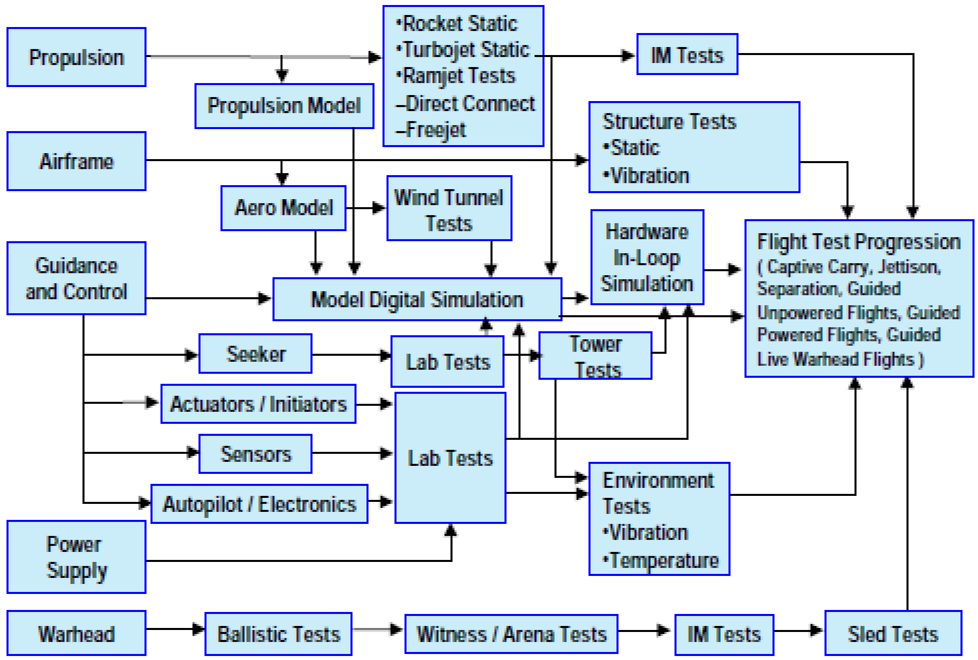


**Figure 5-1 Missile design evaluation block diagram**

Flight test validation is a progressive activity of increasing complexity. The objective of progressive testing is to minimize risk and enhance safety in the flight test activity. A typical progression of flight testing begins with captive carry and ends with live warhead launches. Intermediate tests are store jettison tests, safe separation tests, unpowered guided flights with an inert warhead, powered guided flights with an inert warhead, and finally, all-up powered guided flights with a live warhead.