History of Key Technologies

The Development of Radar Homing Missiles

Mike W. Fossier
Raytheon Company, Bedford, Massachusetts

Introduction

SINCE World War II, guided missiles have played an increasingly important role in warfare. Earlier papers in this series[1,2] reviewed the development of the inertial guidance systems that made possible the accurate delivery of long-range ballistic (and other) missiles for which the target is a known set of Earth coordinates. These systems are not suitable for guidance of missiles against unpredictable targets such as maneuvering aircraft, which requires an ability to sense the target location in real time and respond to rapid changes. To accomplish this, modern air defense missiles use homing guidance, in which an on-board sensor provides the target data on which the guidance is based. Because of the continually improving quality of target information as the missile closes in, homing guidance provides an accuracy that is unmatched by any other form of missile guidance.

This paper is not a comprehensive history of its subject. Rather, it is a summary of the author's first-hand experience, covering the development of continuous wave (CW) radar homing guided missiles, which were pioneered at Raytheon Company. They form an impressive subset of all homing missiles, with nearly 100,000 air- and surface-launched versions produced to date and deployed in some two dozen countries worldwide. Derivatives of these designs form the backbone of the U.S. Navy's ship-based defenses and similar designs are extensively used by the Soviet Union. In reviewing the history, technical descriptions of problems encountered and solutions developed will be given for the three key areas of sensor, navigation, and flight control.

Early History

While the V-1 and V-2 were the missiles with significant military impact on World War II, Germany also did extensive research on missile guidance against targets with unknown ground coordinates, including ships and aircraft. They developed the basic equations and principles of "proportional navigation," which is still in use today in most homing missiles. In proportional navigation, only the angular direction to the target (called the line-of-sight) is employed to accomplish the intercept. The principle is, in a sense, the inverse of avoiding a collision at sea. Novice sailors are taught to steer to cause the sight line to other boats to change in order to avoid collision. If the line of sight does not rotate (in inertial space), a collision will occur eventually (Fig. 1). Proportional navigation operates to force that collision to occur by measuring the rate of change of the line-of-sight angle with time and by commanding a rate of change in the missiles's velocity vector in the direction to reduce the rate at a level proportional to that rate. That is,

$$\dot{\gamma}=N\dot{\lambda} \tag{1}$$

where λ is the spatial line-of-sight angle to the target, γ the angle of missile velocity vector, and N the "navigation ratio."

Fortunately for the Allies, Germany did not succeed in deploying a successful antiaircraft missile during World War II. However, a number of German scientists came to the United States after the war and applied their experience to the fledgling developments here. The principal impetus to U.S. guided missile development near the end of World War II was the horribly effective Kamikaze attacks against our ships. The U.S. Navy initiated the development of the "Lark" missile in 1944 (Fig. 2) with a hoped-for development schedule of six months. That schedule was not realized.

The sensor for the Lark was to be an active CW radar on board the missile. That sensor, in fact, did succeed in guiding

Mike W. Fossier was born in New Orleans in 1928. He attended Louisiana State University, receiving a B.S. in Mechanical Engineering in 1945. He subsequently received an M.S. and Professional Degree in Aeronautics from the California Institute of Technology in 1946 and 1947. He worked as an aerodynamicist for Douglas Aircraft Co. in 1946-1950, specializing in the stability and control of high-speed aircraft. In 1950 he joined Raytheon's missile flight test operation at Oxnard, Calif., where he performed analysis and design of the guidance and control of homing missiles, including many of the fundamental issues discussed in this paper. In 1954 he transferred to the Company's missile R&D center in Bedford, Mass., to participate in the preliminary design of the Hawk missile system. In 1956 he became Project Engineer for the Sparrow missile and supervised the transition from R&D into production. In 1959 he became Chief Engineer of the Missile Systems Division, responsible for all technical activities in the Division. In 1965 he was elected a vice-president of Raytheon and became Assistant General Manager—Technical of Missile Systems Division, a position he continues to hold.

Received Oct. 6, 1983; revision received Feb. 21, 1984.

EDITOR'S NOTE: This manuscript was invited as a History of Key Technologies paper. It is not meant to be a comprehensive study of the field. It represents solely the author's own recollection of events at the time and is based upon his own experiences.

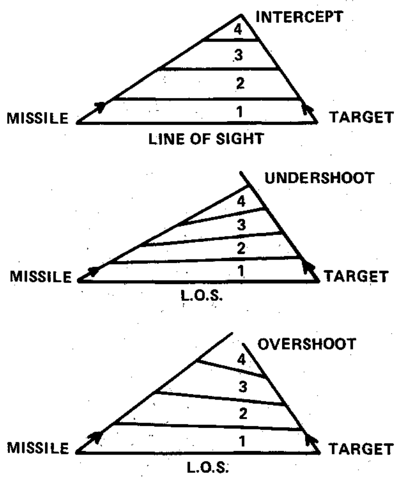

Fig. 1 Line-of-sight motion of intercept.

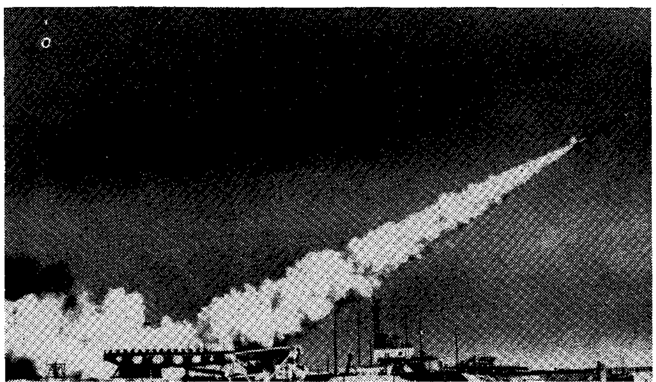

Fig. 2 Lark missile.

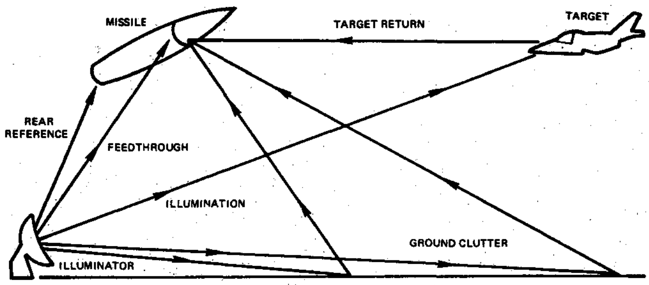

Fig. 3 Signals received by CW homing missile.

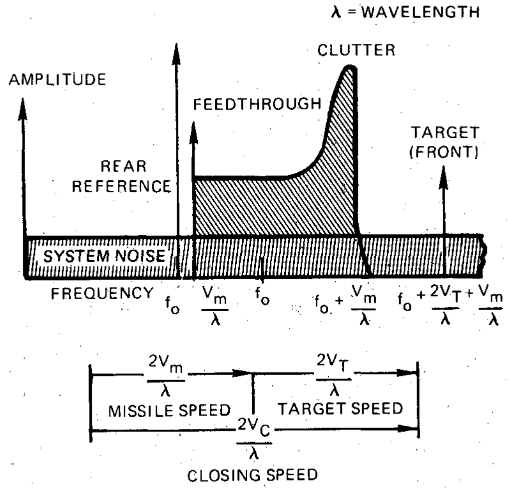

Fig. 4 Missile signals: amplitude vs frequency.

a Lark to the first successful intercept ever made by a missile against a (pilotless) aircraft on Dec. 2, 1950.

In any discussion of CW radar it is necessary to begin with Roy Sanders, the man with the vision of what CW radar could do and who organized and inspired the work at Raytheon for the first seven years.[3] Like many other visions, it was not perfect in detail. The active CW radar that he envisioned eventually flew in a half dozen experimental missiles in 1950 and 1951 and then sank quietly into oblivion. However, its direct descendants, the semiactive CW radar and the semiactive and active pulse-Doppler radars, came to completely dominate antiaircraft missile systems. While these were perfected by others, they were all based on Sanders' pioneering work.

