UNCLASSIFIED

AD NUMBER
AD214252
CLASSIFICATION CHANGES
TO: unclassified
FROM: confidential
LIMITATION CHANGES
TO: Approved for public release, distribution unlimited
FROM: Controlling DoD Organization. Office of the Chief of Naval Operations, Washington, DC 20350
AUTHORITY
CFSTI/DDC-IR, 10 May 1966; CFSTI/DDC-IR, 10 May 1966

THIS PAGE IS UNCLASSIFIED

UNCLASSIFIED

AD 214 252

Reproduced by the

ARMED SERVICES TECHNICAL INFORMATION AGENCY
ARLINGTON HALL STATION
ARLINGTON 12, VIRGINIA

UNCLASSIFIED

AD No. 214 252

214 252

Division 6
Volume 2B
UNCLASSIFIED

FC

REPORT NO. 56
OF THE
OPERATIONS EVALUATION GROUP

UNCLASSIFIED

This volume contains material
originally issued as OEG studies
Nos. 255, 256, 257, 258, 263, 265, 268, 271, and 272.

Manuscript and illustrations for this volume were prepared for publication by the Summary Reports Group of the Columbia University Division of War Research under contract OEMsr-1131 with the Office of Scientific Research and Development. This report was printed and bound by the Columbia University Press and was also issued as Volume 2B of Division 6 in the series of Summary Technical Reports of the National Defense Research Committee.

Distribution of this volume has been made by the Chief of Naval Operations. Inquiries concerning the availability and distribution of this report and microfilmed and other reference material should be addressed to the Office of the Chief of Naval Operations, Navy Department, Washington 25, D. C.

Copy No.

96

OEG REPORT NO. 56

SEARCH AND SCREENING

BERNARD OSGOOD KOOPMAN

OPERATIONS EVALUATION GROUP
OFFICE OF THE CHIEF OF NAVAL OPERATIONS
NAVY DEPARTMENT

WASHINGTON, D. C., 1946

FOREWORD

This volume embodies the results of some of the statistical and analytical work done during the period 1942-1945 by members of the Anti-Submarine Warfare Operations Research Group [ASWORG] of the U. S. Navy, later the Operations Research Group [ORG] and, since January 1946, the Operations Evaluation Group [OEG]. The group was formed and financed by the Office of Scientific Research and Development at the request of the Navy, and was assigned to the Headquarters of the Commander in Chief, U. S. Fleet. The group has been of assistance in:

1. The evaluation of new equipment to meet military requirements.
2. The evaluation of specific phases of operations from studies of action reports.
3. The evaluation and analysis of tactical problems to measure the operational behavior of new material.
4. The development of new tactical doctrine to meet specific requirements.
5. The technical aspect of strategic planning.
6. The liaison for the Fleets with the development and research laboratories, naval and extra-naval.

The material presented in this volume has been compiled from reports and memoranda issued by the group and from the first-hand knowledge and experience which the authors gained during World War II with the techniques and problems discussed.

Specific examples are developed of the applications of the more general *Methods of Operations Research* to the problems of submarine warfare. Also, a mathematical basis is provided for the *Summary of ASW Operations in World War II*, as well as for a wide category of similar investigations. Although the tactical doctrines presented apply to instruments, weapons, and conditions prevailing during World War II, it is believed that the methods and systematic processes of analysis which led to these doctrines have wide application — not only to submarine warfare but to many other military and civilian problems.

It is increasingly evident that no branch of the Service can afford anything less than maximum efficiency in the use of the men and matériel available to it. The realization of this ideal demands that the most advanced scientific knowledge available in the country be focused upon such matters not only in time of war, but especially in time of peace. It is the earnest hope of the OEG that the material contained in this and the companion volumes, by helping to provide a basic understanding of the processes of this important branch of warfare, will materially contribute to this goal.

Philip M. Morse
Director, Operations Evaluation Group

PREFACE

As the Operations Research Group was at work investigating one question after another in the course of its service to the Commander-in-Chief of the United States Navy, in World War II, it became progressively more apparent that large classes of problems were united by common bonds and could be handled by common methods, that there was indeed unity in diversity. And as in other fields of scientific endeavor, where the clarifying influence of general ideas and methods can form a body of isolated facts into a powerful theory—once they exist in sufficient number—so in the work of the Group, methods borrowed from the mathematician and mathematical physicist showed their power and usefulness in those classes of problems in which the body of practical information had sufficiently accumulated. In this regard, one field was pre-eminently ripe for mathematical treatment: the field involving problems of *search*.

In every question of search there are in principle two parts. One involves the targets, and studies their physical characteristics, position, and motion; since from the very nature of the problem the latter are largely unknown to the searcher, a branch of the science of probability is applied, sometimes so simple as to be trivial, at other times involving developments comparable to statistical mechanics. The other part involves the searcher, his capabilities, position, and motion; inasmuch as detection is an event fraught with manifold uncertainties, this part of the question will also appeal to probability, specifically studying the probability laws of detection. But the study does not stop here: having gained fundamental knowledge as to these two parts of the question and their interrelation, it is necessary to make application to the tactical matters in which search is an essential component, such as hunts, barriers, and those defensive types of search known as *screens*.

The book treats these questions from the point of view and in the order indicated above. It is intended to be scientific and critical in spirit and mathematical in method, and while the data upon which its theory rests are practical and experimental and the ultimate application of its conclusions is to naval warfare, the book itself is not a manual of practical information for naval officers. Rather it is intended to serve as a theoretical framework and foundation for more immediately practical studies and recommendations. In particular, it stands in this relation to Volume 3 of the present series (*A Summary of Antisubmarine Warfare Operations in World War II*). On the other hand, its relation to Volume 2A (*Methods of Operations Research*) is in furnishing systematically developed examples, on the analytical side, of the possibilities of operations research foreshadowed in that volume. It is intended for a reader having an interest of a scientific order in the matters treated. While nothing beyond undergraduate physics and mathematics (calculus) is required, a willingness to follow theoretical reasoning of a sometimes rather involved nature is assumed.

The work has been to such a degree the result of a majority of the Operations Research Group that to render adequate acknowledgment would almost be tantamount to giving the roster of the Group: requirements of brevity confine us to the names of those who have been directly involved in writing parts of the book. We wish to express our thanks to Dr. E. S. Lamar for the chapter on visual detection; to Mr. T. E. Phipps for the chapter on radar detection; to Mr. A. M. Thorndike for the chapter on sonar detection and for a part of the chapter on sonar screens; to Dr. J. Steinhardt for important help in the chapter on radar detection and for material on barrier patrols and defense of a landing operation in Chapter 7; to Dr. J. M. Dobbie for material on square searches in the latter chapter; to Mr. Milton Lewis for material on sonar screens; to Mr. J. A. Neuendorffer for material on aerial escort. Finally, it is our great pleasure to vouchsafe our indebtedness to Dr. G. E. Kimball, the pioneer in the theory of search, without whose help and inspiration this enterprise might never have been undertaken.

B. O. Koopman

CONTENTS

Chapter 1

POSITION, MOTION, AND RANDOM ENCOUNTERS

1.1 INTRODUCTION

Among the important functions of any naval operation is the detection (and location) of the enemy. Detection has acquired the stature of a science, and it is to the foundations of this science of detection that the present work is devoted.

A first aspect of any problem of detection concerns the properties of the instrument of detection: the properties of the eye, the characteristics of the radar set, or the nature and capabilities of the sonar equipment, and similarly for any other mechanism of detection which it is proposed to employ. This aspect of the question, which forms the subject of Chapters 4, 5, and 6, is called the *contact problem*. It is essentially a study of engineering or, in the case of the eye, of physiology, but its conclusions have to be given in terms of probability—the probability of detection.

A second aspect of problems of detection concerns the path and motion of the searching unit (termed the *observer*) in its relation to the presumed position and motion of the object of search (the *target*, as it shall be called throughout). This presupposes that the contact problem has been solved to a satisfactory approximation and employs appropriate methods of geometry, relative motion, and probability. This aspect of the question is called the *track-problem* and forms the subject of Chapters 7, 8, and 9, where such matters as searches, barriers, hunts, and screens are studied.

A third aspect—and this will pervade all our later chapters, but particularly Chapter 3—is that of *force requirements* and their economy. As in other naval operations, effectiveness can be increased to perfection if no limit is set to the forces at our disposal, but the realistic problem is to achieve the greatest effect with limited forces, or, equivalently, to achieve a required effect with the greatest economy of forces.

The scope of the book necessitates the omission of such matters as the use and deployment of striking forces in conjunction with detection.

The present chapter provides the general definitions and geometrical and statistical methods which are required in all later parts of the work. Chapter 2 deals with the generalities involved in detection.

Viewed largely and in its ends, the science of detection is a branch of tactics. But like all other branches of tactics it achieves its ends by leading through engineering, physics, physiology, mathematics, and statistics. And it is an ever progressing science: While the main emphasis of this book is to study the existing knowledge and its application to tactical problems, sight is not lost of the inverse process, that of improving the theory by knowledge gained in its application, *the study of operational data.*

1.2 MOTION AT FIXED SPEED AND COURSE

In a most important case both observer and target are moving at constant speeds in straight lines:

v = speed of observer in knots (ocean or true speed),
u = speed of target in knots (ocean or true speed),
w = speed of target relative to observer in knots.

The relationship of w to u and v is best shown by drawing the *vector* velocities $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, whereupon it is seen that $\mathbf{w}$ is simply the vector difference $\mathbf{w} = \mathbf{u} - \mathbf{v}$ (Figure 1). For, $\mathbf{w}$ being the target's velocity with

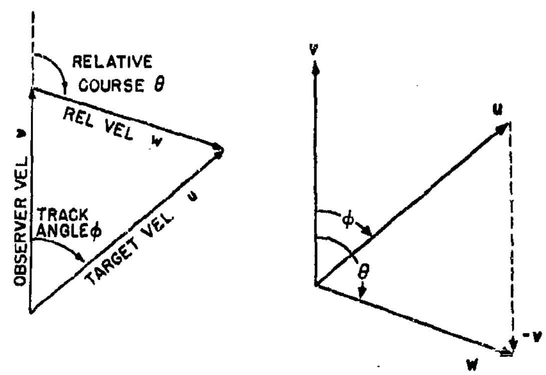


Figure 1. True and relative velocities and angles.

respect to a reference system, itself moving over the ocean at velocity $\mathbf{v}$, the vector sum $\mathbf{w} + \mathbf{v}$ must equal the target's ocean velocity $\mathbf{u}$; hence the above equation. Figure 1 also shows two important angles, the target's *track angle* ϕ and *relative course* θ, with respect to the observer, where

ϕ = angle between **v** and **u** measured from the former to the latter in the clockwise sense.

θ = angle between **v** and **w** measured from the former to the latter in the clockwise sense.

Throughout the present chapter, and later unless the contrary is explicitly stated, these angles are measured in radians and $0 \leqq \phi < 2\pi$, $0 \leqq \theta < 2\pi$.

A convenient method for showing the dependence of the relative quantities **w** and θ upon the angle ϕ (the speeds u and v remaining fixed) is by drawing the circular diagrams A, B, and C of Figure 2, corresponding to the cases $v < u$, $v = u$, and $v > u$, respectively. In each case the radius of the circle around which the extremities of **u** and **w** move is u, and the distance of its center from the origin O of **w** (the extremity of **v**) is v. As ϕ goes from 0 to 2π, **v** stays fixed, **u** rotates with its length remaining constant, and **w** changes both in length and direction. It is to be noted that while in case A ($v < u$) all directions of **w** are possible ($0 \leqq \theta < 2\pi$), in the other cases ($v \geqq u$) this is untrue, and we have:

When $v = u$, $\frac{\pi}{2} \leqq \theta \leqq \frac{3\pi}{2}$;

When $v > u$, $\pi - \sin^{-1}\frac{u}{v} \leqq \theta \leqq \pi + \sin^{-1}\frac{u}{v}$.

This corresponds with the fact that when the searcher is faster than the target, relative approach of the latter to the former is restricted (see Section 1.3). When $v > u$, two values of w correspond to general values of θ for which approach is possible, one for the target approaching the observer, the other for the overtaking of a target headed away from the observer. When $\theta = \pi \pm \sin^{-1} u/v$, there is just one value of w; for other θ's, no value.

The relative speed w can be found from the law of cosines or else by projecting **v** and **u** on **w** and using the law of sines; similarly for ϕ. This expresses relative quantities in terms of true:

$$w = \sqrt{u^2 + v^2 - 2uv\cos\phi},$$
$$= -v\cos\theta \pm \sqrt{u^2 - v^2\sin^2\theta}, \qquad (1)$$
$$\sin\theta = \frac{u}{w}\sin\phi.$$

In addition to the speeds and angles just considered, it is necessary to have further quantities to

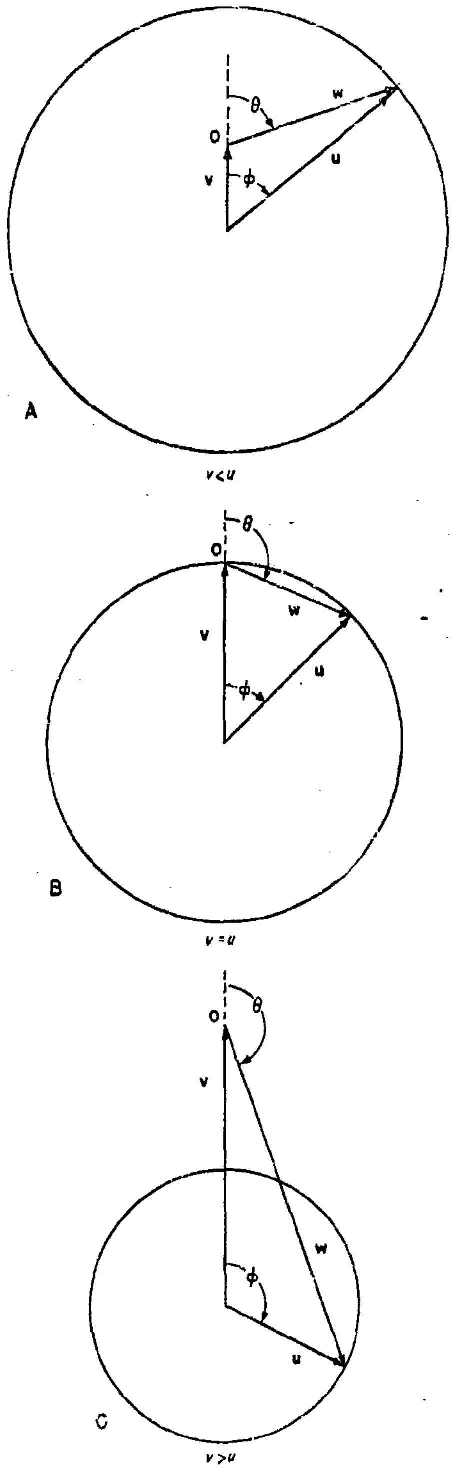


FIGURE 2. Circle of relative velocities.

specify a particular contact between observer and target; one must be able to state the position of the target relative to the observer at any given instant of time (epoch) t. One method of accomplishing this is to give the *target range* r and relative bearing β; these are shown in Figure 3 (at the arbitrary epoch t), together with the *target angle* α, which depends on quantities previously introduced. The definitions are as follows:

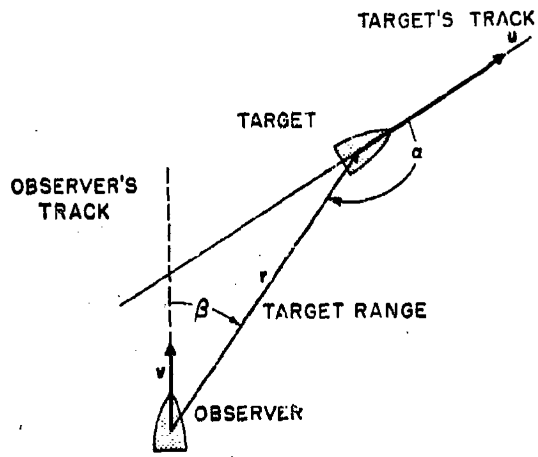


FIGURE 3. Geographic tracks, ranges, and bearings.