The choice of CW radar is worth some discussion. The Kamikaze aircraft usually attacked from low-altitude flight. When a missile directs its radar energy to the aircraft target, only a very small fraction of this energy reaches the target and is reflected, and an even smaller fraction of that is intercepted by the missile antenna as a signal. When the missile looks down at an aircraft, it also "sees" the surface of the Earth, which reflects a much larger signal than an aircraft (Fig. 3). Unless some way can be found to distinguish between them, the ground clutter (or sea clutter), in this case) hides the small target signal.

A CW radar transmits a pure tone of microwave energy. Reflections from targets in its beam will be displaced in frequency by the Doppler effect. In principle, all the missile has to do is select the frequency corresponding to the moving aircraft target and spectrally reject the frequency that corresponds to the stationary ground (Fig. 4). It took a while, however, to understand and overcome some fundamental problems in doing this.

First, the Doppler shift is not very large—20 Hz/ft/s of relative velocity at a radiated frequency of 10 GHz (10^{10} Hz). To separate the target signal from the interference, a frequency resolution of about 10^3 Hz (50 ft/s) is desired. The implementation of this narrow-band filter was accomplished by beating the received signal down to video frequencies, using a sample of the transmitted signal as a reference for the incoming signal. This spectrum was analyzed automatically by heterodyning various parts of the spectrum sequentially into a highly selective fixed-frequency 1 KHz wide filter, using a variable frequency oscillator as the mixer reference. The selection of the center frequency of this filter was originally based on the availability of coils from Radio Shack. While environmental sensitivity eventually forced the engineers to design their own, the frequency remained the same for decades.

There remained two major problems. The first of these is termed "feedthrough." A CW radar transmits at the same time that it is trying to listen for a received signal on the order of 10^{16} times smaller than the transmitted signal. This problem has been likened to that of trying to hear a pin drop a mile away while yelling at the top of your lungs. Obviously, extreme care must be taken to prevent any of the transmitted power from leaking into the receiver. By careful design, this eventually proved to be possible to do to an adequate degree in the absence of environmental factors, but vibration-induced modulation of the feedthrough aggravated the problem significantly. In practice, the active seeker developed for the Lark could eventually track an aircraft to a range of 4 miles when the seeker was mounted on the ground; in "captive flight" on an airplane, the range was reduced to 2 miles; and in free missile flight, a range of about 1 mile was achieved.

The second problem, generically similar but different in detail, is due to the magnitude of the clutter return. If the transmitted signal is not a pure tone, its frequency side bands will modulate the large clutter signal, resulting in components at frequencies within the receiver bandwidth. Modulations at a frequency equal to the difference frequency between the target and clutter Doppler frequencies will modulate the clutter signal to create a side band at the same frequency as the desired target signal. Since the clutter power may be as high as 10^7 times that of the desired target signal, the noise on the transmitted signal (as well as noise introduced on the received signal from any source) must be severely limited. Therefore, the development of low-noise microwave sources became a critical part of CW radar development and remains so to this day.

While factors such as these were debated at the conceptual level and governed the actions of the designers who were trying to successfully fly missiles, the battle was really decided in the trenches of equipment reliability. Missiles are characterized by a combination of a hostile environment and a requirement that all hardware must operate within acceptable limits without the benefit of human intervention. In the late 1940s, the electronics state-of-the-art was not ready for this and much had to be learned before success could be achieved.

There eventually developed a rigid discipline in missile checkout and flight operations that came to be adopted universally. In the case of Lark, it was decided that no missile would be flown before its time—specifically, until the flight hardware had passed 10 sequential hours of successful testing. When a reliability failure occurred, the clock was restarted. A key element in this test was a ground simulator, informally known as the "Bug" (Fig. 5). The Bug was a rebuilt automobile frame (a Ford, as I recall) on which the missile seeker electronics were mounted. The seeker antenna gimbal axis was coincident with the axis of the steered wheels. The wheels were steered in response to the real-time guidance commands to simulate the missile motions in flight (scaled down by two orders of magnitude), allowing a two-dimensional intercept to be made against a simulated moving target (Fig. 6) with an appropriate (simulated) Doppler frequency. This early "hardware-in-the-loop" simulator built confidence in the guidance accuracy of the design as well as its reliability.

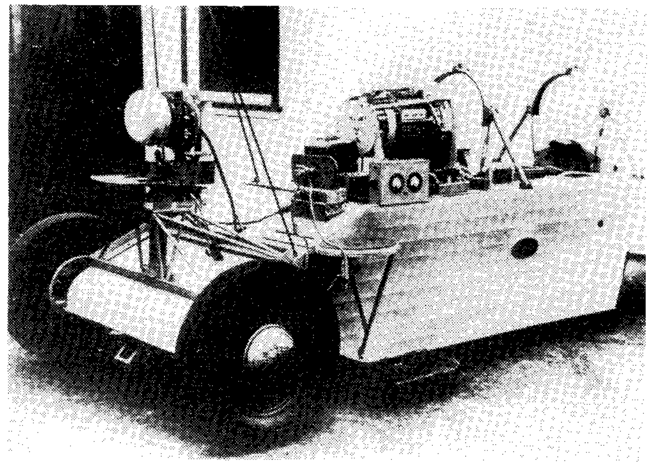

Fig. 5 Hardware in the loop missile simulator.

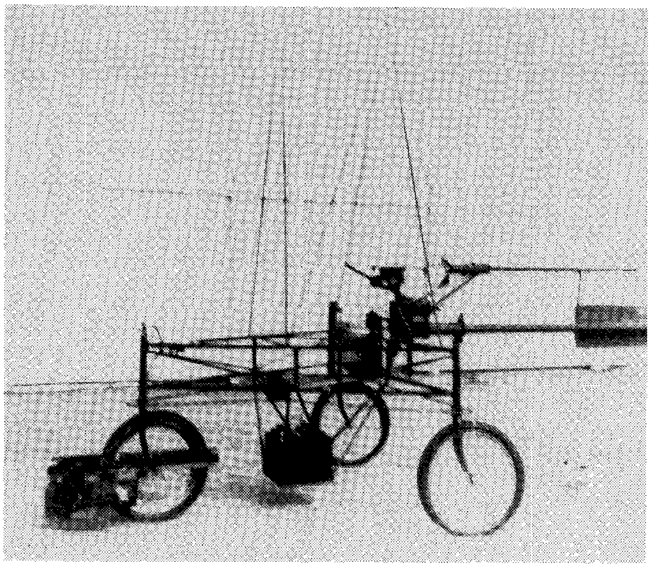

Fig. 6 Moving target simulator.

The 1950s

The December 1950 flight ushered in the crucial decade in homing missile history. By the end of that decade, both air-to-air and surface-to-air homing missiles were operationally deployed and the pattern for future developments was firmly set. The major steps taken in that decade are reviewed here.

An active seeker—that is, one in which the transmitter as well as the receiver is carried in the missile—has the advantage of being able to operate autonomously after the missile is launched. However, achieving the necessary isolation of the receiver from the transmitter in an active CW seeker was too difficult a task for an early successful solution. Accordingly, shortly after the successful flight described above, the Navy opted for a semiactive approach—that is, one in which the transmitter remains at the launch point and only the receiver is flown in the missile. Further, it selected an air-to-air application and contracted with Raytheon in June 1951 to develop a version of the Sparrow missile family using the semiactive seeker. The Sparrow family then consisted of Sparrow I with radar "beam rider" guidance, Sparrow II with an active pulse radar homing seeker operating at a much shorter wavelength, and the semiactive CW Sparrow III. A common airframe was the unifying element of the family.

Sparrow

In approaching the development of Sparrow III, a number of problems of immense difficulty presented themselves. The 8 in. diameter body of Sparrow contained much less packaging volume than the 20 in. diameter of Lark, creating a need to redesign the seeker that had been developed so laboriously. Fortunately, the removal of the transmitter from the missile proved to be very straightforward, and the use of newly available "subminiature" vacuum tubes with the necessary electrical characteristics made the packaging problem solvable. While not easy, the seeker design was at least a continuation of the work that had been going on for the previous 7 yr.

The problems of navigation and control, however, were almost totally new, requiring a series of innovations across a broad spectrum of technologies. The Lark airframe used a liquid-propellant rocket engine, which sustained the missile at a nearly constant velocity of about Mach 0.8. Both the wings and tails of the missile were controlled to maintain zero angle of attack during maneuvers, with lift provided by the incidence angle of these surfaces. The nearly constant speed and altitude of the flight tests permitted the achievement of an angle of attack very close to zero. In turn, this made it possible to approximate the inertial line-of-sight measurements with body-fixed measurements. The known head-on aspect of attack against a fixed (and relatively slow) velocity F6F drone target contributed to the ease of the intercept problem, and it proved to be possible to "can" all parameters of the early intercepts using precalculated values.

With Sparrow, the problem was much more complex. The missile was boosted in about 2 s to a velocity more than Mach 1 greater than the (variable) launch velocity and it then glided without thrust to the intercept point, slowing down all the way. Only the wings were controlled, resulting in angles of attack up to 10 deg. "Around-the-clock" aspects of attack were required to cope with the tactical needs against a wide range of target velocities and maneuvers. Finally, intercept altitudes from sea level to (initially) 60,000 ft were necessary, at the pleasure of the intended target.

Let us examine the way in which these problems were solved. First, it was clear that proportional navigation should be used—it was all we knew and extensive analysis performed by Hughes Aircraft Company[4,5] showed that it could achieve the required accuracy. But as we began to understand all of its implications, it quickly became apparent that the navigation ratio N used by the Germans was not the proper gain term. In the easily derived linear differential equation representing the engagement, the missile's lateral acceleration history is invariant, not with N, but with what we called N', or effective navigation ratio, defined by

$$N' = N(V_M/V_C)$$

where V_M is the missile velocity; V_C the closing velocity with the target, $-\dot{R}$; and R the missile to target range.

As shown in the examples of Figs. 7 and 8 for an initial heading error and a target maneuver, the lateral acceleration history is proportional to the disturbance, with a shape that depends only on N'. Miss distance can be determined analytically for these cases for a simplified two-time lag control system as shown in Figs. 9 and 10.

From analyses of this type, it was concluded that N' should be kept in the range of 3-4 to avoid saturation in the end game, and that when this is done a time of flight equal to at least 10 time constants is sufficient to reduce the miss caused by any disturbance to a very small value, assuming no saturation of missile maneuvering capability.

We were concerned of course, about the validity of linearizing the problem in this way. However, what we discovered in time was the threshold nature of homing missile guidance. So long as the target maneuver (or any other disturbance) is below the level that demands more maneuver from the missile than it can deliver, the miss distance will be quite small. As soon as saturation occurs, however, the miss distance quickly escalates to intolerable levels. The designer, then, makes linear operation of the guidance system a major objective. When this is achieved, the linearized analysis is quite accurate.

The input to the proportional navigation guidance law is the time rate of change of the line-of-sight angle to the target $\dot{\lambda}$. A spatial reference is required, which in the other U.S. homing missiles of the 1950s (the Hughes Falcon and the

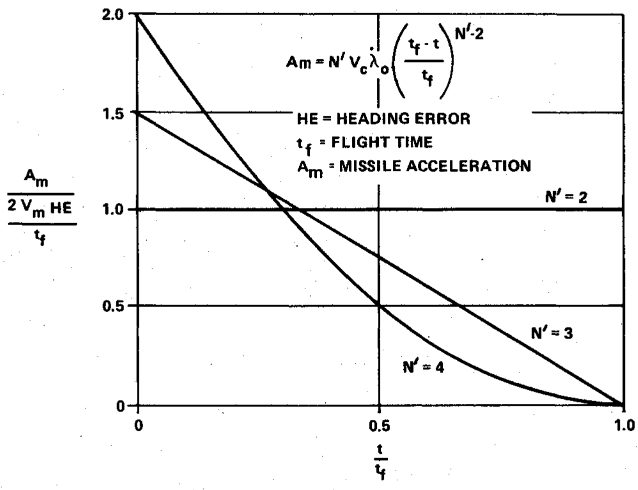

Fig. 7 Missile acceleration required to overcome initial heading error.

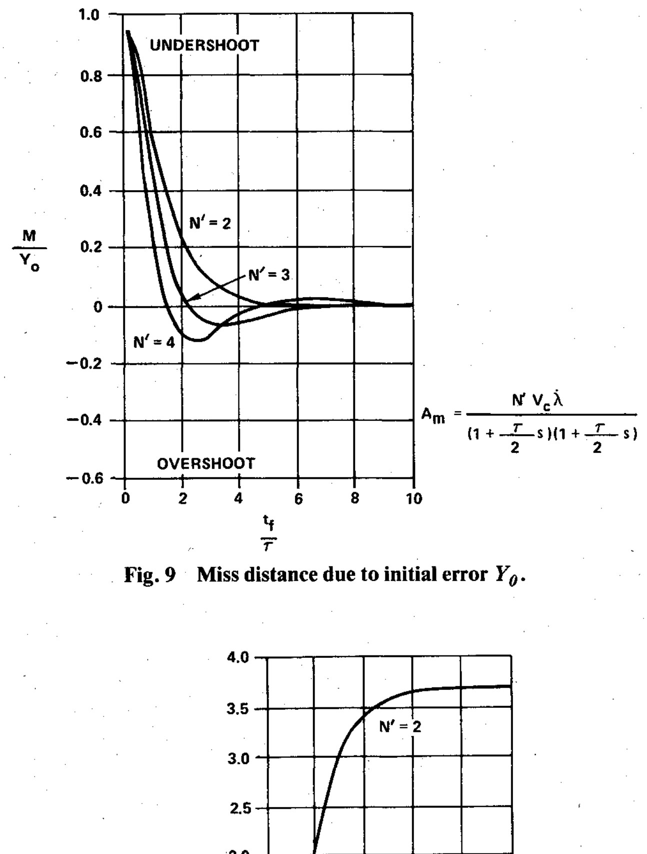

Fig. 9 Miss distance due to initial error Y_0.

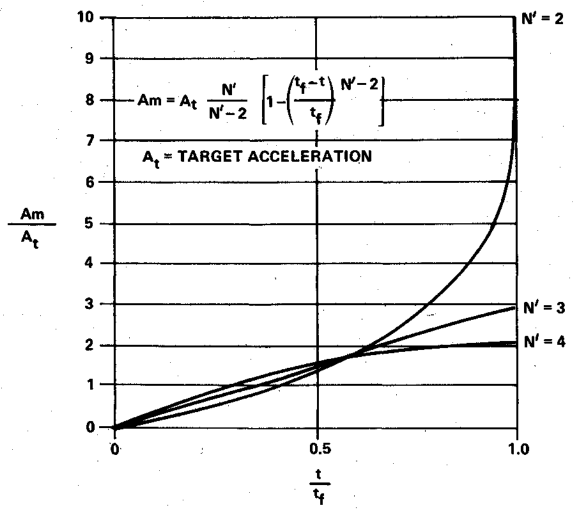

Fig. 8 Missile acceleration required to overcome step function target maneuver.

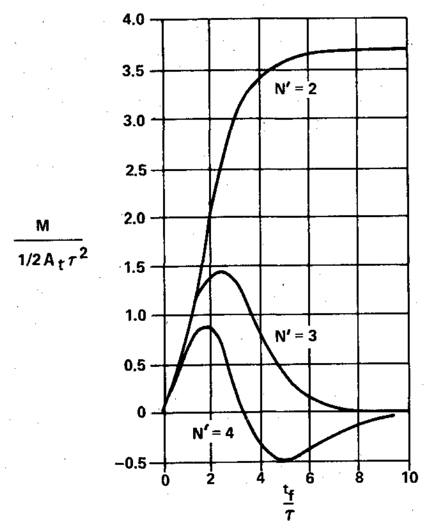

Fig. 10 Miss distance due to target maneuver.

Naval Weapons Center Sidewinder) was achieved by making the antenna itself a gyro. Since spinning the antenna of a CW radar receiver offers untold opportunities for modulating the target signal with bearing noise from the gyro, it was decided to build a hydraulically actuated gimbal system, space stabilized with subminiature rate gyros mounted on the back of the antenna. The development of the gyro, gimbals, and hydraulic actuators and valves was an ambitious but necessary undertaking.