$\mathbf{r}$ = vector from observer to target,
r = length of $\mathbf{r}$ in miles,[a]
β = angle from $\mathbf{v}$ to $\mathbf{r}$ measured clockwise,
α = angle from $\mathbf{u}$ to $-\mathbf{r}$ measured clockwise.

As usual, angles are in radians and lie between 0 and 2π, except when in later chapters the contrary is explicitly stated. It is evident that when the rôles of target and searcher are interchanged, those of β and α are likewise, vector $\mathbf{r}$ being replaced by the reversed vector $-\mathbf{r}$.

The situation relative to the observer is given in Figure 4, which shows the target's track, etc., in a plane in which the observer is fixed and which moves over the ocean with the velocity $\mathbf{v}$. It is seen that (r,β) are the polar coordinates of the target referred to observer's position and heading. The target's track is altogether different from his geographic track of Figure 3; it is described with the velocity $\mathbf{w}$, but the target's heading is in the direction of $\mathbf{u}$ and hence not along its relative track. It is seen that with the particular angles of Figure 4, $\alpha = \pi + \beta - \phi$. It is sometimes convenient to use rectangular coordinates (ξ,η), the η axis being along the observer's heading. ξ and η are in miles; they are related to (r,β) by the equations

$$r^2 = \xi^2 + \eta^2, \quad \xi = r \sin \beta, \quad \eta = r \cos \beta. \tag{2}$$

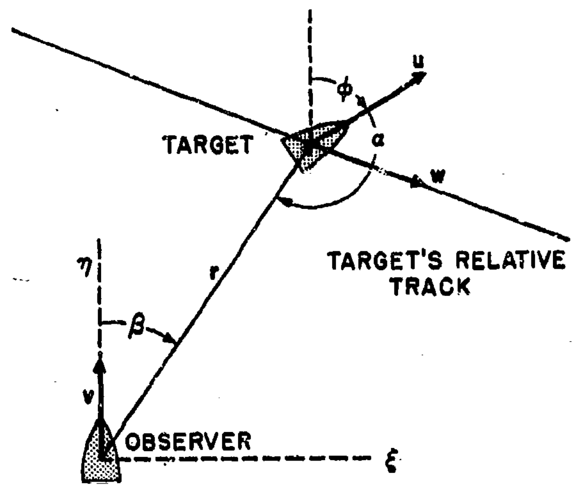


FIGURE 4. Target's track relative to observer.

In the course of time (as t increases) $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, u, v, w, ϕ, θ stay constant, while $\mathbf{r}$, r, β (and α) change. If at the epoch $t = t_0$, $\mathbf{r}$, r, β have the values $\mathbf{r}_0$, r_0, β_0, their values at a general epoch t are found by noting that relative to the observer the target undergoes the vector displacement $(t - t_0)\mathbf{w}$, and thus

$$\mathbf{r} = \mathbf{r}_0 + (t - t_0)\mathbf{w}, \tag{3}$$

from which the equations expressing (r,β) in terms of $(r_0,\beta_0,t - t_0)$ are found by trigonometry, and similarly for α. Equation (3) or its equivalent in terms of (r,β) are the equations of the target relative to the observer. In the very special case when target and observer have the same speed and direction, i. e., when $\mathbf{u} = \mathbf{v}$ so that $\mathbf{w} = 0$, equation (3) reduces to $\mathbf{r} = \mathbf{r}_0$, corresponding with the fact that the target remains fixed relative to the observer. In all other cases, there is a least distance between the target and the observer. This distance is called the lateral range of the target.

Let x = lateral range of the target in miles, and
y = distance in miles traveled by the target relative to the observer since its closest approach (negative prior to closest approach).

If t_0 is now used to denote the epoch of closest approach, evidently $y = (t - t_0)w$. Figure 5 shows the

[a] Throughout, a "mile" shall mean a "nautical mile." In numerical examples, a nautical mile is taken as 2,000 yards.

relation between x,y and the earlier quantities. Clearly $r^2 = x^2 + y^2$, and with the particular angles of Figure 5, $\beta = \theta - \cot^{-1}(y/x)$.

Thus it appears that, in addition to u, v, ϕ (and their dependent w, θ), in order to specify a *particular* contact we can use either (r,β) (at a standard epoch) or

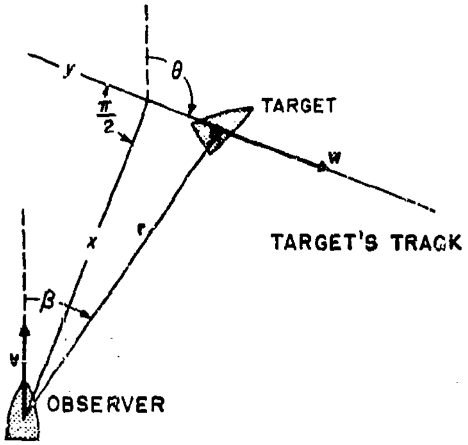


FIGURE 5. Lateral range of the target.

(x,t_0) (or, indeed, any convenient independent functions of either pair); (r,β) can be found where (x,t_0) are given, and vice versa, and either pair can be used as independent variables. On the other hand, only one quantity is needed to specify a *type* of contact, for example, x: Given u, v, ϕ and the range of closest approach x, the configuration (searcher and target and their tracks) is determined, but not the time at which the contact occurs.

1.3 THE REGION OF APPROACH

In Figure 6A, A is a plane region fixed with respect to the observer; A, then, is moving straight ahead over the ocean surface with the velocity **v**. It may or may not be possible for a target capable of moving with the speed u and starting outside A to enter A. It is understood that the target is restricted to the *speed* u but can choose any *direction*, and has all the time it needs to try to enter A. Evidently if $u > v$, the target can always enter A; but if $u \leq v$ this is no longer necessarily true, as for example when the target starts behind A. In order to be able to enter A the target must have its starting point in a certain region B called the *region of approach*. Of course B is also attached to the observer and moves over the ocean with the velocity **v**. Figure 6A shows the construction of B; when **v** points up the page, a line inclined at the angle $\sin^{-1} u/v$ with the vector **v** to the right is drawn to the right of A and is moved toward A until it touches A; the part of the line above the lowest point of contact forms the right-hand boundary of B. Similarly for the left-hand boundary, the inclination being to the left and the contact occurring on the left of A. The rear boundary is the forward boundary of A between the two rear points of contact.

The justification of this construction is based on Figure 6B: In the time t a point starting at P and moving with the observer's velocity will arrive at Q, $PQ = vt$; the circle centered at Q and of radius ut is the locus of positions from which the target must start if it is to close this point after the time t. By

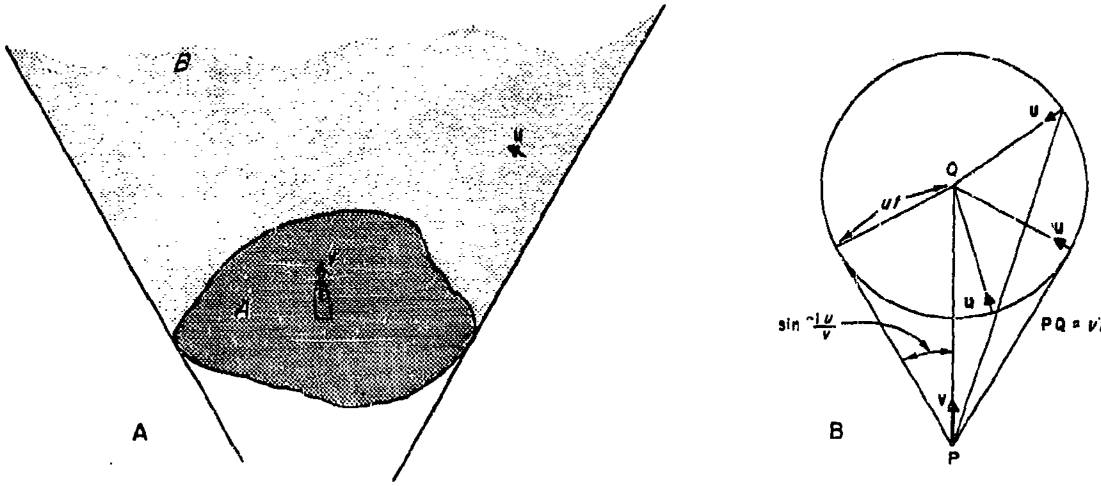


FIGURE 6. Construction of the region of approach.

considering all positive values of t and observing how the circle varies in both center and radius, it is seen that it will sweep out the whole angular region between the two (fixed) tangents drawn from O. Thus the whole upper angular space between the two tangents to the circle is the locus of all starting positions of the target if it is to close the moving point at *any* time after the latter leaves P. B is constructed by letting P take on all positions in A, whereupon the upper angular space attached to P sweeps out B (in addition to A itself).

When $u = v$, the tangents coalesce into a horizontal line tangent to A in the rear, and B is the upper half-plane above it (exclusive of A).

An important naval application of the region of approach is to submarine warfare, where v is the speed of the convoy and A the region within which a torpedo must be fired to be in range of a ship of the convoy. What we have been terming the "target" may be thought of as a submerged submarine having the underwater speed $u < v$. Then evidently the submarine must be in B in order that it may be able to approach, submerged, to a torpedo firing position.

The angle $\sin^{-1} u/v$ is called the *limiting approach angle*—the "limiting submerged approach angle" in the example of the submarine.

Quite a different figure for B is obtained if the target is assumed to have a limited time (T hours) to make its approach to A. Then even when $u > v$ approach is not always possible. The construction of A in such a case is given in Figure 7A and B, when $u > v$; Figure 7B shows a circle of radius uT centered at Q and vT units ahead of the starting point P; the circular region is the locus of starting positions from which the target can close the point. If P is given all positions on A, the points of the attached circle cover B (thus B is bounded by the envelope of the circles). A similar construction is made when $u < v$, where the starting point region is that bounded by the two tangents as well as the larger intercepted arc of a circle disposed as in Figure 6B.

An example of this second type of region A is the case in which account is taken of the limited endurance time of submerged run of the submarine in the previous example.

1.4 RANDOM DISTRIBUTIONS OF TARGETS

In the two preceding sections the motion of observer and target were given or precisely specified and the conclusions were exact. In Section 1.2 one observer and one target of stated speeds and tracks were assumed; in Section 1.3 the same was true for the observer and for the target's speed, but a precisely defined class of targets (those which enter A) was considered in defining B (the locus of their starting points) and the conclusion was the precise one: "*The target can enter A if and only if it starts in B.*" Fundamentally different is the state of affairs in the present section, in which the notion of *random* is introduced and conclusions are stated in terms of *probability*. Instead of saying, "Under such and such conditions the target will necessarily do so and so," we shall be saying, "Under such and such conditions the probability that the target will do so and so has this value," or, equivalently, "This percentage of targets will on the average do so and so." The importance of arriving at probabilities and statistical results in naval matters should be self-evident.

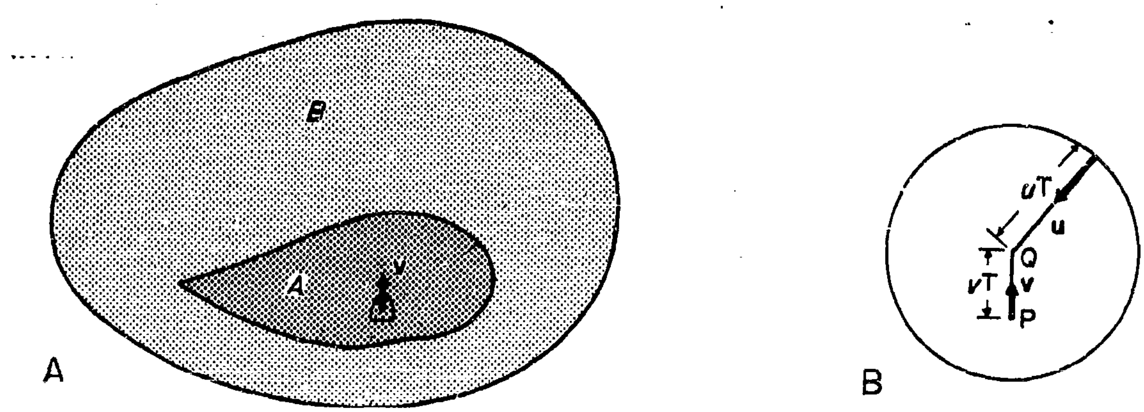


FIGURE 7. The region of approach with limited time.

To specify a target (always assumed in this chapter to be moving at constant speed and course) it is necessary to give its speed and heading (the vector **u**) and also its position at a particular epoch (i.e., when $t = t_0$). This requires in principle four independent quantities, such as u, ϕ, r, β. By a *random distribution* of targets is meant either of the two following situations:

I. There are present a very large number of different targets, and what is known is not their velocities and positions, but the *proportion* or *percentage* which have the various possible velocities and positions.

II. There is present only one target; its velocity and position are not precisely known, but the *probabilities* that it have the various possible velocities and positions are known.

In these statements, the proportion or probability of targets "having such and such a velocity and position"[b] must be interpreted to mean "having a velocity and position within a stated closeness of such and such a velocity and position." Thus if the above choice of quantities is made, it is the proportion or probability of targets having a speed between u and $u + du$, track angle between ϕ and $\phi + d\phi$, range between r and $r + dr$, and bearing between β and $\beta + d\beta$ which is in question; in many cases (but not all! see below) it can be represented to terms of first order in the differentials by $p(u,\phi,r,\beta)du d\phi dr d\beta$, and the function $p(u,\phi,r,\beta)$ is the *mean relative density* (I) or *probability density* (II) of the distribution. Then the proportion or probability for a large class of velocities and positions is obtained by integrating $p(u,\phi,r,\beta)$ over all values of the class considered—a quadruple integration in the "space" of the "coordinates" (u,ϕ,r,β).

While the situations in I and II above appear to be quite different, they are in reality equivalent, or rather either one leads to the other. Thus from the very large number of targets in I we can think of an individual target chosen at random, all targets having the same chance of being chosen; this target will then be the single target to which the situation II applies. Reciprocally, if a very large number of targets is constituted from individual targets to each of which the state of affairs of II applies, the resulting swarm will be as described in I. The mean relative density (I) and the probability density (II) are equal. All this is a consequence of the law of large numbers in the theory of probability. Situation I is generally used to give a pictorial representation of II which might otherwise seem too abstract; but I has the disadvantage of being rather unrealistic—if so many targets were actually all present they would be apt to interfere with one another physically.

What we have termed a *random distribution* should more properly be called a *known random distribution*, to distinguish it from the case where the values of the probabilities are partly or wholly unknown: a frequent problem of importance in operational analysis is to find them by theoretical calculations or statistical methods.

One of the simplest and most important cases of random distribution of targets is that of the *uniform distribution* of targets of given speed u. It is one which complies with the three following requirements.

a. The probability that the track angle ϕ be between ϕ_1 and ϕ_2 is proportional to $\phi_2 - \phi_1$ [and hence is equal to $(\phi_2 - \phi_1)/2\pi$ when $\phi_1 < \phi_2$].

b. The probability that at any chosen epoch the target be in the area A (fixed in the ocean or else fixed relative to the observer; the two situations are here equivalent) is proportional to A (and hence, if the target is known to be in a larger area B containing A, the probability that it be in A is A/B).

c. The event of ϕ being between ϕ_1 and ϕ_2 on the one hand and the event of the target being in A on the other are *independent events*. If one of them is known to have occurred, the probability of the occurrence of the other is the same as before.

In the case of such a distribution, the probability density will (when the target is given to be in B) have the value $p(\phi,r,\beta) = r/2\pi B$; this is because $r dr d\beta$ is the element of area corresponding to a position of range and bearing between r and $r + dr$ and β and $\beta + d\beta$ respectively. It is to be noted that the probability is $p(\phi,r,\beta)d\phi dr d\beta$ and not $p(u,\phi,r,\beta)du d\phi dr d\beta$, as in the earlier example, i.e., neither u nor du occur; this is because the value of the target's speed is supposed to be known.

The case considered is of importance in naval operations, since it corresponds to the situation in which the target is an enemy unit presumed to be running at the known speed of about u knots, but is in such a large area of ocean with so many possible intentions that nothing concerning its position or heading can be regarded as known. The chances that the observer will make various kinds of contacts with such a unit are studied in the following section.

[b] Such probabilities might all be zero. With many distributions the probability that the target be exactly at a pre-stated position is always zero. It is the probability that it lie in a pre-stated *area* which is of interest.

It is noted that in the foregoing example the language of II is used. This is permissible in view of the equivalence of II with I, and there is manifestly no difficulty in rewording things to correspond to I. Whichever of the two terminologies will be employed in the succeeding pages will be purely a matter of convenience: this involves absolutely no inconsistency.

1.5 RANDOM ENCOUNTERS WITH UNIFORMLY DISTRIBUTED TARGETS

When an observer is progressing on its course at the constant velocity **v** among a uniform random distribution of targets of speed u, it is frequently important to know the proportions of targets which pass within the stated range of R miles of the observer. In some cases R may be the range within which the observer can sight the target (horizon distance); in others, the range within which the target can detect the observer's presence; again, R may be effective gunfire range of observer against target, or vice versa. If a circle of radius R is pictured centered on the observer and moving along with it at velocity **v** over the ocean, the question becomes that of the proportion of targets entering the circle, or entering it at various specified bearings, or the chance that a target of given starting point shall enter the circle—a question of *probability*, in contrast with that of Section 1.3 which was one of *possibility*. The three problems will be solved in turn.

Problem 1. Let there be on the average N targets per square mile (N will usually be far less than unity; it is an "expected value" in the sense of probability). On account of the uniform distribution of track angles and their independence of position [(a) and (c) of Section 1.4], the average number with track angle between ϕ and $\phi + d\phi$ will be $Nd\phi/2\pi$. Fixing our attention exclusively on targets of a particular track angle ϕ, it is easy to find how many enter the circle per unit time. The relative speed and course are found as in Section 1.2 and the circle of radius R is drawn about the observer as shown in Figure 8. Since the target is moving at velocity **w** with respect to the observer, if it is to enter the circle in a unit of time (one hour) it must be in the large shaded region of Figure 8 between the circle and the circle moved through the displacement $-$**w**, and between their tangents parallel to **w**. The area being $2Rw$, it is seen that the number of targets of track angle between ϕ and $\phi + d\phi$ which enter the circle per unit time is (to quantities of first order in the differential) $2RwNd\phi/2\pi$. Hence the total number N_0 is given by integration:

$$\begin{aligned} N_0 &= \frac{2RN}{2\pi}\int_0^{2\pi} w\,d\phi, \\ &= \frac{RN}{\pi}\int_0^{2\pi}\sqrt{u^2 + v^2 - 2uv\cos\phi}\,d\phi, \\ &= \frac{4RN}{\pi}(u+v)\int_0^{\pi/2}\sqrt{1 - \sin^2\sigma\sin^2\psi}\,d\psi, \\ &= \frac{4RN}{\pi}(u+v)E(\sigma),\ \sin\sigma = \frac{2\sqrt{uv}}{u+v} \end{aligned} \tag{4}$$

Here the second equation results from equation (1), the third by introducing σ and the new variable of integration $\psi = (\pi - \phi)/2$, and $E(\sigma)$ is the complete elliptic integral of the second kind.

Note that equation (4) is left unchanged if u and v are interchanged. This corresponds with the fact that

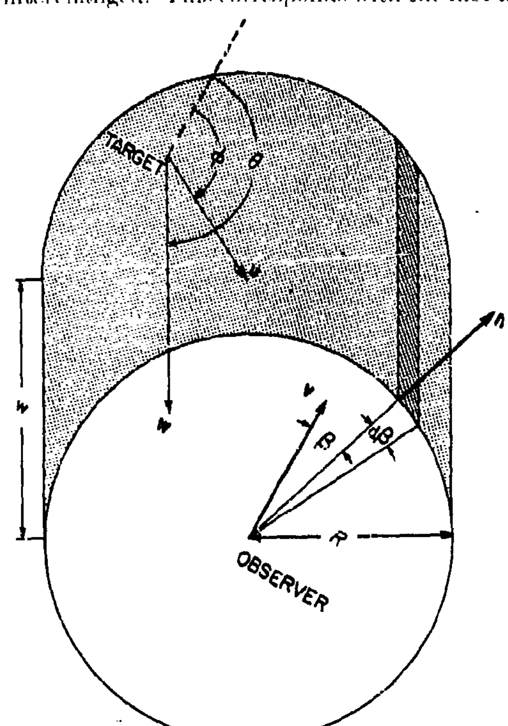


FIGURE 8. Area of targets entering a circle.

whenever the target comes within range R of the observer, the observer automatically comes within this range of the target, etc.

As an example, if there are 20 vessels distributed at random in an area of 10,000 square miles, so that $N = 0.002$, if they are cruising at 10 knots in various directions and if the observer is traveling at 15 knots, the number per hour arriving within the range of $R = 25$ miles of the observer is found to be 1.67, contrasted with 1.5, which would be the number in case the targets remained stationary. That the first number is greater than the second is due to the fact that the target's motion tends to bring more of them into the range than escape from getting within range —a fact which would not have been self-evident without calculation.

The preceding example illustrates the general principle that the contact rate on the random targets increases with increase of the motion of the targets. To show this, we have but to prove that $\partial N_0/\partial u$ is positive. Using the second form for N_0 in (4), but with the integral taken over half the interval of integration and doubled (a permissible change, in view of the symmetry of the integrand),

$$\frac{\partial N_0}{\partial u} = \frac{2RN}{\pi}\frac{\partial}{\partial u}\int_0^\pi \sqrt{u^2 + v^2 - 2uv\cos\phi}\, d\phi,$$

$$= \frac{2RN}{\pi}\int_0^\pi \frac{u - v\cos\phi}{w}\, d\phi,$$

$$= \frac{2RN}{\pi}\int_0^\pi \cos\omega\, d\phi,$$

where ω is the angle between the vectors **u** and **v** (cf. the various cases of Figure 2, with obvious constructions). The integrand is always positive when $v \leqq u$ (cases A and B), so that the required inequality $\partial N_0/\partial u > 0$ is evident. To show that this continues to be the case when $v > u$ [case C], decompose the interval of integration $(0,\pi)$ into the two halves $(0,\pi/2)$ and $(\pi/2,\pi)$, and then replace the variable of integration in the second half by the supplement of the ϕ of the first, thus recombining the integrals

$$\frac{\partial N_0}{\partial u} = \frac{2RN}{\pi}\int_0^{\pi/2}(\cos\omega + \cos\omega')d\phi;$$

here ω' denotes the value which ω assumes when ϕ is replaced by its supplement. Since a simple construction based on Figure 2C shows that ω' is less than the supplement of ω, we have $\cos\omega' > -\cos\omega$; i.e., $\cos\omega + \cos\omega' > 0$, and hence $\partial N_0/\partial u > 0$, as was to be proved. Application: If an enemy is passing in our vicinity but along an unknown path, we must cut our speed until he passes, if we wish to remain undetected.

Problem 2. Find the number of targets which enter the circle considered above between the bearings β and $\beta + d\beta$, per unit time. Their number is given by an expression of the form $N_0(\beta)d\beta$, where $N_0(\beta)$ is a density related to N_0 by the equation

$$N_0 = \int_0^{2\pi} N_0(\beta)d\beta. \tag{5}$$

To find $N_0(\beta)$, again we begin by considering only those targets of a particular track angle ϕ. They can enter the circle only if the direction of the vector **w** points *into* the circle, i.e., if the angle γ between the reversed vector $-\mathbf{w}$ and the unit vector **n** normal to the circle and pointing outward is *acute* (γ is defined as measured between 0 and π). Figure 8 shows that in this case the targets in question all come from the small heavily shaded region of area $Rw\cos\gamma\, d\beta = (-\mathbf{w}\cdot\mathbf{n})Rd\beta$ where $(-\mathbf{w}\cdot\mathbf{n})$ denotes the scalar product. The number per unit time is obtained by multiplying this expression by the density $Nd\phi/2\pi$. Hence, finally, the number $N_0(\beta)$ for targets of all track angles is given by

$$N_0(\beta) = \frac{RN}{2\pi}\int(-\mathbf{w}\cdot\mathbf{n})d\phi, \tag{6}$$

where the integration is over all those values of ϕ between 0 and 2π for which the integrand is *positive*.

For the evaluation of (6), observe that in view of the vector equation $\mathbf{w} = \mathbf{u} - \mathbf{v}$, (Section 1.2) we have

$$\begin{aligned}-\mathbf{w}\cdot\mathbf{n} &= \mathbf{v}\cdot\mathbf{n} - \mathbf{u}\cdot\mathbf{n}\\ &= v\cos\beta - u\cos(\phi - \beta).\end{aligned}$$

It remains to insert this value into (6), then to determine those values of ϕ for which $v\cos\beta - u\cos(\phi - \beta) > 0$, and finally to integrate over such values. The details of this straightforward computation are omitted. As a result, the following expressions are obtained.

When $v \leqq u$:

$$N_0(\beta) = \frac{NR}{\pi}\left[v\cos^{-1}\left(-\frac{v}{u}\cos\beta\right)\cos\beta + \sqrt{u^2 - v^2\cos^2\beta}\right]. \tag{7}$$

When $v > u$:

$$N_0(\beta) = NRv\cos\beta, \text{ when } -\cos^{-1}\frac{u}{v} \leqq \beta \leqq \cos^{-1}\frac{u}{v};$$

$$N_0(\beta) = \frac{NR}{\pi}\left[v\cos^{-1}\left(-\frac{v}{u}\cos\beta\right)\cos\beta + \sqrt{u^2 - v^2\cos^2\beta}\,\right], \quad (8)$$

when $-\cos^{-1}\left(-\frac{u}{v}\right) \leqq \beta \leqq -\cos^{-1}\left(\frac{u}{v}\right)$

or when $\cos^{-1}\left(\frac{u}{v}\right) \leqq \beta \leqq \cos^{-1}\left(-\frac{u}{v}\right)$;

$$N_0(\beta) = 0,$$

when $\beta \leqq -\cos^{-1}\left(-\frac{u}{v}\right)$

or $\beta \geqq \cos^{-1}\left(-\frac{u}{v}\right)$.

This result can be stated in terms of probabilities. Suppose that it is known that a target has entered the circle; *where* is it likely to have entered? If $p(\beta)d\beta$ is the probability that it entered between the bearings β and $\beta + d\beta$, the expected number $N_0(\beta)d\beta$ which enter per unit time is the number entering in unit time N_0 times the probability $p(\beta)d\beta$; thus

$$p(\beta) = \frac{N_0(\beta)}{N_0} = \frac{\pi N_0(\beta)}{4RN(u+v)E(\sigma)}, \quad (9)$$

in virtue of equation (4). Thus equations (7) and (8) give $p(\beta)$ at once.