The gyro was a straightforward but imaginative miniaturization of conventional technology. A spinning wheel, driven by an electric motor with 1000 Hz excitation, was contained within a partially floated gimbal structure that had a jewel bearing at one end and a spring restraint at the other end. The antenna gimbal was a "four-bar linkage," in which the antenna was kept close to the center of rotation, allowing as large an antenna diameter as possible—a key element in any radar design. Finally, a two-stage, four-way valve was developed that was characterized by large control forces from the second stage and low leakage in the pilot stage. The low leakage allowed use of a simple open hydraulic system in which the used oil is dumped overboard rather than having to be pumped back to high pressure.

To accomplish proportional navigation with a fixed value of N' requires a control system in which

$$\dot{\gamma}=N\dot{\lambda}$$

$$N=N'\ (V_C/V_M)$$

so

$$\dot{\gamma}=N'V_C\dot{\lambda}/V_M$$

or

$$A_M=V_M\dot{\gamma}=N'V_C\dot{\lambda}$$

where A_M is the missile lateral acceleration.

The principal outputs of the missile seeker are line-of-sight rate and Doppler frequency, which is proportional to closing velocity. The seeker thus provides the information required to establish the product $V_C\dot{\lambda}$ and a constant value of N' can be implemented by commanding missile acceleration proportional to that product. This can be done readily by using an acceleration feedback autopilot with an accelerometer as the principal feedback element. Because such an acceleration feedback autopilot was also the most direct way to provide a missile response sufficiently invariant with missile altitude and velocity, it was adopted. The total dynamic range was too great to allow satisfactory performance with a single set of autopilot gains; however, satisfactory performance was achieved with gain changes based on the altitude of the launch aircraft. Over a range of 2:1 in static pressure and at least that in velocity, each set of gains operated over a dynamic pressure range of about 10:1. The accelerometer was built by replacing the rotating wheel of the head stabilization gyro with an offset mass, giving an output signal proportional to the linear acceleration.

The miss distance produced by a target maneuver was seen (Fig. 10) to be proportional to (and indeed to have a peak value roughly equal to)

$$M=\tfrac{1}{2}A_T\tau^2$$

which is the distance a target can maneuver in one guidance time constant τ. This places an upper limit on the acceptable value for autopilot response of about 0.5 s for targets of that time. In order to achieve this response, it was necessary to stabilize the acceleration feedback loop with the body pitch rate, which was sensed by an additional pair of gyros mounted directly to the missile body.

Some comments on the airframe and its implications for autopilot design should be made. The original Sparrow configuration is shown in Fig. 11; it was a cruciform design, with identical planforms in the orthogonal pitch and yaw planes. For a variety of reasons, the forward surfaces (called the wings) were chosen to be the movable surfaces. This obviously made it easier to package the control servos, since there was no rocket blast tube taking up the center space as would be required with tail control. Perhaps more important to the decision was the necessity for fast response. When the wing is controlled, the initial lift is in the desired direction, permitting rapid achievement of the commanded maneuver. In a tail control missile, the tail is initially deflected in the opposite direction in order to develop the angle of attack that eventually provides the desired lift.

Two difficulties were found in the original Sparrow configuration. First, the static stability of the airframe was quite nonlinear because of the shadowing of the tail surfaces at combined pitch and yaw angles of attack, a situation not found in aircraft that maneuver only in pitch. Second, roll control variations with the vector angle of attack were very large. Roll control was provided by the differential motion of the forward wings responding to the roll motion sensed by an electronically integrated roll rate gyro identical with the pitch and yaw gyros. Unfortunately, the downwash from the wings creates forces on the tails in a direction opposite to the wing forces. Because the spans of the wings and tails were nearly equal, the net control rolling moment was often very small (and sometimes even negative), leading to the possibility of uncontrollable transient roll motions that could seriously degrade missile accuracy.

While this was the configuration actually deployed in Sparrow I, Douglas Aircraft Company developed an improved airframe configuration for Sparrow II, which was then adopted for the tactical version of Sparrow III. This configuration is also shown in Fig. 11. It is seen that the leading-edge sweep of the wings and tails is essentially reversed from the earlier configuration, giving a wing span much larger than the tail span. This gave much more positive roll control and improved the linearity in pitch somewhat.

A more elegant solution to pitch stability was found, however. By electronically integrating the pitch rate gyro output, an autopilot signal proportional to body attitude (which is equal to angle of attack at high frequencies) was developed. Its use in the autopilot made the accelerometer loop operate as though the airframe had an increased static stability M_α. With this "synthetic stability," even a statically unstable airframe could be stabilized in the closed loop. It should be recalled that the center of mass of the rocket propellant is normally well aft of the missile center of gravity (since the seeker uses the front of the missile and the rocket exhaust uses the rear). As a result, most missiles are statically unstable at launch. This is one of the reasons that missiles were initially not guided during the boost phase. Synthetic stability overcame this concern and Sparrow was allowed to begin guiding as soon as lock-on occurred. This dramatically reduced the minimum launch range, a factor whose significance was not recognized until later, as will be discussed below.

With this autopilot configuration, all of the requirements for a tactical design could be met. There remained the need to

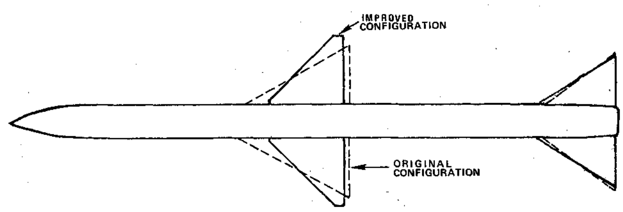

Fig. 11 Sparrow configuration improvements.

apply servo design theory to maximize performance (i.e., speed of response) while maintaining loop stability in the presence of various so-called "parasitic" feedback paths. These paths arise because of imperfections in the system described above. The imperfections include randome bore-sight error, antenna stabilization error, acceleration sensitivity of the head gyros, and bending of the missile body.

The radome problem arises because the nose of the missile should be pointed for aerodynamic reasons and no radome material is completely transparent to radar energy. As a result, as the antenna is gimballed inside the radome, a small angular displacement in the apparent location of the target is induced by rotation of the electromagnetic wave front as it passes through the radome material. This creates a feedback loop in which the missile responds to a target line-of-sight change by maneuvering; the resulting rotation of the missile body causes an apparent additional change in the line-of-sight angle, which closes the loop.

The effect of this feedback on the control system depends on the sign of the radome error. When the error is degenerative (i.e., tending to reduce the input line-of-sight rate), it slows the system response and reduces the stability margin of the autopilot. When the error is regenerative, it creates the possibility of a guidance instability at very low frequency. While either can degrade the miss distance, the regenerative sign is generally the more significant because low-frequency oscillations are characterized by large displacements of the missile.

The effect can be seen mathematically by considering that the first-order effect of the radome is to introduce an additional term in the closed-loop response of the control system. That is,

$$A_m = \frac{N' V_C \dot{\lambda}(1+R)/(1+NR)}{1+[(\tau+NR\alpha/\dot{\gamma})/(1+NR)]s} \tag{2}$$

where τ is the first-order time lag of the control system; N the actual (not effective) navigation ratio; $\alpha/\dot{\gamma}$ the "pitch turning time constant" of the airframe, a measure of the maneuvering capability of the missile airframe; and R the radome bore sight error slope in degrees of bore sight error per degree of gimbal angle change.

It is seen from Eq. (2) that either sign of R can degrade performance; a positive sign for R (degenerative error) reduces the effective navigation ratio and increases the response time of the system; a negative sign reduces the damping of the system when higher-order terms are considered. A typical miss distance sensitivity to R is shown in Fig. 12. It is seen that the miss increases catastrophically for negative values greater than some threshold that is a function of the flight condition (through $\alpha/\dot{\gamma}$) and of the control system design (through τ). Since $\alpha/\dot{\gamma}$ varies with altitude, the designer must increase τ in a corresponding manner to maintain acceptable performance.

The approach to designing low-distortion radomes evolved rapidly during the early years. The first radomes were of conventional sandwich design, with thin fiberglass walls covering a lightweight foam core. It was found that inserts made out of the foam material could be placed inside the radomes to "straighten out" the warped beam. As analysis became more sophisticated, however, it was found that solid fiberglass (later ceramic) walls approximately one-half a wavelength thick performed inherently better. Complex tapers in the wall thicknesses were designed into the domes to minimize the errors.