Figure 9 gives the polar diagram showing the de-

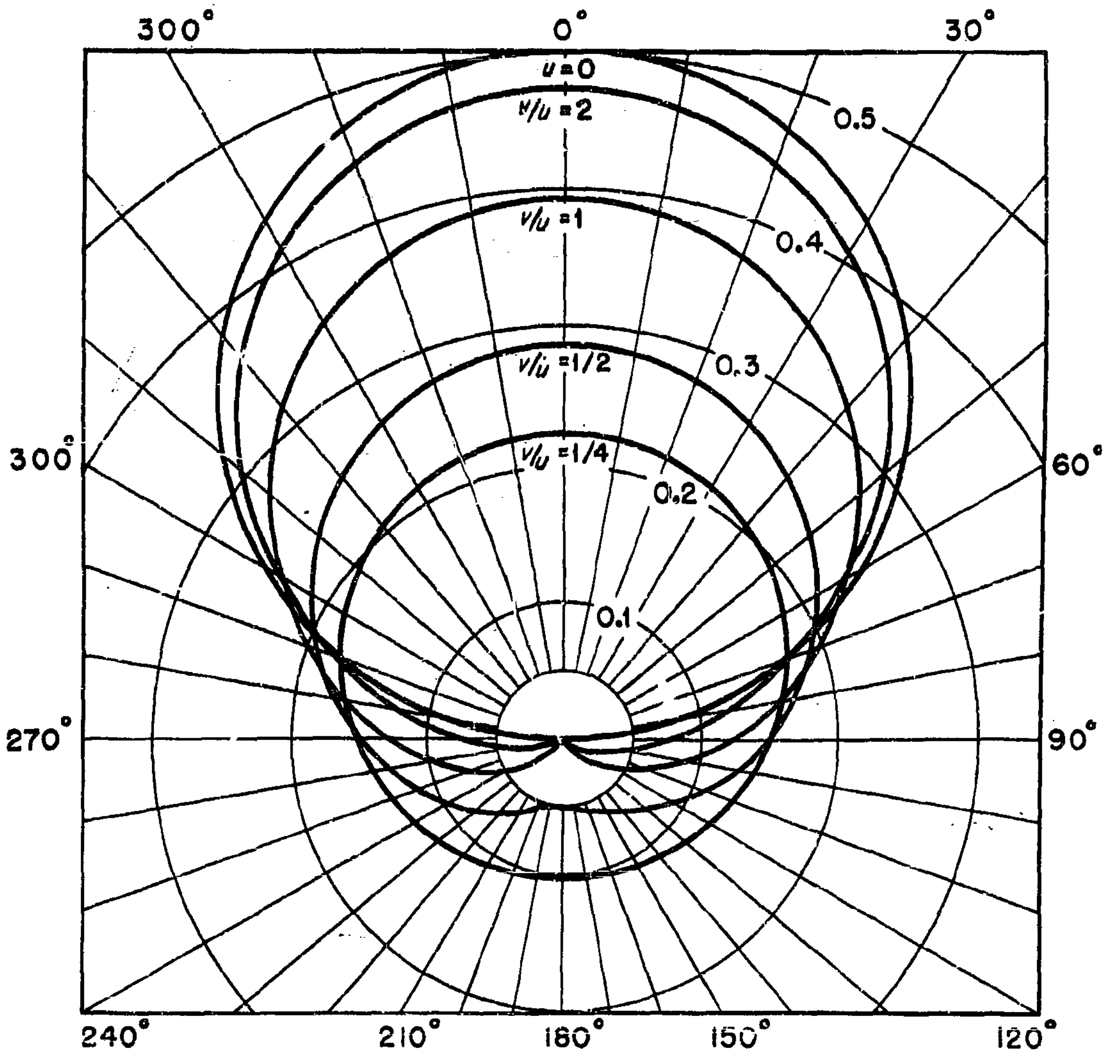


FIGURE 9. Polar diagram showing the dependence of $p(\beta)$.

pendence of $p(\beta)$ on β for different values of u/v. At one extreme, $u/v = 0$: The targets are at rest and the dependence on β is as the cosine, and the diagram is a circle with the observer at the circumference. At the other extreme $u/v = \infty$: the targets move but the observer is at rest; in this case the number entering at all bearings is the same and we have a circle centered at the target. In the intermediate cases, as long as target speed u is less than observer speed v, there is a certain angular range aft over which no contacts are made. As u becomes greater than v, however, the number of contacts made aft increases rapidly until in the limit as many are made aft as ahead.

There is a sort of inverse to this problem which it is useful to consider. Suppose that a contact has actually been made at the known bearing β (and range R), what is the distribution of values of the track angle ϕ? In other words, we have seen the target—what is its heading likely to be? This is essentially a problem in the "probability of causes" and is solved by Bayes' formula:[c]

$$\Pi_\beta(\phi) = \frac{\tilde{\omega}(\phi) f_\phi(\beta)}{\int \tilde{\omega}(\phi) f_\phi(\beta) d\phi},$$

where $\tilde{\omega}(\phi)d\phi$ is the "a priori probability" (i.e., as estimated before the contact was obtained) of a track angle between ϕ and $\phi + d\phi$; $f_\phi(\beta)$ is the "productive probability" of the effect observed (i.e., of a contact between the bearings β and $\beta + d\beta$), given that the target actually has the track angle ϕ; and, finally, $\Pi_\beta(\phi)d\phi$ is the "a posteriori probability" (i.e., as estimated after the contact at bearing β has been observed) that the target's track angle lies between ϕ and $\phi + d\phi$. In other words, $\Pi_\beta(\phi)$ is the quantity we want.

As before, $\tilde{\omega}(\phi) = 1/2\pi$. To obtain $f_\phi(\beta)d\beta$, observe that it equals the average number of targets detected in unit time between bearings β and $\beta + d\beta$, divided by the average number detected in unit time at all bearings (both averages for targets of given track angle ϕ). This quotient is calculated at once by means of the reasoning used before (based on Figure 8); it has the value

$$\frac{(-\mathbf{w}\cdot\mathbf{n})d\beta}{2\mathrm{w}} = \frac{v\cos\beta - u\cos(\phi-\beta)}{2\mathrm{w}}\,d\beta.$$

Thus
$$\Pi_\beta(\phi) = \frac{(-\mathbf{w}\cdot\mathbf{n})}{\int(-\mathbf{w}\cdot\mathbf{n})d\phi},$$

where the region of integration must be determined as in problem 2, since here again values of ϕ for which the integrand is negative are excluded. The results are,

when $v \leqq u$,

$$\Pi_\beta(\phi) = \frac{1}{2}\cdot\frac{\cos\beta - \frac{u}{v}\cos(\phi-\beta)}{\cos^{-1}\left(-\frac{v}{u}\cos\beta\right)\cos\beta + \sqrt{\left(\frac{u}{v}\right)^2 - \cos^2\beta}};$$

when $v > u$,

$$\Pi_\beta(\phi) = \frac{1}{2\pi}\left[1 - \frac{u}{v}\cdot\frac{\cos(\phi-\beta)}{\cos\beta}\right],$$

when $-\cos^{-1}\frac{u}{v} \leqq \beta \leqq \cos^{-1}\frac{u}{v}$;

$$\Pi_\beta(\phi) = \frac{1}{2}\frac{\cos\beta - \frac{u}{v}\cos(\phi-\beta)}{\cos^{-1}\left(-\frac{v}{u}\cos\beta\right)\cos\beta + \sqrt{\left(\frac{u}{v}\right)^2 - \cos^2\beta}},$$

when $-\cos^{-1}\left(-\frac{u}{v}\right) \leqq \beta \leqq -\cos^{-1}\frac{u}{v}$,

or when $\cos^{-1}\frac{u}{v} \leqq \beta \leqq \cos^{-1}\left(-\frac{u}{v}\right)$.

Detection is impossible when

$$\beta \leqq -\cos^{-1}\left(-\frac{u}{v}\right)$$

or
$$\beta \geqq \cos^{-1}\left(-\frac{u}{v}\right).$$

Problem 3. Given the relative position (r,β) of a target at a particular epoch t; find the probability P that the target will enter the circle of radius R centered on the observer. Find the curves of constant probability.

Evidently P depends on (r,β): $P = P(r,\beta)$, and $P = 1$ if $r \leqq R$. When $r > R$ the target will enter the circle if, and only if, its vector relative velocity $\mathbf{w}$ points into the circle (i.e., when $\mathbf{w}$ is produced in the direction it is pointing). The situation is illustrated in Figures 10A ($\mathrm{v} > \mathrm{u}$) and 10B ($\mathrm{v} < \mathrm{u}$), which show the angular range of vectors $\mathbf{w}$ pointing

[c] See, for example, *Probability and Its Engineering Uses*, T. C. Fry, D. Van Nostrand Co., New York.

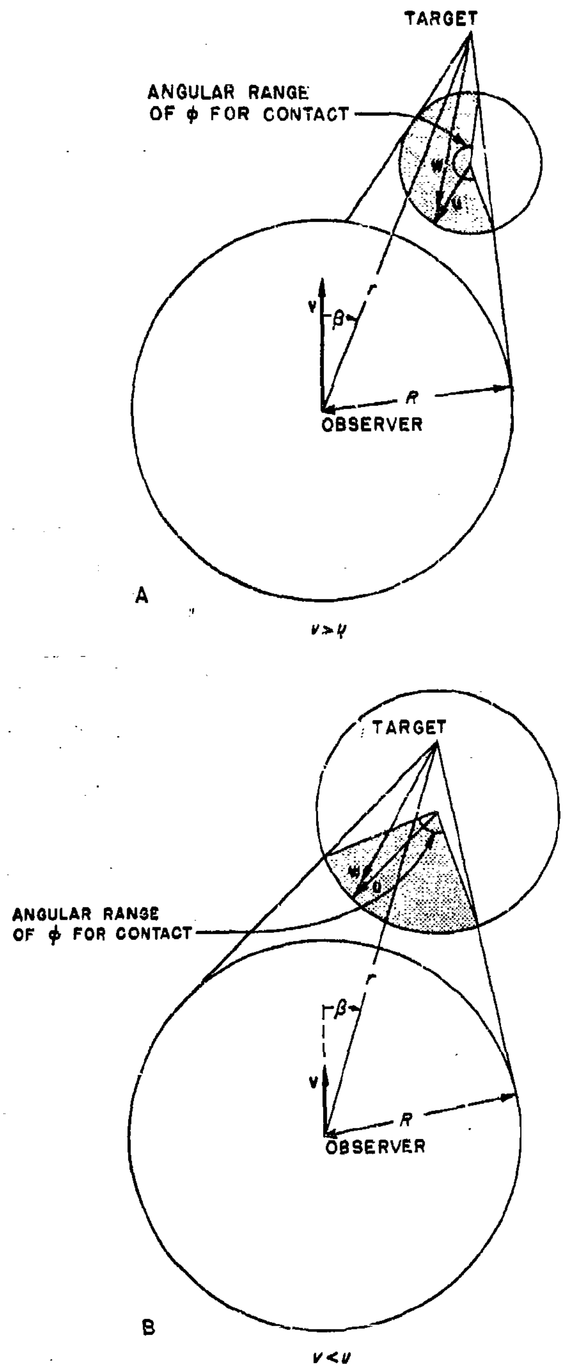


FIGURE 10. Probability of entering circle.

into the circle, i.e., the range of angles θ. Corresponding to this angular range of θ the angular range of ϕ is constructed immediately (shaded angle in Figure 10). On account of the uniformity of the distribution [in particular, of Section 1.4, (a) and (c)], the probability that ϕ lie in this angular range is the magnitude of the range divided by 2π. This is the required value of $P(r,\beta)$. The problem is thus reduced to the geometry of Figure 10, and the formula for $P(r,\beta)$ is obtained by straightforward trigonometry. There are, however, a number of different cases to be considered. For example, in Figure 10B the angular range of θ is between the two tangents from the target to the circle of radius R, while in Figure 10A it is between one such tangent and the tangent to the velocity diagram circle, corresponding to the restricted orientations of **w** when $v > u$ (limiting approach angle); moreover, in this case the θ range counts multiply: To one θ there are two ϕ's, one for the target moving toward the observer and the other for the target moving away and being overtaken. There are other cases not shown in Figure 10.

The expression of $P(r,\beta)$ is as follows:

$$P(r,\beta) = \frac{\Phi(r,\beta)}{2\pi},$$

where $\Phi(r,\beta)$ is the total radian length of the range or ranges of values of ϕ $(0 \leqq \phi < 2\pi)$ which satisfy the inequality

$$r^2[u \sin(\beta - \phi) - v \sin \beta]^2 \leqq R^2[u^2 + v^2 - 2uv \cos \phi],$$

subject to the condition

$$u \cos(\beta - \phi) \leqq v \cos \beta.$$

The second of the above inequalities is needed to insure that the target enter the circle of radius R after the reference epoch t. It is automatically satisfied for those values of ϕ which satisfy the first inequality when $v \geqq u$.

The curves of constant probability are symmetrical with respect to the course of the observer. In the following discussion, only the half-plane to the right of the observer's course is considered. First, consider the case when $v \geqq u$ and $k = v/u$ (Figure 11). Outside the circle of radius R and between the tangents to this circle which are inclined to the right and the left of the course of the observer at the limiting approach angle $\sin^{-1} u/v$, the curve of constant probability P is a straight line tangent to the circle and inclined to the observer's course at the angle $\sin^{-1} [(u/v) \cos \pi P]$. When $v = u$, this latter angle reduces to $(\pi/2)(1 - 2P)$. On and below the lower tangent line inclined at the angle $\sin^{-1} u/v$, $P = 0$.

Above the upper tangent line inclined at the angle $\sin^{-1} u/v$, the equation of the curve of constant probability P is

$$r^2 = -\frac{R^2 v^2 \csc^2\left(\frac{\pi P}{2}\right)\left[\sin^2\beta - \cos^2\left(\frac{\pi P}{2}\right)\right]}{v^2 \sin^2\beta - u^2 \cos^2\left(\frac{\pi P}{2}\right)}.$$

When $v = u$, this equation reduces to $r = R \csc(\pi P/2)$.

Second, consider $v < u$ and $k = v/u$ (Figure 12). Outside the circle of radius R, the equation of the curve of constant probability P is

$$v^2 \sin^2\beta\,[\cos^2(\psi + \pi P) - \cos^2\psi]$$
$$= \sin^2(\psi + \pi P)\,[u^2\cos^2(\psi + \pi P) - v^2\cos^2\psi],$$

where ψ is the positive acute angle $\cos^{-1}(R/r)$. Figure 13 shows how the equiprobability curve $P = 0.25$ varies with $k = v/u$. Figure 14 is for later reference.

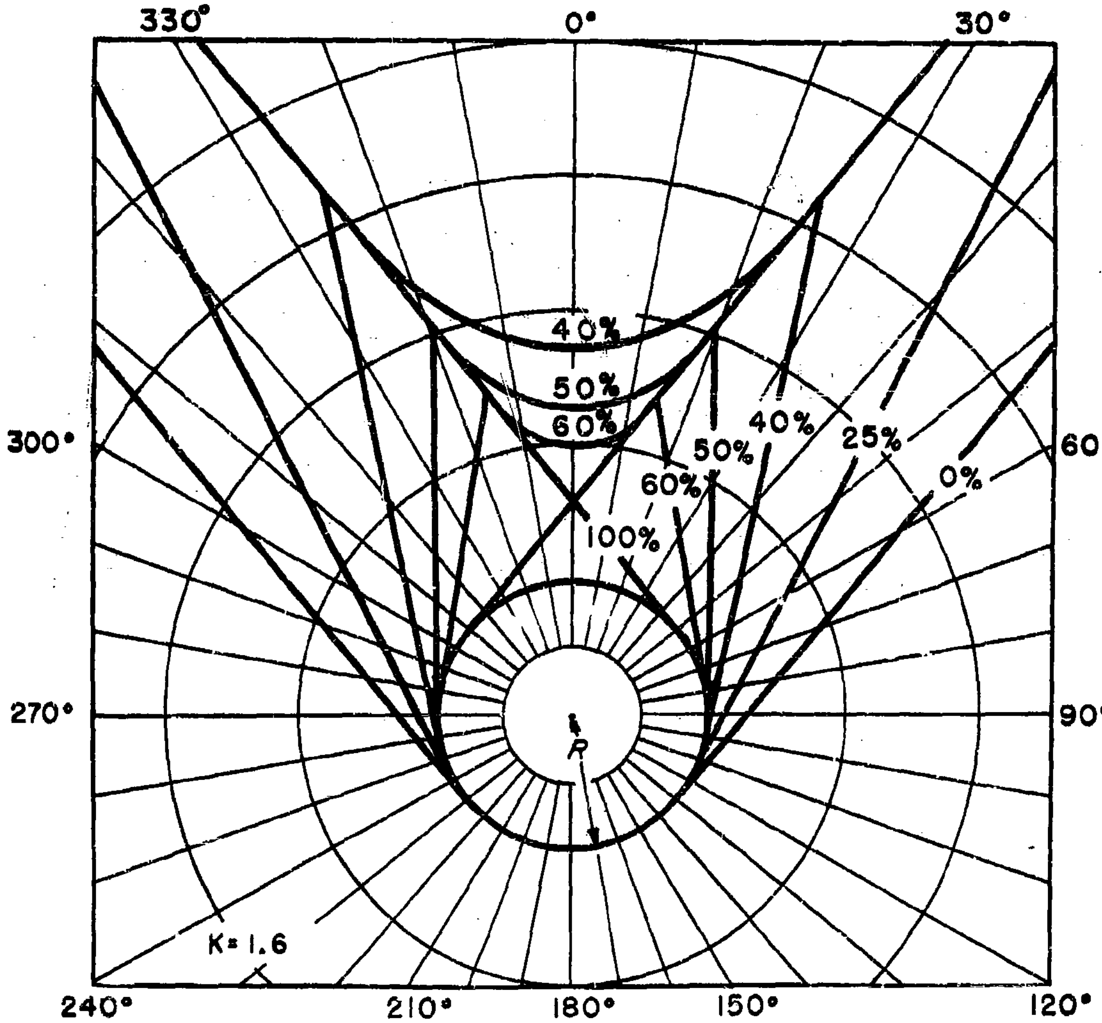


FIGURE 11. Contact probability curves. $k = 1.6$.

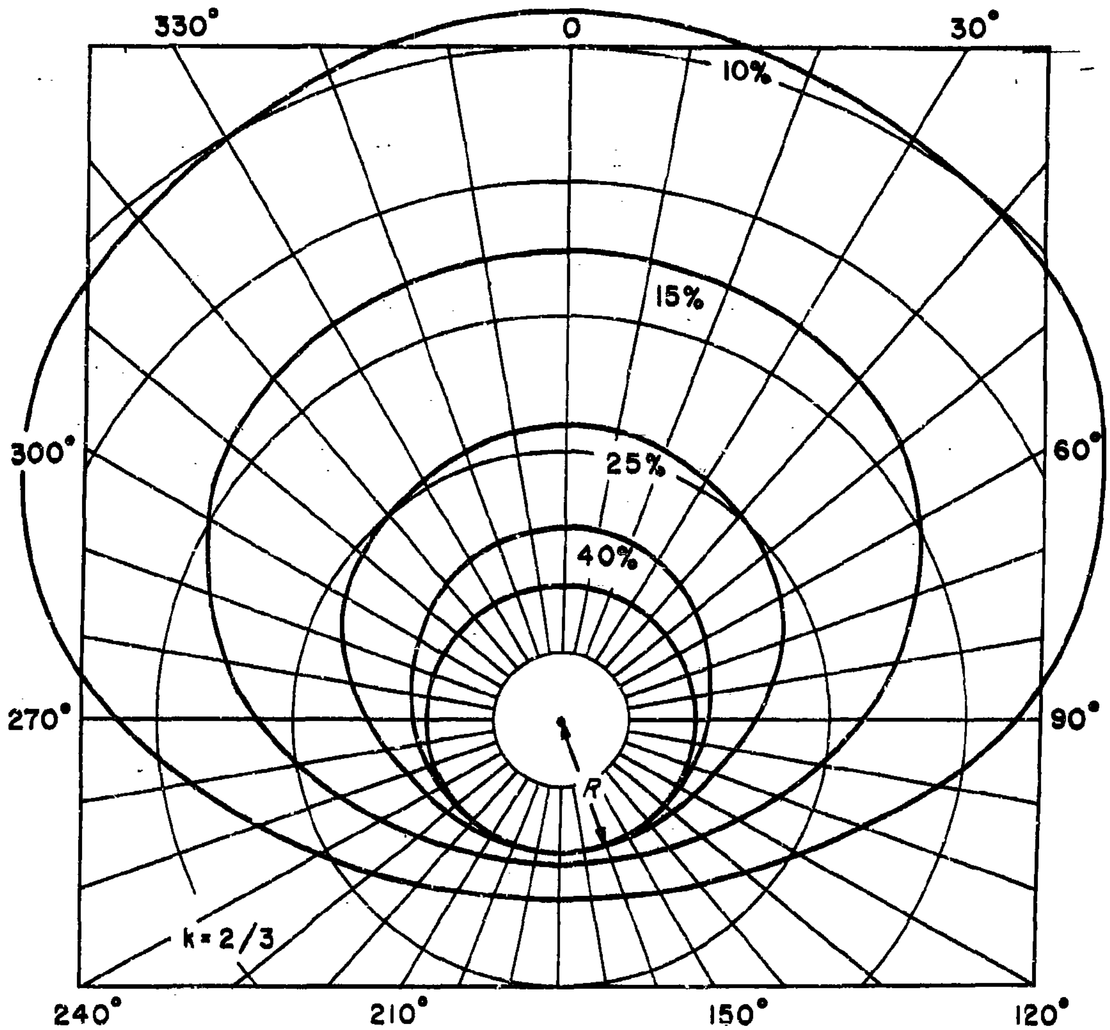


FIGURE 12. Contact probability curves $k = 2/3$.

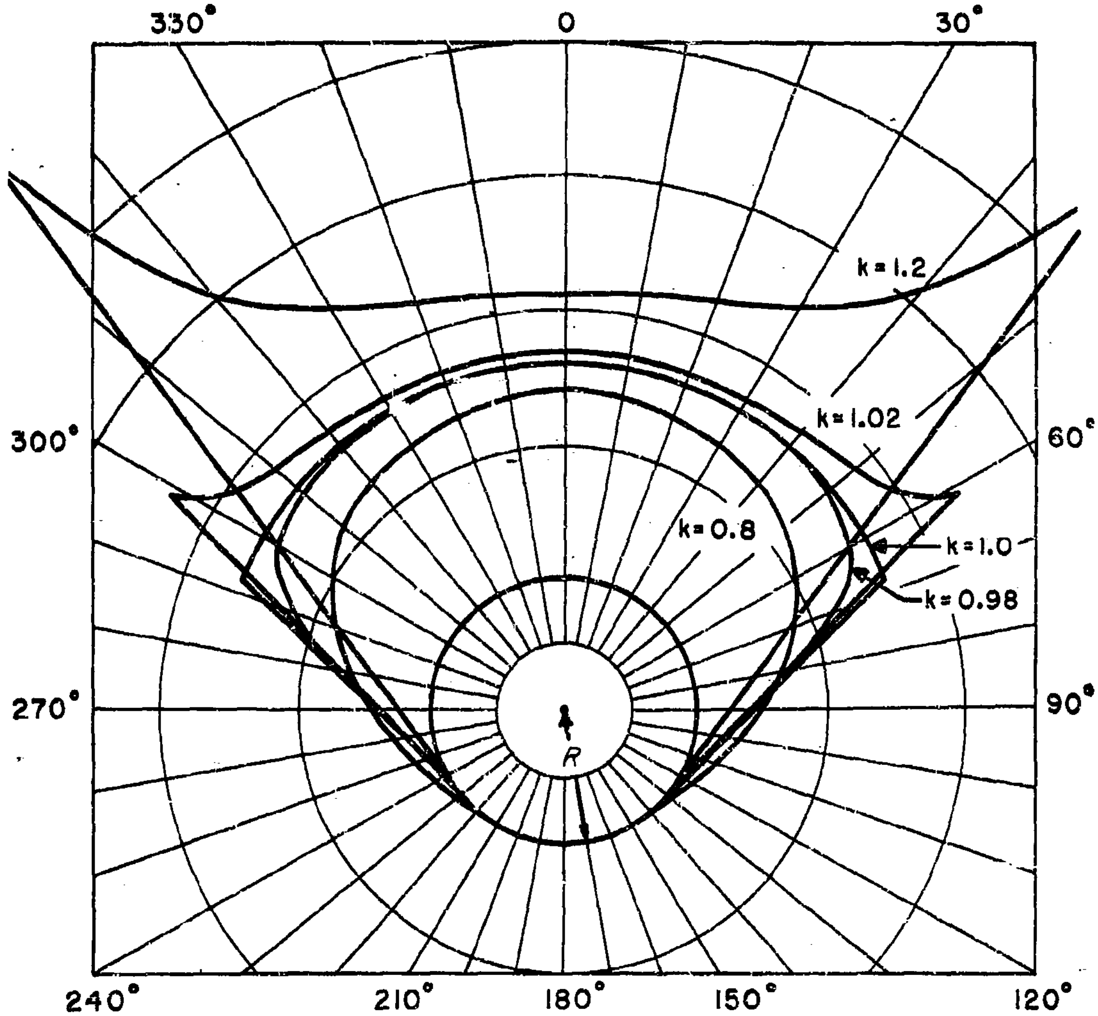


FIGURE 13. 25 per cent contact probability curve.

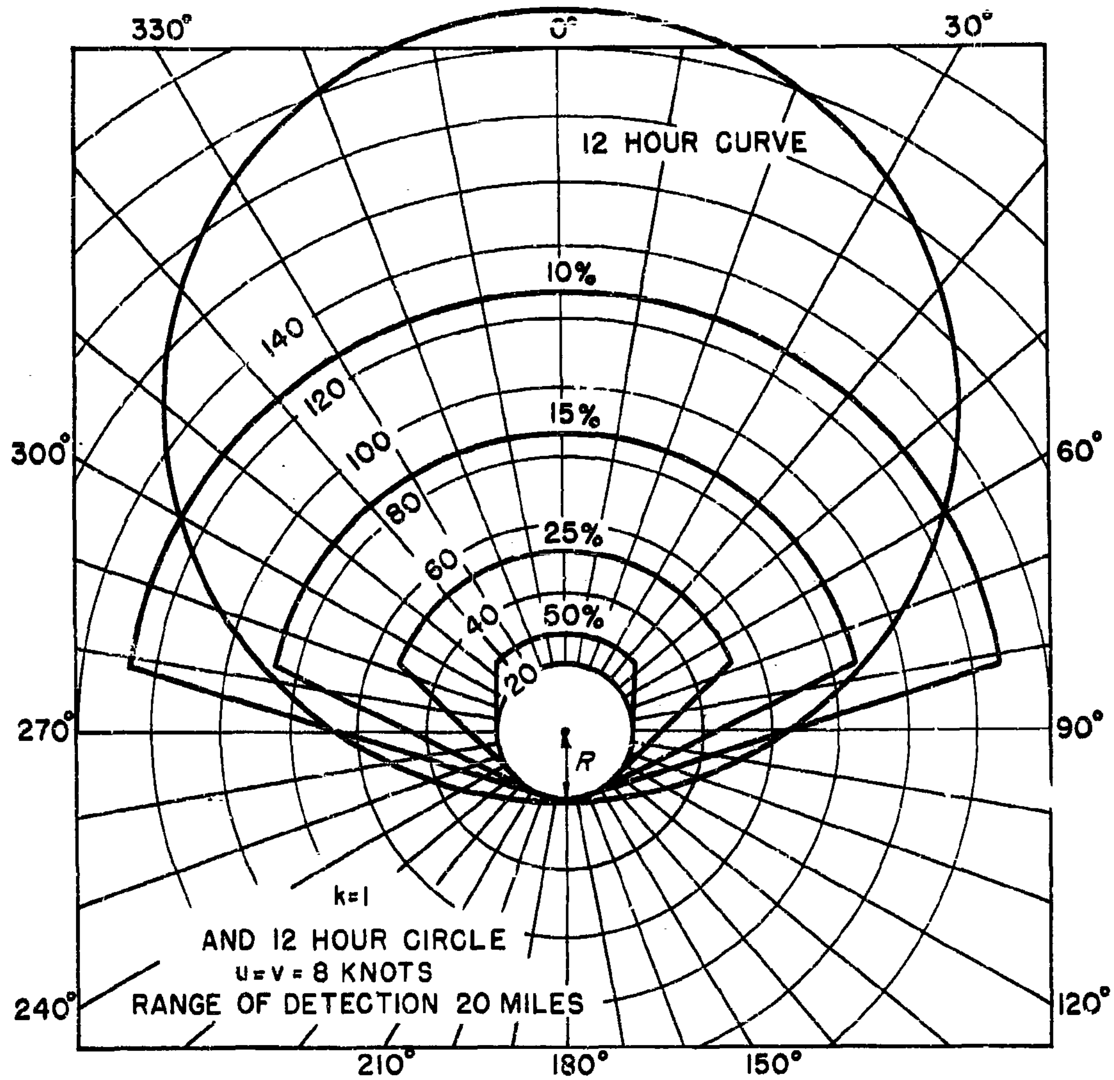


FIGURE 14. Contact probability curves. $k = 1$, and 12 hour circle. $u = v = 8$ knots. Range of detection 20 miles.