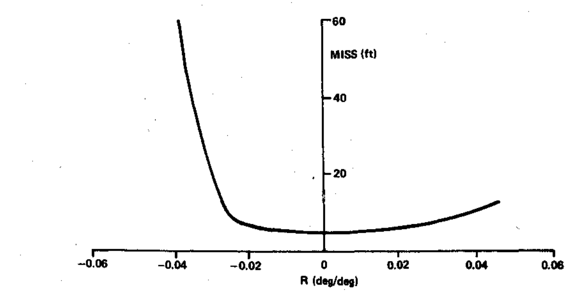

Fig. 12 Representative effect of radome bore sight error slope on miss distance.

In an analogous manner to the radome error, there is a feedback loop formed by imperfections in the antenna stabilization system, which result in a change in the antenna direction as the missile body attitude (or missile acceleration) changes. The stabilization loop was formed by rate gyros mounted on the back of the antenna, with their outputs electronically integrated to drive a hydraulically actuated gimbal system. Because of the finite frequency response of the servo, the antenna does not remain perfectly stationary for body motions at frequencies of primary concern; the motion of the antenna causes a bore sight error that is equivalent to that caused by positive radome slope. Considerable difficulty was experienced in closing the head stabilization loop to a high enough frequency to reduce this effect to acceptable levels, the limitation being the frequency response of the hydraulic valves.

In parallel with improving valve response, an electronic change was introduced that significantly eased the problem. It was noted (somewhat belatedly around 1956) that the close-out frequency of the stabilization loop is important only insofar as it affects the gain in the stabilization loop at the frequencies of interest, typically 1-3 Hz. It was realized that this could be increased by changing the electronic integrator to provide a -2 slope instead of a -1 slope, to the degree that loop stability considerations allowed.

The similarity to radome slope was further highlighted when it was noted that purposely limiting the dc gain of the stabilization loop (instead of allowing perfect integration of the rate gyro) produced the same effect as a positive radome slope, thus compensating for the negative radome slope at the very low frequencies that resulted in a large miss distance. This very simple compensation for a serious accuracy problem was therefore implemented.

The other stabilization loop problem arose because of the acceleration sensitivity of the rate gyros employed. Any unbalance of the gyro mechanism about the output axis will result in an erroneous rate signal when the gyro is subjected to linear acceleration normal to that axis. This rate signal will cause a tracking error that in turn commands an additional lateral acceleration to close the loop. This effect was unfortunately initially discovered during a flight test in which the missile flew a helical path to an excessive miss distance. The large miss was immediately explained by analysis and improved mass balance procedures were initiated; in addition, the gyro was mounted in a direction that minimized the miss distance impact.

The final item in the list of parasitic feedbacks is the effect of body bending. This is not really a separate parasitic loop, but simply a high-frequency autopilot instability in which body bending is detected by the autopilot as a motion of the missile. This problem was also discovered as a flight test instability, but fortunately in a Sparrow I flight before the Sparrow III autopilot design was completed. A test was immediately designed for Sparrow III in which the missile was plucked like a violin string and allowed to vibrate in a free-free mode. An early test fixture to accomplish this is shown in Fig. 13. This provided experimental verification of the body natural frequency and also allowed determination of the shape of the mode, which is very difficult to determine analytically with any accuracy. The autopilot rate gyros, which in Sparrow III were the most sensitive to this effect, were then mounted as close as possible to the nonrotational point for the first bending mode. Together with the electronic filtering employed, this avoided an autopilot instability.

Hawk

In 1952, while the Sparrow development was still struggling with fundamental problems, the U.S. Army initiated a program to develop the technology needed for a battlefield SAM system to protect friendly troops from low-altitude aircraft attack; this program was called "Project Hawk." The eventual award of an engineering development contract in 1954 to Raytheon for the Hawk system came about from an interesting series of events.

As a result of preliminary discussions with U.S. Army people, T.C. Wisenbaker, then Chief Engineer of Raytheon's flight test operation at Point Mugu, Calif., was invited to present a paper on low-altitude missile guidance at an Army Low Altitude Symposium in early 1953. Mr. Wisenbaker was confident he could convince the Army that CW radar was the unique solution to the problem of seeing moving targets hidden by ground clutter. However, there remained grave concern that the effect on guidance accuracy of radar reflections from the Earth (the so-called image or multipath problem) was fundamental and would cause the missile to home somewhere between the target and its image and therefore always strike the ground before reaching the target (Fig. 14).

About two weeks before the symposium, Wisenbaker turned to his design engineering staff, which consisted of Dr. Harold Rosen assisted by the author, with a request for help with this problem. We spent the remainder of that first day researching the nature of radar reflections from the ground. From Vol. I of the MIT Radiation Laboratory series[6] we learned that horizontally polarized microwave energy can reflect almost completely from a smooth Earth in the forward scatter direction at the grazing angles of interest, potentially creating the serious image problem that was feared. However, vertically polarized energy is significantly attenuated (cf., Figs. 15 and 16), leading to the possibility that the problem might be solvable if the grazing angle could be controlled. The reflection coefficient shown in Figs. 15 and 16 is simply the fraction of the incident energy reflecting in the forward scatter direction.

The next morning, we developed and solved in closed form the simplified equations representing the intercept of a low-flying target with no multipath disturbance. We were thus able to determine the time histories of the grazing angle (which leads to the determination of the reflection coefficient and thus the amplitude of the image-induced disturbance) and the interferometric lobe cutting rate (which is the frequency at which the disturbance would be felt by the missile seeker). The interferometric lobes are formed by the geometry-induced relative phase of the direct and reflected path signals; as the relative phase changes through 360 deg, the signals alternately reinforce and cancel and the apparent target elevation location goes through a complete cycle.

That afternoon, we quantified the analysis and found that launching at elevation angles in the range of 15-20 deg accomplishes the following:

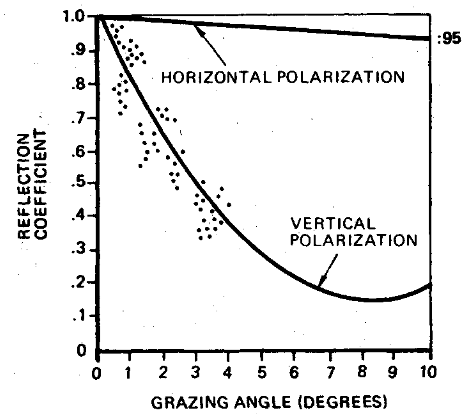

Fig. 15 Reflection coefficient over smooth sea (λ = 3 cm).

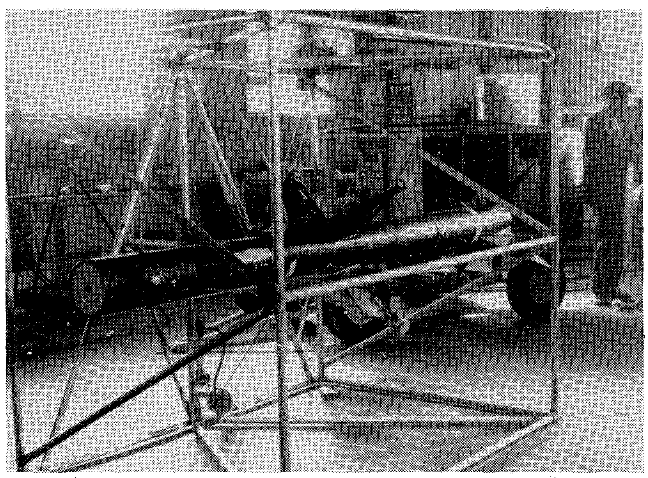

Fig. 13 Vibrational test fixture.

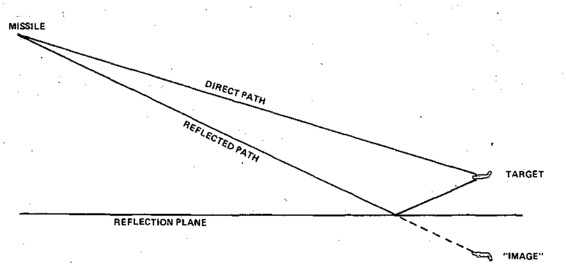

Fig. 14 Low-altitude geometry.