1.6 NONUNIFORM DISTRIBUTIONS OF TARGETS

The uniform distribution of targets considered hitherto is by no means the only important case which arises in connection with naval operations. One example will illustrate this point; it will be used later on in the subject of antisubmarine hunts.

A target has been detected inaccurately. All that is known is that it is more likely to be at a certain point O (the "fix") than at any other point, but may not be at O but only within a short distance of O, all points at the same distance r from O being equally likely, and the probability falling rapidly to a negligible value as the distance r increases beyond a few miles. If $f(r)dA$ denotes the probability that the target be in the area dA r miles from O, the graph of $f(r)$ against r will be of the character shown in Figure 15.

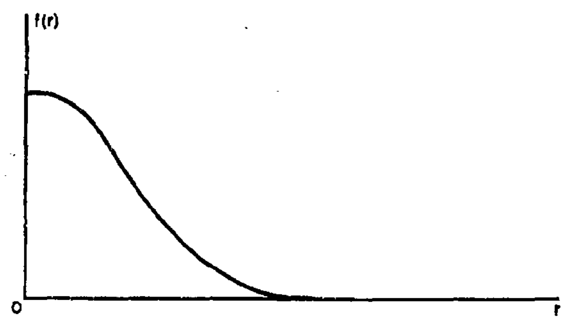


FIGURE 15. The distribution of targets about a point of fix O.

Under such conditions it is almost always possible to approximate to the situation with sufficient accuracy by assuming a *circular or elliptical normal distribution;* we shall assume the former, i.e., that $f(r)$ is proportional to $e^{-r^2/2\sigma^2}$, where σ is a constant (the standard deviation) which increases the more the graph of $f(r)$ is spread out, that is, the more indefinite the knowledge of the target's position. The constant of proportionality is found by the fact that, since the target is surely somewhere, the integral $\iint f(r)dA \left(= \iint f(r) r dr d\psi\right)$ extended over the whole surface has the value unity. Thus one obtains

$$f(r) = \frac{1}{2\pi\sigma^2} e^{-r^2/2\sigma^2}.$$

Now suppose that t hours elapse after the time of the observation. The target will have moved and the distribution will no longer be the same. Assume that the speed of the target can be estimated with satisfactory accuracy, but not its direction: u is known, but not $\mathbf{u}$, i.e., not ϕ. It is natural to assume further that all directions are equally likely and are independent of the actual position of the target. Thus the distribution complies with Section 1.4 (a) and (c), for a uniform distribution, but not with Section 1.4 (b). The problem is to find the new distribution $f(r,t)$ after t hours.

Consider first the case of a target whose vector velocity $\mathbf{u}$ makes a given angle γ with the direction from O to the contemplated position [γ is measured as usual from vector $\mathbf{r}$ (from O to reference point in dA) to vector u]. This target will be in dA if and only if it had initially been in a region congruent to dA and situated ut miles away in the direction of

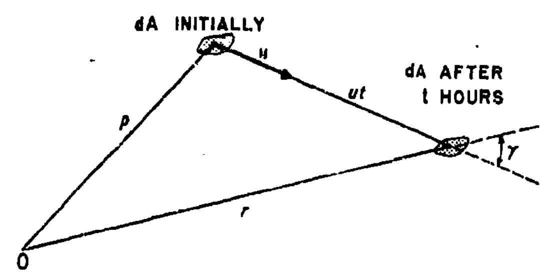


FIGURE 16. Entry of target into dA.

the reversed vector $-\mathbf{u}$, as shown in Figure 16. The probability of this event is

$$\frac{1}{2\pi\sigma^2} e^{-\rho^2/2\sigma^2} dA = \frac{1}{2\pi\sigma^2} e^{-(r^2 + u^2t^2 - 2rut\cos\gamma)/2\sigma^2} dA.$$

Now the probability that $\mathbf{u}$ make an angle with $\mathbf{r}$ between γ and $\gamma + d\gamma$ is $d\gamma/2\pi$. The probability of both these events is the product of these two probabilities, and to obtain the required total probability $f(r,t)dA$ this product is added (integrated) over all possible initial positions of dA, i.e., over all values of γ from 0 to 2π:

$$f(r,t) = \frac{1}{2\pi\sigma^2} \cdot \frac{1}{2\pi} \int_0^{2\pi} e^{-(r^2 + u^2t^2 - 2rut\cos\gamma)/2\sigma^2} d\gamma$$

$$= \frac{1}{2\pi\sigma^2} e^{-(r^2 + u^2t^2)/2\sigma^2} \cdot \frac{1}{2\pi} \int_0^{2\pi} e^{2rut\cos\gamma/2\sigma^2} d\gamma.$$

Now we have

$$\frac{1}{2\pi}\int_0^{2\pi} e^{rut\cos\gamma/\sigma^2}d\gamma = \frac{1}{2\pi}\int_0^{2\pi} e^{-rut\cos\gamma/\sigma^2}d\gamma$$

$$= \frac{1}{2\pi}\int_0^{2\pi} e^{i(irut\cos\gamma/\sigma^2)}d\gamma$$

$$= J_0\left(i\frac{rut}{\sigma^2}\right) = I_0\left(\frac{rut}{\sigma^2}\right),$$

where $i = \sqrt{-1}$ and J_0 denotes the ordinary Bessel function of zeroth order, I_0 its value for pure imaginary values of the argument. Thus the equation

$$f(r,t) = \frac{1}{2\pi\sigma^2}e^{-(r^2+u^2t^2)/2\sigma^2}I_0\left(\frac{rut}{\sigma^2}\right). \tag{10}$$

The graph of $f(r,t)$ for different values of t is shown in Figure 17. It is seen how the probability spreads outward with time, so that the target is most likely to be in an expanding ring about O.

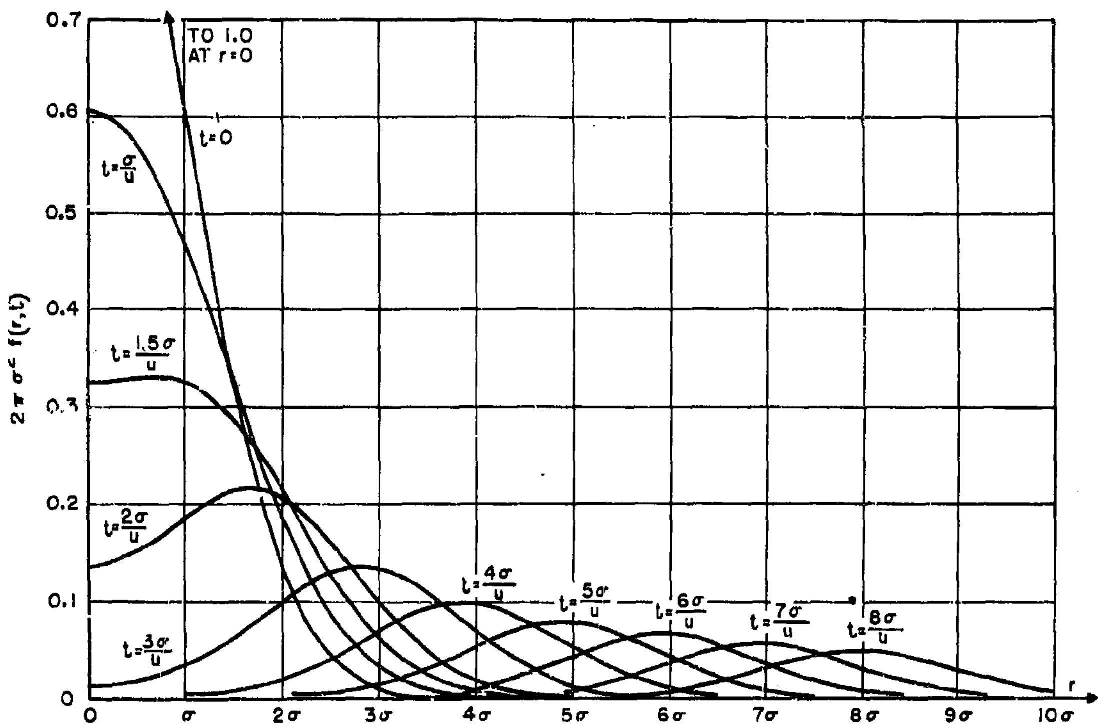


FIGURE 17. The distribution of moving targets about a point of fix O.

Chapter 2

TARGET DETECTION

2.1 GENERALITIES CONCERNING DETECTION

The first chapter has dealt with the positions, motions, and contacts of observer and target, but has left entirely out of consideration the act or process whereby the observer gains knowledge of the presence and position of the target. Contacts have been considered as purely geometrical events and their probabilities have been simply the probabilities that the target reaches a specified position in relation to the observer. The present chapter will be concerned, on the other hand, with the act of *detection*, that event constituted by the observer's becoming aware of the presence and possibly of the position and even in some cases the motion of the target by visual sighting, by radar detection, by hydrophonic listening, by echo ranging, or by any other means whatsoever. There are certain general ideas and methods which apply to all cases of detection, and their study is the object of this chapter. But before quantitative results of immediate practical applicability can be obtained, a detailed study must be made of the special instrumentality of detection; this is done for the visual case in Chapter 4, for radar in Chapter 5, and for sonar in Chapter 6.

Two basic facts underlie every type of detection:

(i) *There is a certain set of physical requirements which have to be met if detection is to be possible, and which if met will in fact make detection possible, though not necessarily inevitable.* Thus targets must obviously not be too far away: their view from the observer must not be completely obstructed; to be seen there must be some illumination; the radar will not reveal them if the atmospheric conditions or background echoes are too bad; sonar detection requires that the sound path be not completely bent away from the observer by water refraction; etc.

(ii) *Even when the physical conditions make detection possible, it will by no means inevitably occur: Detection is an event which under definite conditions has a definite probability, the numerical value of which may be zero or unity or anything in between.* Thus when the target just barely fulfills the physical conditions for possible detection, the probability of detection will be close to zero (at least when the time for observation is very limited). As the conditions improve the chance of detection increases, and it may become close to or equal to unity: detection becomes practically certain. Experience in everyday life shows that we may be looking for an object in plain sight and yet sometimes fail to find it. Cases are known where observational aircraft flying on clear sunny days on observational missions have passed close over large ships and yet failed to detect them. And a host of operational statistics give further confirmation of this point. It must be constantly realized that every instrumentality of detection is based in last analysis on a human being, and its success is accordingly influenced by his attention, alertness, and fatigue, and the whole chain of events which occur between the impact of the message on his sense organs and his mental response thereto. Furthermore, even under physical conditions which are as fixed and constant as it is practicable to make them, innumerable rapid fluctuations in them are still apt to occur (a radar target changes its aspect from moment to moment with the continual rocking of the ship, sonar ranges experience short-term oscillations about their mean, etc.); and as a result, a target which may not be detected at one instant may be detected if sought a moment later.

In view of (i), one part of the study of detection requires the physical conditions for detection to be explored; in view of (ii), the other part requires the probabilities of detection, when the former conditions are given, to be obtained.

In Section 2.9 the effect of statistically combining observations made under operational conditions in which the physical situation is not constant is considered.

2.2 INSTANTANEOUS PROBABILITIES OF DETECTION

Suppose that the physical conditions (distances, etc.) remain fixed and that the observer is looking for the target (by "looking" shall be meant trying to detect with the means considered, visual, radar, sonar, etc). There are two possibilities: First, the observer may be making a succession of brief "glimpses"; a typical case of this is in the echo-ranging procedure in which each sweep or scan affords one opportunity for detection (glimpse), successive ones occurring two or three minutes apart. Second, the observer may be looking continuously; a typical case is the observer fixing his eyes steadily on the position where he is trying to detect the target. The case of radar is intermediate; on account of the scanning it would belong to the first case, but if the scanning is very fast, and especially when there is persistency of the image on the scope, it may be treated as in the second. Likewise, visual detection by a slow scan through a large angle belongs to the first rather than the second case. Very often the decision to regard a method of detection in the first or in the second way depends simply on which affords the closest or most convenient approximation. This will be made clear on the basis of examples in Chapters 4, 5, and 6.

In the case of separated glimpses, the important quantity is the *instantaneous probability g of detection by one glimpse.* When n glimpses are made under unchanging conditions the probability p_n of detection is given by the formula

$$p_n = 1 - (1 - g)^n. \tag{1}$$

This is because $1 - p_n$ is the probability of failing to detect with n glimpses, and for this to occur the target must fail to be detected at every single one of the n glimpses; each such failure having the probability $1 - g$ and the n failures being independent events, we conclude that $1 - p_n = (1 - g)^n$; hence (1). When $g = 0$, obviously $p_n = 0$, but if $g > 0$ and even if g is very small, p_n can be made as close to 1 as we please by increasing n sufficiently; in other words, once the physical conditions give some chance, however small, of detecting on one glimpse, enough glimpses under the same conditions will lead with practical certainty to eventual detection.

To find the mean or expected number $\bar{n}$ of glimpses for detection we must first find the probability P_n that detection shall occur precisely at the nth glimpse (and not before). This is the product of the probability that it shall not occur during the first $n - 1$ glimpses, $(1 - g)^{n-1}$, times the probability that a detection shall occur on a single glimpse (the nth), g; it is accordingly $P_n = (1 - g)^{n-1}g$. The required mean number $\bar{n}$ is, according to the theory of probability, $1P_1 + 2P_2 + 3P_3 + \cdots$, and thus

$$\begin{aligned} \bar{n} &= \sum_{n=1}^{\infty} n(1-g)^{n-1}g \\ &= g + 2(1-g)g + 3(1-g)^2 g + \cdots \\ &= -g\frac{d}{dg}\left[1 + (1-g) + (1-g)^2 + \cdots\right] \\ &= -g\frac{d}{dg}\frac{1}{1-(1-g)} \\ &= -g\frac{d}{dg}\left(\frac{1}{g}\right) \\ &= \frac{1}{g}. \end{aligned} \tag{2}$$

Turning to the case of continuous looking, the important quantity is the *probability γdt of detecting in a short time interval of length dt.* The quantity γ is called the *instantaneous probability density* (of detection). When the looking is done continuously during a time t under unchanging conditions, the probability $p(t)$ of detection is given by

$$p(t) = 1 - e^{-\gamma t}. \tag{3}$$

To prove this, consider $q(t) = 1 - p(t)$, the probability of failure of detection during the time t. For detection to fail during the time $t + dt$ [probability $= q(t + dt)$], detection must fail both during t [probability $= q(t)$] and during dt (probability $= 1 - \gamma dt$), and multiplying these probabilities of independent events we obtain

$$q(t + dt) = q(t)(1 - \gamma dt)$$

which is equivalent to the differential equation

$$\frac{dq(t)}{dt} = -\gamma q(t).$$

The solution of this equation on the assumption that $q(0) = 1$ (no detection when no time is given to looking) is $q(t) = e^{-\gamma t}$; whence (3). Again it is true that if there is the least chance of detection in time dt (i.e., if $\gamma > 0$) the chance of detection increases to virtual certainty as the looking time t becomes sufficiently large. It may be observed that the quantity γt in the exponent of (3) represents the mean or expected number of targets detected by an observer passing through a swarm of unit density of targets uniformly distributed over the ocean.