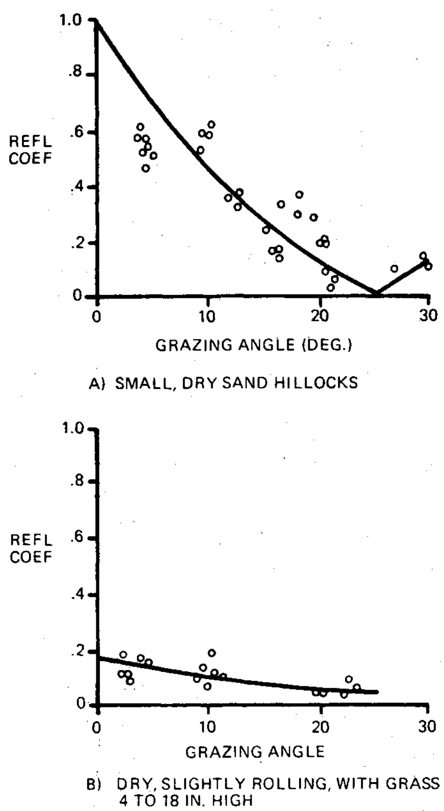

Fig. 16 Reflection coefficient over land, vertical polarization (λ = 3 cm).

1) It keeps the reflection coefficient below 0.5 in all circumstances, insuring first that the real target will be dominant, resulting in an average pointing direction of the antenna at the real target rather than the image, and second that the magnitude of the disturbance will be well bounded.

2) It keeps the frequency of the disturbance well above the autopilot bandwidth, so that the bounded disturbance established above does not significantly affect the homing trajectory and thus verifies the simplifying assumptions on which the analysis was based.

3) As we discovered later, it improves the aerodynamic range performance of the missile somewhat because the lower air density at the higher altitudes more than offsets the longer path to the target.

On the next day, we documented the analysis, which became the focus of Wisenbaker's presentation.[7] This was sufficiently convincing to result in Army acceptance of our proposal to perform 13 "critical tasks" considered essential to the successful development of the low-altitude missile system. The effort on this contract was led by Tom Phillips, now Chairman of the Raytheon Company.

The tasks took 10 months to complete. Toward the end of that period, the results were used as the basis of a preliminary design of a tactical weapon system capable of meeting the Army's requirements. The Army liked the preliminary design and in mid-1954 awarded Raytheon a contract for development of the Hawk system.

The image analysis was one of a whole series of activities that were necessary, but in themselves not sufficient, to ensure development of a successful missile system. Nevertheless, it was a good example of the contribution that system analysis can make in the solution of practical problems. The hardware was designed in accordance with the dictates of the theory and the analysis was eventually validated in all respects by subsequent flight tests. Indeed, the approach developed continues to be used in all low-altitude homing missiles to this day. An example of captive flight results is shown in Fig. 17 for horizontal and vertical polarization, with the white spot showing the seeker pointing direction.

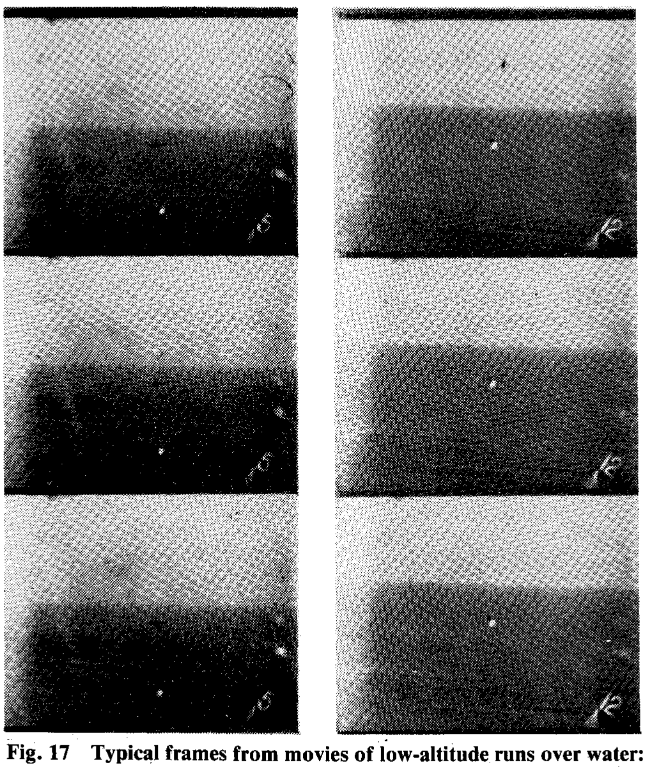

Fig. 17 Typical frames from movies of low-altitude runs over water: H-H polarization (left) and V-V polarization (right).

The development of Hawk proceeded with much greater ease than that of Sparrow because it lagged by 3 years and was able to benefit from the lessons learned in that period. A complete repackaging was necessary, of course, because the missile diameter was 14 in. instead of the 8 in. of Sparrow. However, the similarities were certainly much greater than the differences. Because Hawk was intended primarily for low-altitude defense (the Sparrow was considered an all-altitude missile), the radome fineness ratio (length-to-diameter ratio) was increased from 2.25 to 3.5, thus favoring aerodynamic drag in the tradeoff with the electrical characteristics. The hydraulically servoed gimbal system was scaled to the larger diameter, using the same hydraulic valves as the Sparrow wing servos.

The initial design of the electronics used circuits identical to those of Sparrow, although the two tended to drift apart in detail in response to detailed changes in requirements (for example, the initial conditions for Hawk included launch from exactly zero velocity). The hydraulic accumulator was increased in size, but remained functionally identical. The control surface hydraulic actuators at the rear of the missile were repackaged to allow the rocket engine exhaust to pass down the center, but retained the push-push antibacklash design and the same basic valve design.

The greatest change, and the only fundamental one, was in the airframe configuration. Because the Hawk started from zero velocity, it required a considerably larger rocket engine than the Sparrow to reach flight velocities. There was no feasible place to install servos in a wing control configuration with the performance required for Hawk, leading to selection of the tail control "dart" configuration shown in Fig. 18, which is similar to the Air Force/Hughes Falcon air-to-air missile.

That was not an accident. Hughes Aircraft Company published volumes of detailed information on the Falcon in the early 1950s, which were very useful to the entire industry. As a result, it was possible to prepare a preliminary design of a similar configuration with high confidence in the predicted aerodynamic characteristics. After Raytheon was selected to develop Hawk, one of the first tasks was to choose a subcontractor to conduct the detailed design of the airframe. In the less formal atmosphere of that era, the evaluation team consisted only of Tom Phillips and the author. One of the first companies visited was Hughes. In those discussions, Dr. Allen Puckett, then Chief of Aerodynamics and now Chairman of Hughes, gave a half-hour "chalk talk" on why they selected the dart configuration. It was a very effective talk, confirming our configuration choice and leading to an offer to Hughes to design and build the airframe.

After several days' consideration, the offer was declined on the basis that competition would be best served if we went our

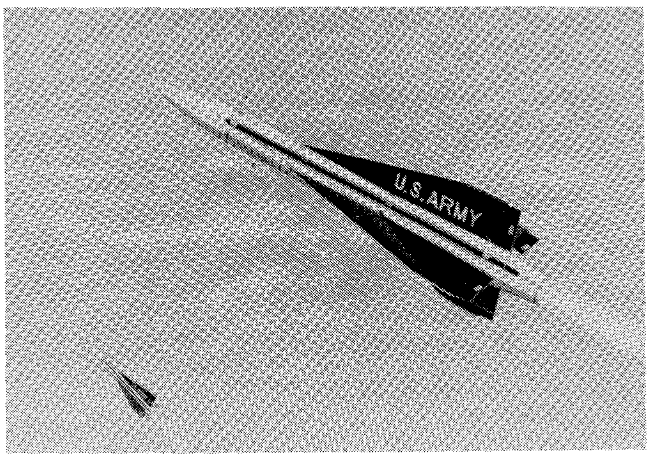

Fig. 18 Hawk configuration.

separate ways. Northrop Aircraft was subsequently selected to do the airframe. Although second choice, that relationship worked very well because of the superb engineering team assigned to the job, headed by Walt Fellers. Detailed studies confirmed the airframe choice, leading to only minor changes to optimize the design for this application. It is interesting to note that, as the design progressed, it was found that the presumed response time advantage of a wing control design is insignificant when the autopilot loops are closed; thus, all new missiles are tail controlled.