To find the mean or expected time $\bar{t}$ at which detection occurs, observe that the probability $P(t)dt$ of detection between t and $t + dt$ (when looking has been continuing from the initial time 0) is the product of probability of no detection before t times probability of a detection during dt, i.e., $P(t)dt = e^{-\gamma t}\gamma dt$. $\bar{t}$ is found by integration

$$\bar{t} = \int_0^\infty t e^{-\gamma t}\gamma dt = \frac{1}{\gamma}. \tag{4}$$

Figure 1 shows the graphs of the probability $p(t)$ of detection *during* the time t and $P(t)$ of detection *at* the time t and gives the construction of $\bar{t}$ as the abscissa of the intercept with the horizontal line of unit ordinate of the tangent to $p(t)$ at the origin.

Since equation (3) reduces to equation (1) when γ is taken as $\gamma = -\log(1 - g)$ and $t = n$ (glimpses one unit of time apart), Figure 1 serves to show the quantitative behavior of p_n and P_n: the difference is that only discrete points ($t = 1, 2, 3, \cdots$) on the curve

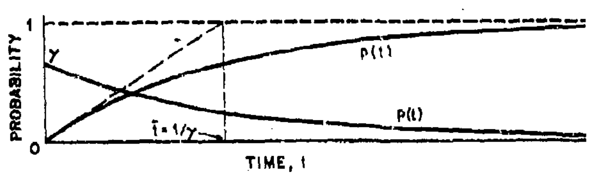


FIGURE 1. Probabilities of detection under fixed conditions.

are used, and n is no longer given by the tangent intercept but rather by a secant intercept.[a]

When, as usually occurs in actual search, the distances and hence the probability quantities g or γ change as time goes on, (1) must be replaced by

$$p_n = 1 - \prod_{i=1}^{n}(1 - g_i) = 1 - (1 - g_1)(1 - g_2)(1 - g_3)\cdots \tag{5}$$

which takes into account the fact that g will change from glimpse to glimpse: g_i is the probability of detection for the ith glimpse. And (3) must be replaced by

$$p(t) = 1 - e^{-\int_0^t \gamma_t dt}, \tag{6}$$

where in γ_t the possible change in the probability density of detection as time goes on is put into evidence by the subscript. The reasoning leading to these equations is precisely similar to that in the earlier case. But the probabilities p_n, $p(t)$ do not necessarily approach unity as n or t increase, thus when $\int_0^\infty \gamma_t dt$ is finite, the chance of detection $p(t)$ never exceeds $1 - e^{-\int_0^\infty \gamma_t dt} < 1$. Here again the quantity $\int_0^\infty \gamma_t dt$ in the exponent represents the expected number of targets detected as the observer passes through a swarm of targets distributed uniformly at unit density over the ocean.

The instantaneous probabilities quantities g and γ depend, as we have said, on the sum total of physical conditions. For example, in visual detection γ depends on the range r from target to observer, on the meteorological state (illumination and haze), on the size and brightness of target against the background, on the observer's facilities, altitude, etc. And corresponding lists can be made out for radar and sonar detection. Throughout the remainder of the present chapter, only the dependence on range will be explicitly considered, i.e., we shall write

$$g = g(r), \quad \gamma = \gamma(r). \tag{7}$$

It will be legitimate to apply the results either when all the other conditions remain practically unchanged during the operation considered, or when the other conditions have been shown not to influence the results to the degree of approximation that is accepted.

Since the instantaneous probability quantities tend to decrease to zero as the range r increases and to be large when the range is small, their graph against r will be of the character shown in Figure 2. Case A is when the instantaneous probability density reaches a finite maximum at zero range (probability of detecting target when flying over target is less than unity). In case B this maximum is infinite (probability of detection when flying over target is unity). In case C the effect of sea return on radar diminishes the probability of detection when over target. In case D, the instantaneous probability is *infinite* when $r < R$: detection is sure to occur as soon as the target gets within this critical range R.

The last case, while not altogether realistic, is often not very far from the truth. A very useful rough approximation is to assume further that the instantaneous probability is *zero* for $r > R$. Then detection is sure and immediate within the range R and is impossible beyond R. This assumption shall be called the *definite range law of detection*.

An important example showing the evaluation of the function $\gamma(r)$ is in the case when the following assumptions are made.

1. The observer is at height h above the ocean, on which the target is cruising.

[a] Throughout this book, log is used to denote "natural logarithm," and $\log_{10}$ to denote "common logarithm."

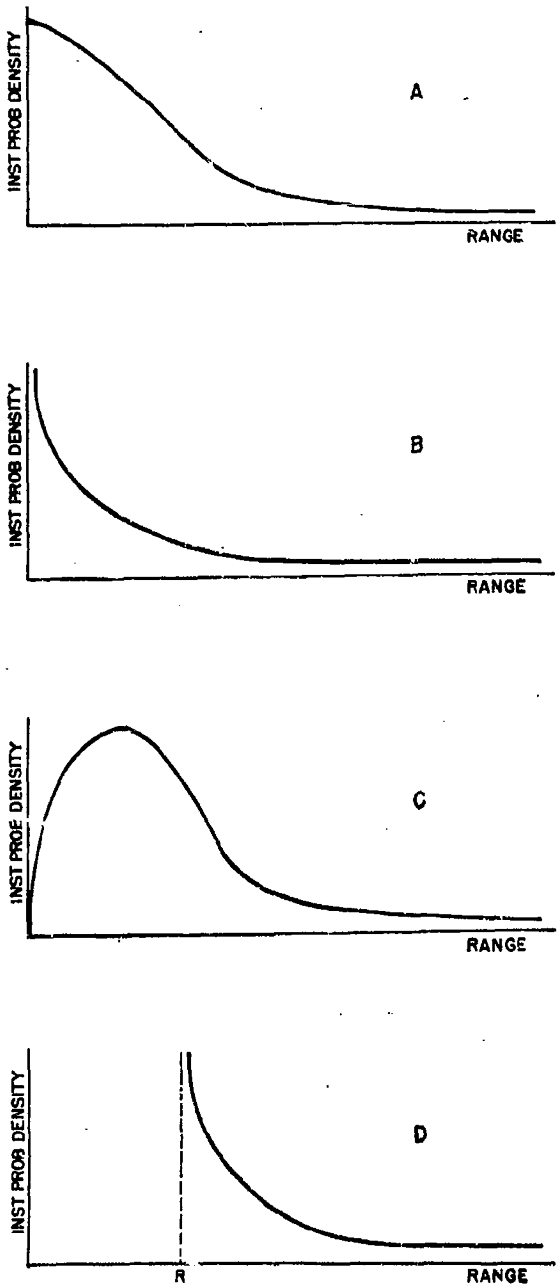


FIGURE 2. Instantaneous probability at various distances.

2. The observer detects the target by seeing its wake.

3. The instantaneous probability of detection γ is proportional to the solid angle subtended at the point of observation by the wake.

The calculation of the solid angle is shown in Figure 3 for an area of ocean which is a rectangle of length a toward the observer and width b perpendicular to the direction of observation (perpendicular to the page in Figure 3A). The infinitesimal solid angle is the product of the angle α subtended by a, and the angle β subtended by b. The radian measure of α is c/s. By similar triangles, $c/a = h/s$ and hence

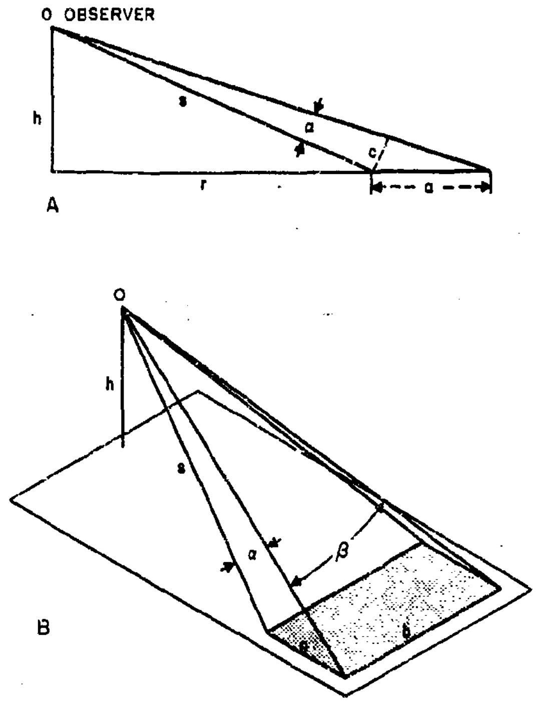


FIGURE 3. Solid angle subtended by wake.

$\alpha = ah/s^2$. And the radian measure of β is obviously b/s. Hence solid angle $= \alpha\beta = abh/s^3 =$ area of rectangle times h/s^3. The actual area A of the target's wake is not rectangular, but can be regarded as made up of a large number of rectangles like the above, the solid angle being the sum of the corresponding solid angles. Hence, when the dimensions of A are small in comparison with h, r, and s, we have the formula

$$\text{Solid angle} = \frac{Ah}{s^3} = \frac{Ah}{(h^2 + r^2)^{3/2}}. \qquad (8)$$

Since γ is assumed to be proportional to the solid angle, we obtain

$$\gamma = \frac{kh}{s^3} = \frac{kh}{(h^2 + r^2)^{3/2}}, \qquad (9)$$

where the constant k depends on all the factors which we are regarding as fixed and not introducing explicitly, such as contrast of wake against ocean, observer's ability (number of lookouts and their facilities), meteorological conditions, etc.; and of course k contains A as a factor. Dimensionally, $k = [L^2T^{-1}]$. In the majority of cases r is much larger than h, and (9) can be replaced by the satisfactory approximation

$$\gamma = \frac{kh}{r^3}. \tag{10}$$

Formulas (9) and (10) lead to cases A and B respectively of Figure 2; the property of detection which they express shall be called the *inverse cube law* of sighting. When the subject of vision is studied in Chapter 4 it will be found that many changes in this law have to be made to obtain a high degree of approximation under the various conditions of practice. Nevertheless the inverse cube law gives a remarkably useful approximation. Its use in the present chapter is chiefly as an illustration of the general principles.

2.3 DEPENDENCE OF DETECTION ON TRACK

When the observer and target are moving over the ocean in their respective paths, which may be straight or curved and at constant or changing speeds, the continuous change in their relative positions constantly changes the instantaneous probability of de-

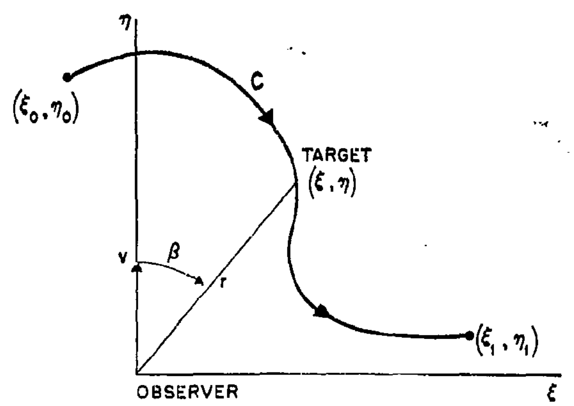


Figure 4. Target's relative track.

tection; we have to deal with the functions g_t and γ_t and calculate probabilities of detection by means of formulas (5) and (6). It is convenient to draw the target's track C (Figure 4) relative to the observer. The latter need not be moving in fixed course and speed over the ocean, although that is very often the case. The coordinates used have been described in Section 1.2 (see Figure 4 and equation (2) of Chapter 1.

The target is at (ξ,η) at the time t, so that the equations of the target's relative motion are

$$\xi = \xi(t), \quad \eta = \eta(t), \tag{11}$$

where initially $(t = t')$ $\xi_0 = \xi(t')$, $\eta_0 = \eta(t')$, and finally $(t = t'')$ $\xi_1 = \xi(t'')$, $\eta_1 = \eta(t'')$. The target describes the relative track C. Accordingly, (7) becomes [writing $\xi^2(t)$ for $\{\xi(t)\}^2$]:

$$\begin{aligned} g &= g\left(\sqrt{\xi^2(t) + \eta^2(t)}\right) = g_t \\ \gamma &= \gamma\left(\sqrt{\xi^2(t) + \eta^2(t)}\right) = \gamma_t. \end{aligned} \tag{12}$$

Hence according to equations (5) and (6) the probabilities p_C of detection are given by either of the following

$$p_C = 1 - \prod_{i=1}^{n}\left[1 - g\left(\sqrt{\xi^2(t_i) + \eta^2(t_i)}\right)\right], \tag{13}$$

$$p_C - 1 - \exp\left[-\int_{t'}^{t''} \gamma\left(\sqrt{\xi^2(t) + \eta^2(t)}\right)dt\right]. \tag{14}$$

In (13) t_i is the time (epoch) of the ith "glimpse" or scan, and n the number of glimpses between t' and t'':

$$t' \leqq t_1 < t_2 < \cdots < t_n \leqq t''.$$

In (14), the integral is actually a *line integral* along C; if w is the relative speed (not necessarily constant), we may write [with s = arc length of C from (ξ_0, η_0)]:

$$p_C = 1 - e^{-\int_C \gamma(r)ds/w}. \tag{15}$$

Formulas (13) and (14) may be united into

$$p_C = 1 - e^{-F[C]}, \tag{16}$$

where for the case of *separate glimpses*

$$F[C] = -\sum_{i=1}^{n} \log\left[1 - g\left(\sqrt{\xi^2(t_i) + \eta^2(t_i)}\right)\right], \tag{17}$$

and in the case of *continuous looking*

$$F[C] = \int_C \gamma(r)\frac{ds}{w}. \tag{18}$$

This quantity $F[C]$ shall be called the *sighting potential.* It has the important property of *additivity:* If C_1 and C_2 are two tracks and $C = C_1 + C_2$ is their combination or sum, and if $p_C = p_{C_1+C_2}$ is the probability of sighting on at least one track, $p_{C_1+C_2}$ is still obtained by formula (16) and

$$F[C_1 + C_2] = F[C_1] + F[C_2]. \tag{19}$$

This is an immediate consequence of the usual equation for combining probabilities of events which may not be mutually exclusive:

$$p_C = 1 - (1 - p_{C_1})(1 - p_{C_2}) = p_{C_1} + p_{C_2} - p_{C_1}p_{C_2}.$$

The additivity applies, of course, to the sum of any number of paths. One application is to the calculation of p_C when C is complicated, but made up out of a sum of simple pieces such as straight lines. Another application is in the case of two or more inter-communicating observers; C_1 can be the path of the target relative to the first and C_2 that relative to the second, etc.