Studies showed that a boost-sustain rocket engine was preferable to an all-boost engine because it kept the radome temperatures caused by high speed at low altitude within the capabilities of the materials available at that time. The initial design consisted of two engines in tandem, with the forward sustainer exhausting through a blast tube centered in the rear booster. This design was replaced early in development with a much simpler single-chamber, dual-thrust design with the two radial burning propellants inside one another.

In the development program for the missile, most of the technical problems were encountered and solved on Sparrow before they occurred in Hawk (although a few items went in the other direction). This was not the case in the ground support equipment, which included important items not found in Sparrow. Examples of these are the illuminating radar, which was a tracking CW radar also used to illuminate the target for missile guidance, and a CW acquisition radar used to detect low-flying targets and designate them to the illuminator. While these had their own interesting history, they fall outside the scope of this paper and will not be addressed further here.

Other Related Activities

It may be worth a few words to summarize the parallel activities at other organizations, especially as they impacted the events described above.

The excellent work at Hughes Aircraft has already been alluded to with respect to homing guidance analysis and the Hawk airframe configuration. As was true for most organizations within the small missile community of the 1950s, they were very generous in making their data available to other groups. The Falcon missile and its related avionics system became the primary Air Force system on its interceptors for defense of the continental United States. The adoption by Hughes of pulse-Doppler techniques for radar and missile guidance in the late 1950s permitted them to extend their strong position in these areas up to the present time.

In the area of shipboard defense, the problem that led to the original Lark program, the Applied Physics Laboratory (APL) of Johns Hopkins University was the principal Navy advisor in the 1950s. The first generation ship-launched missiles, Terrier and Talos, were designed to be radar beam-rider guided to avoid the problems of homing guidance discussed earlier. Beam-rider guidance, in which the missile senses its position relative to the center of the shipboard radar beam that is tracking the target, is inherently inaccurate for low-flying targets because of reflections from the sea (the multipath problem discussed previously). Unlike the homing missile situation, the beam-rider multipath problem is fundamental and cannot be resolved satisfactorily.

Therefore, as homing missiles began to demonstrate the required technology, APL began an investigation of the applicability of homing guidance. This led in 1953 to contracts with Raytheon for system engineering support and for the procurement of a number of Sparrow seekers for evaluation by ground and flight tests in modified Terrier missiles. As a result of this program, semiactive CW guidance was selected for a new missile, Tartar, and a homing version of Terrier, which were developed and produced by General Dynamics Corporation. These missiles were combined in the "Standard" missile program in the 1960s, which remains the primary Navy air defense missile to this day.

Finally, there was considerable U.S. interaction in the 1950s with related British efforts. The British were also pursuing semiactive CW guidance and were doing high-quality work in a number of critical component and theoretical analysis areas. The transmitter for the original Raytheon Sparrow work was an ultraquiet 60 W magnetron.The Hawk program was started with a higher-power 200 W version of this tube. At about this time, the British were developing a 2 kW klystron amplifier with low-noise characteristics. Part way through the Hawk development program, this tube was adopted for the Hawk tracking illuminator and continues in use today.

The 1960s

Both Sparrow and Hawk were successfully developed, put into production (a task nearly comparable in difficulty to the development task and also beyond the scope of this paper), and deployed (Sparrow in 1959 and Hawk in 1960). It seemed to be an appropriate way to close out the decade that saw the beginning of the homing missile era. To some of us who had lived through it, it appeared as if all the interesting problems had been solved; this lack of foresight is not uncharacteristic of people who are too busy to step back to view their situation from a distance. As it happened, circumstances kept intruding on our reverie, raising new problems to be solved. I will review here only those bearing directly on the decisions described above.

Sparrow

The major event of the 1960s for Sparrow was the combat experience in Vietnam. As the history of operations accumulated, it quickly became clear that Sparrow's performance record was not living up to expectations. The author was a member of a joint Navy/contractor team sent to Vietnam in 1967 to gather first-hand knowledge of the air-to-air problems and recommend both short- and long-term fixes. We found that the reasons for the poor performance fell into two categories: 1) the way in which the missile was employed, to some extent peculiar to the unique situation in Vietnam 2) the missile performance requirements, generally applicable to any combat.

When Sparrow was designed, it was envisioned that the missile would be loaded on its carrying aircraft in time of war and fired at bomber aircraft approaching the fleet. Reliability was always a key consideration, but the critical period was thought to be the half-minute of missile flight. In Vietnam, these conditions did not apply. The North Vietnamese did not use air power to attack the fleet or other U.S. positions; instead, they used interceptors at isolated times and under conditions of their own choosing to try to disrupt U.S. air raids on the North. As a result, months would go by without air combat, during which the Sparrows would be carried daily in a fully operating state of F-4s conducting ground attack missions of several hours each.

Equally important, the avionics system (that is, the airborne radar and associated electronics) required to support the missile had to survive these months of punishing operation. The ground crews devoted full attention to keeping the aircraft in the air, but aircraft elements not vital to dropping bombs were not adequately maintained. When the MIGs suddenly appeared, the pilot would try to use the air-to-air system, usually unsuccessfully.

The primary lessons from this experience in missile employment lay in the areas of reliability and maintainability. As far as the missile was concerned, Raytheon applied the many lessons to the developmental version of "Advanced Sparrow," now called the AIM-7F, a far more rugged missile than its predecessors. For example, the mean time between failure in captive flight of the AIM-7F was more than an order of magnitude greater than that of the version used in Vietnam.

The second area of difficulty was in-flight performance. Sparrow was designed to shoot down bombing aircraft attacking a Naval task force. It was assumed that these attacks would generally be made beyond visual range. Since 95% of the aircraft over Vietnam were friendly, the United States very sensibly established a requirement that no target could be attacked until it was visually identified. Thus, old-fashioned "dogfights" reappeared and the launching ranges were often shorter than the missile could cope with. In addition, there was great surprise to discover that pilots will go to extreme lengths to survive. The fighter aircraft used by North Vietnam could maneuver at more than the 3 *g* for which Sparrow was designed. In addition, U.S. pilots, flying in North Vietnam against the Soviet-made SA-2 missile, found that they could outmaneuver the missile by pulling 5-6 *g* rather than the 3 *g* for which the SA-2 was apparently also designed. Even with the benefit of hindsight, it seems astonishing today that missiles would be designed to overcome only 3 *g* target maneuvers; but, in fact, that was what every missile in the world was designed to do in the 1950s.

The U.S. response to Sparrow's performance limitations was immediate and effective and it provides an interesting view of the psychology of that time. The minimum intercept range of the missile was established by the control system speed of response. Under the conditions representative of the Vietnam dogfights, the response time was longer than necessary, being determined by the desire to limit control surface movement in response to noise rather than by fundamental stability considerations. This filtering reduced the induced drag, thus preserving missile velocity for long-range intercepts, and also reduced the amount of hydraulic oil consumed in long-range flights. When it was realized that short flight times against very agile targets also formed an important missile application, a straightforward solution suggested itself: provide a "dogfight mode" with a minimum response time that could be selected when the time of flight was expected to be short. Thus, the long-range performance could be retained by using the normal mode when appropriate. In the dogfight mode, the minimum range was about half that of the normal mode.

A second change involved increasing the maximum maneuverability of the missile, a change that one might think would involve redesign of the whole missile; instead, it involved changing the value of four resistors. The reason it was so easy derives from the way in which design optimization involving both the missile and its launching aircraft was accomplished. The first production model of Sparrow III (AIM-7C) was designed for launch from a pylon-mounted rail under the wing of the Navy F3H-2 fighter. The rail guided the missile for the first 3 ft of travel, essentially getting it started in the right direction. This external carry contrasted with the internal stowage in a missile bay used for Air Force missiles, in which the launch rail and missile were lowered from the stowed position onto a trapeze before launch.

A relatively small number of AIM-7C missiles were produced before changing over to the AIM-7D, or "supersonic launch" version, compatible with launching from the F-4 fighter. In their installation studies for the F-4, McDonnell Aircraft found that a large positive increment in aircraft performance could be realized by using a "semisubmerged" carriage in which the underside of the fuselage is contoured to exactly match the upper half of the missile body. Carrying the missile in this way, with only its lower half projecting from the aircraft, significantly reduces the drag compared with the pylon/rail/missile combination in a wing installation; it also eliminates the need for designing the fuselage large enough to provide the volume needed for the internal carriage. In this concept, separation of the missile from the aircraft fuselage was accomplished using "ejection launch," in which a propellant charge is burned to provide a controlled force to eject the missile from its stowed position just before rocket ignition.

The principal problem with the semisubmerged installation was a concern that the missile might break up directly in front of the launch aircraft in the event of a control system failure. To avoid this, the missile structure was designed to withstand loads generated by maneuvers of about 60 *g*, a value greatly in excess of the 15 *g* established by electronic limiters on the guidance command signals for normal operation. This, of course, imposed a weight penalty of a few pounds, but it provided a rugged design better able to withstand the unanticipated rigors of field handling by military personnel, and it possessed the growth capability needed to cope with higher than anticipated target maneuvers. It was retained in the AIM-7E, an updated version that was the principal weapon employed in Vietnam.

The net result of this long story is that, by changing a few component values establishing the autopilot filters and the level of electronic limiters, Sparrow was transformed from a fleet defense missile against bombers to one that was also a dogfight missile against fighters. Within about four months, the changes were tested and proved and retrofitted missiles, designated AIM-7E-2, were on their way overseas.

Hawk

Following the trend of the 1950s, the Sparrow lessons were immediately applied to the Hawk design. An update of the Hawk missile, called Improved Hawk, was already well into development by 1967. Its special feature was a monopulse seeker with greatly improved resistance to electronic countermeasures. Following the Sparrow lead, the Improved Hawk acceleration limits were also increased and a fast-response autopilot mode was provided. This was more difficult to accomplish because the Hawk airframe, unlike Sparrow, was designed for only the in-flight maneuver level. With the required modifications to the autopilot and airframe, however, the Improved Hawk missile also became capable of coping with the highly maneuverable target.

In one sense, the maneuvering target was only the tip of the iceberg. The real lesson of Vietnam, which has been reconfirmed in the various Middle East conflicts, is that an attacking air force will undertake those actions that it believes are necessary to permit it to survive and carry out its mission. Maneuvering at uncomfortably high levels is certainly preferable to going down in flames. The whole range of offense options, known collectively as "defense suppression" when applied against ground-based defenses, includes direct attack with iron bombs (used very effectively by the Israeli Air Force); attack with standoff weapons such as antiradiation missiles (extensively used by the U.S. Air Force in Vietnam); use of standoff jamming aircraft to mask quiet penetrators with broad-band noise (used by the U.S. Navy in Vietnam); use of sophisticated jamming by the aircraft under attack; use of chaff to hide penetrating aircraft (extensively used in the Soviet invasion of Czechoslovakia); penetration at very low altitudes to underfly the defenses; and, looking ahead to the future, use of "stealth" techniques to reduce the electromagnetic signature of piloted and pilotless aircraft.

In the 1950s, many of these were recognized and allowed for in the design. For example, the clutter rejection features of Hawk apply equally to chaff, and an ability to angle track on noise radiated from the target being attacked provided a "home-on-jam" mode that defeats the most common electronic countermeasure. What was not appreciated in the 1950s was the tenacity with which the attackers will apply the techniques open to them when their survival is at stake.[8] Much of the effort since Vietnam has been devoted to closing these avenues. As would be expected, a generic approach to this problem—such as the monopulse seeker used in the Improved Hawk and later adapted for Sparrow (to show that information did flow in both directions)—should be used whenever possible. In those cases in which this is not possible, specific fixes must be applied. The modern use of general-purpose hardware under software control allows very sophisticated fixes to be done with minimum hardware change.

Conclusions

In this article, I have given my recollections of the principal events and technical accomplishments in the development of the first continuous-wave radar homing missiles. "Recollections" is, indeed, a suitably descriptive word. While I occasionally made reference to a document to clarify a point, the article has been written for the most part from my own memory. I would like to conclude by adding two non-technical points regarding the recollections.

The first has to do with my view of the flavor of the times. On joining Raytheon as a young engineer, I found an intensity of spirit that I had never experienced before. I felt myself carried along in what seemed almost a crusade, sharing a burning commitment to succeed against an immense challenge.

As a result of this shared feeling, each triumph led to a broad-based feeling of great elation. I still vividly recall the first flight test by Lark against a low-flying drone aircraft in 1951 over the ocean at Point Mugu. The intercept was at relatively short range and was the first to be in full view of the handful of observers permitted on the beach. When the missile homed unerringly to a spectacular direct hit on the drone, the human explosion matched the one in the air. The formal celebration ended in the wee hours of the morning, but the emotional wave lasted for months.

On the other hand, I can still feel the despair that resulted when a technician inadvertantly connected B+ (250 V) to the filament string, blowing out every tube in an early Sparrow being prepared for flight test. At that time, the best flight test engineers we had took about a month of 12-hour days to check out a missile. The resulting delay was felt at a personal level in every corner of the organization.

For years I attributed this intensity to the relative youth of most of the participants. However, I found that age has served mainly to mute their exuberance, but not their intensity, and I was forced to seek another source. My current view is that the intensity was (and is) a result of the intellectual challenge of the unknown. It is characteristic of the engineering profession that there are always new problems waiting in the wings to replace the old ones as they are solved. The constant element is the challenge itself, which demands that we do our utmost and rewards us only when we do.

My final point is a caution regarding the accuracy of the factual information. While I recall the events described, the perception of reality even as the events occur is never perfect and the intervening 30 years have created a selective filter with unknown coefficients through which my memory must pass. If this combination has distorted reality, I can only fall back on a defense borrowed from the Li'l Abner comic strip. When Pappy Yokum's recollection of past events drifted (as it usually did) into highly colored flights of embellishment, Mammy Yokum would bring him back to Earth with a recital of the hard facts. While outwardly acknowledging the facts, Pappy would still cling to his dreams, as people are wont to do. With a show of bravado, he would reply, "Well, mainly it were so."

References

[1] Haeusserman, W., "Developments in the Field of Automatic Guidance and Control of Rockets," *Journal of Guidance and Control*, Vol. 4, May-June 1981, pp. 225-239.

[2] Draper, C. S., "Origins of Inertial Navigation," *Journal of Guidance and Control*, Vol. 4, Sept.-Oct. 1981, pp. 449-463.

[3] Fossier, M. W., "Tactical Missile Guidance at Raytheon," *Electronic Progress*, Vol. XXII, Fall 1980, pp. 2-9.

[4] Muchmore, R. G. et al., "Summary Report—Guidance and Control I," Hughes Aircraft Co. Research and Development Laboratories, Tech. Memo. 235, July 1, 1950.

[5] Bennett, R. R. and Mathews, W. E., "Analytical Determination of Miss Distance for Linear Homing Navigation Systems," Hughes Aircraft Co. Tech. Memo. 260, March 1952.

[6] Kerr, D. E., "Propagation of Short Radio Waves," *MIT Radiation Laboratory Series*, McGraw Hill Book Co., New York, 1951, pp. 443-434.

[7] Wisenbaker, T. C., "Image Effects in Missile Seekers," Paper presented at Symposium on Low Altitude Target Problem, U. S. Army Research and Development Board, Redstone Arsenal, Ala., Jan. 6-8, 1953.

[8] Fossier, M. W., "The Role of SAMs in Tactical Warfare," Paper presented at Conference on the Impact of New Technologies on U.S. Defense Planning for Non-Nuclear Conflict," International Security Studies Program, The Fletcher School of Law and Diplomacy, Washington, D.C., Sept. 16-18, 1974.