A most important case, and one which will chiefly concern us in this book, is when both observer and target are moving at constant speed and course. The results of Chapter 1 become applicable. Track C is a straight line, and the speed w is a constant (as long as C is not turned). It is convenient to make the calculations with the aid of the coordinates (x,y) of Chapter 1 (Figure 5), where x is the lateral range. The equations of motion which take the place of (11) are $x =$ constant, $y = wt$, where t is measured from the epoch of closest approach, and where, furthermore, the positive direction of the y axis is that of the target's relative motion; this convention is used throughout this chapter. The potential $F[C]$ is given by the appropriate one of the formulas

$$\begin{aligned} F[C] &= -\sum_{i=1}^{n} \log\left[1 - g\left(\sqrt{x^2 + w^2t_i^2}\right)\right], \\ &= -\sum_{i=1}^{n} \log\left[1 - g\left(\sqrt{x^2 + y_i^2}\right)\right]; \end{aligned} \tag{20}$$

$$\begin{aligned} F[C] &= \int_{t'}^{t''} \gamma\left(\sqrt{x^2 + w^2t^2}\right)dt, \\ &= \frac{1}{w}\int_{y'}^{y''} \gamma\left(\sqrt{x^2 + y^2}\right)dy, \end{aligned} \tag{21}$$

where y_i is the distance of target at the ith glimpse to its closest position, and (x',y') and (x'',y'') are the

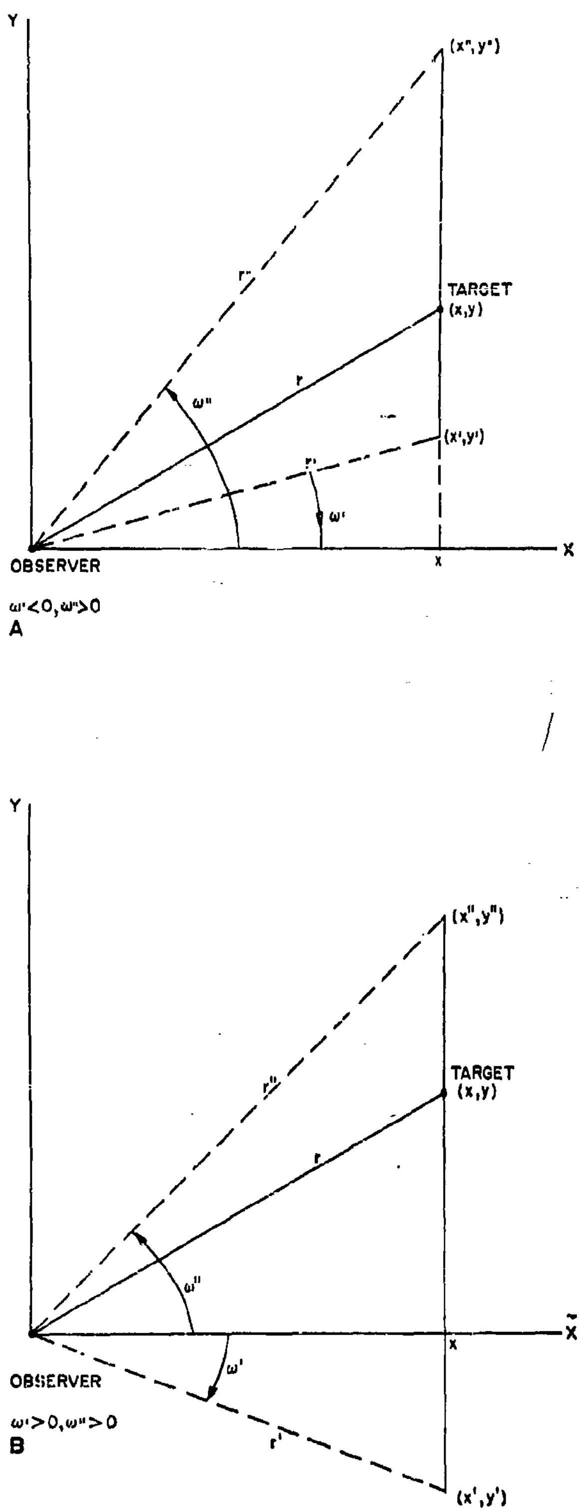


FIGURE 5. Detection at fixed speed and course.

extremities of C: $x' = x'' = x =$ constant, $y' = wt'$, $y'' = wt''$.

In the case of the inverse cube law (9),

$$F[C] = \frac{kh}{w}\int_{y'}^{y''}\frac{dy}{(h^2 + x^2 + y^2)^{3/2}}$$
$$= \frac{m}{h^2 + x^2}\left(\frac{y''}{\sqrt{h^2 + x^2 + (y'')^2}} - \frac{y'}{\sqrt{h^2 + x^2 + (y')^2}}\right). \qquad (22)$$

And for (10),

$$F[C] = \frac{kh}{w}\int_{y'}^{y''}\frac{dy}{(x^2 + y^2)^{3/2}} = \frac{m}{x^2}\left(\frac{y''}{r''} - \frac{y'}{r'}\right)$$
$$= \frac{m}{x^2}(\sin\omega' + \sin\omega''), \qquad (23)$$

where in each case

$$m = \frac{kh}{w}, \qquad (24)$$

and where r' and r'' are the ranges of the extremities of C, and ω' and ω'' the angles they subtend with the normal to C.

2.4 THE LATERAL RANGE DISTRIBUTION

When the observer and target are on their straight courses at constant speeds for a long time before and after their closest approach, the probability $p(x)$ of detection is a function of the lateral range x. The graph of $p(x)$ against x is called the *lateral range curve* and expresses the *distribution in lateral range*. In consequence of (16) $p(x)$ is given by

$$p(x) = 1 - e^{-F(x)}, \qquad (25)$$

where $F(x)$ is the value of $F[C_x]$, C_x being an infinite straight line at the perpendicular distance x from the observer. The value of $F(x)$ is found by applying equation (20), summing over all integral values of i, or equation (21) with $y' = -\infty$ and $y'' = \infty$, in the glimpse or the continuous looking cases, respectively.

With continuous looking (21) applies. For the definite range law, $p(x) = 1$ or 0 according as $-R < x < R$ or not, and the lateral range curve is Figure 6A. For the inverse cube law,

$$p(x) = 1 - e^{-2m/(h^2+x^2)} \text{ or } p(x) = 1 - e^{-2m/x^2}, \qquad (26)$$

according to whether (22) or (23) is used; the curve is shown in Figure 6B in the former case.

With intermittent glimpses taking place T units of time apart, equation (20) applies. For the definite range law, $p(x) = 0$ when $x > R$ or $x < -R$, and $p(x) = 1$ when the length $2\sqrt{R^2 - x^2}$ of relative track during which the target is within range R of the observer is greater than wT, i.e., when

$$-\sqrt{R^2 - w^2T^2/4} \leqq x \leqq \sqrt{R^2 - w^2T^2/4}\,;$$

but
$$p(x) = \frac{2\sqrt{R^2 - x^2}}{wT}$$

in intermediate cases, this being the probability that the target be glimpsed while within range R. The lateral range curve is shown in Figure 6C.

Other typical lateral range curves are those of Figure 6D and E. The dip at $x = 0$ in Figure 6E shows the effect of sea return (radar) or pinging over the target (sonar).

The area W under the lateral range curve is called the *effective search (or sweep) width*:

$$W = \int_{-\infty}^{+\infty} p(x)dx. \qquad (27)$$

It has the following interpretation. If the observer moves through a swarm of targets uniformly distributed over the surface of the ocean (N per unit area on the average) and either all at rest or all moving with the same vector velocity $\mathbf{u}$, the average number N_0 detected per unit time is

$$N_0 = NwW. \qquad (28)$$

For suppose that t is such a long period of time that the length of time during which a target is within range of possible detection is small in comparison with t. Then the number of targets passing during the period t through detection range (i.e., exposing themselves to detection) and having the lateral range between x and $x + dx$ is $Nwtdx$ (since such targets are in an area of $wtdx$ square miles). On the average $p(x)Nwtdx$ of these will be detected. Hence the average total number detected is

$$\int_{-\infty}^{+\infty} p(x)\, Nwtdx.$$

Dividing this by t and applying equation (27), equation (28) is obtained. Since for continuous looking with a definite range law, $W = 2R$, we may describe W as follows:

The effective search width is twice the range of a *definite range law of detection which is equivalent to the given law of detection in the sense that each of the two laws detects the same number of uniformly distributed targets of identical velocity.*

The product wW is called the *effective search (or sweep) rate.*

When the distribution of targets is uniform in the sense of Section 1.3, i.e., when their speed is given so that the search width is proportional to the square root of the altitude and inversely proportional to the square root of the target's relative speed. Furthermore, if there are n aircraft flying the same path without mutual interference (or if there are n observers having the same facilities operating independently of one another in the same aircraft), W is replaced by $W\sqrt{n}$.

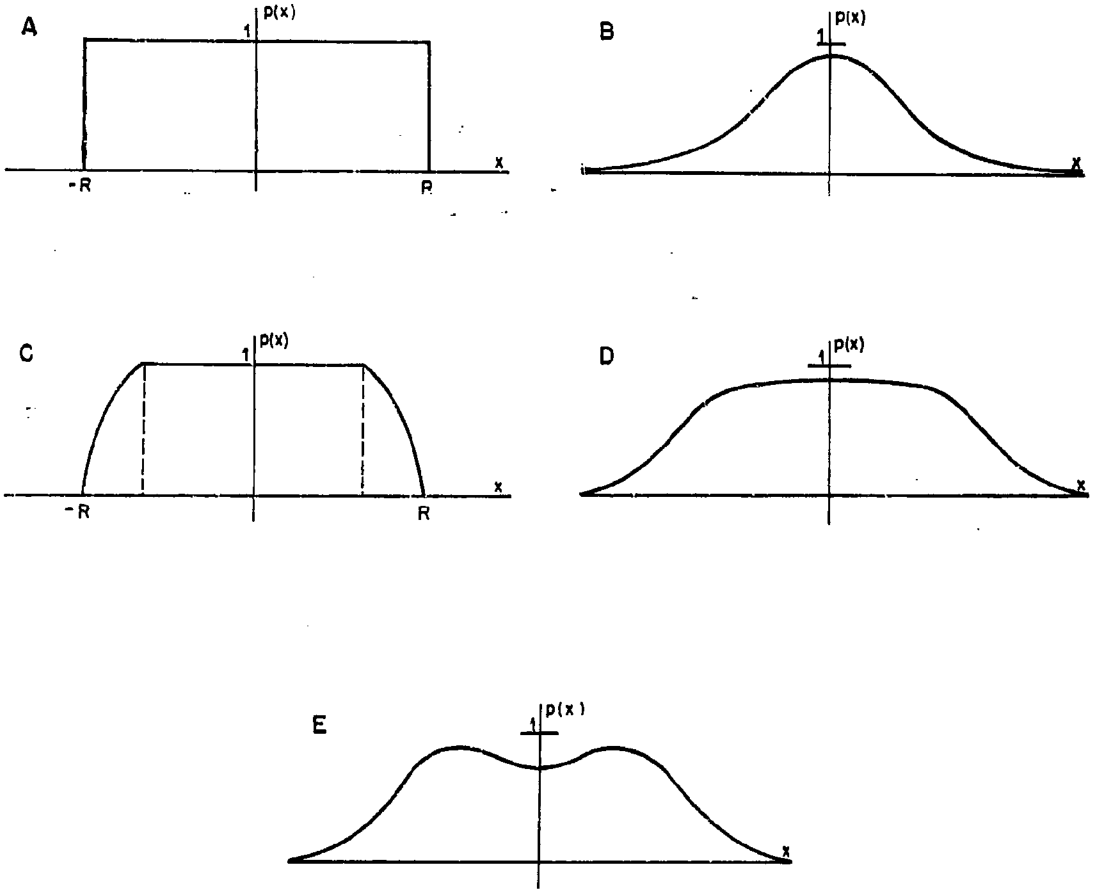


FIGURE 6. Lateral range curves.

but their course is not, w has to be replaced by its average $\bar{w}$ (taken as uniformly distributed in track angle ϕ) i.e., we must write $w = \frac{1}{2\pi}\int_0^{2\pi} w\,d\phi$, so that (28), $N_0 = N\bar{w}W$, may hold ($W = 2R$). [See Chapter 1, equations (1) and (4)].

In the case of the simplified inverse cube law equations (26) and (27) give, by carrying out the integration (see below):

$$W = 2\sqrt{2\pi m} = 2\sqrt{\frac{2\pi kh}{w}}, \qquad (29)$$

This results from the additivity of the potentials Section 2.3, which has the effect that k is replaced by nk in equations (22) and (23), and thus that m is replaced by nm [see (24)]. Thus the statement that W is replaced by $W\sqrt{n}$ is a consequence of equation (29).

The integration leading to (29) is performed by introducing equation (26) into (27) and changing to the new variable of integration:

$$z = \frac{\sqrt{2m}}{x};$$

and then integrating by parts. Use is made of the well-known equation

$$\int_0^\infty e^{-z^2} dz = \frac{\sqrt{\pi}}{2}.$$

By its definition, $p(x)$ is the probability (*not* probability density) that a target, known to have the lateral range x, be detected. On the other hand, $p(x)dx/W$ is the probability that a target, known to have been detected, have a lateral range between x and $x + dx$ (in this case $p(x)/W$ *is* a probability density). This fact (actually a consequence of Bayes' theorem in probability) is easily seen, as follows: The detected target may be thought of as chosen at random from the set of all detected targets; the chance that its lateral range be between x and $x + dx$ is equal to the proportion of targets in this set which have such a lateral range; from the previous calculations, this proportion is seen to be

$$\frac{Nwp(x)dx}{NwW} = \frac{p(x)dx}{W}.$$

2.5 THE DISTRIBUTION IN TRUE RANGE

Again we suppose that the observer makes constant speed and course and that the targets do likewise and are distributed uniformly over the surface of the ocean with the density N (average number per unit area). Relative to the observer, the targets all move parallel to the y axis in the direction of

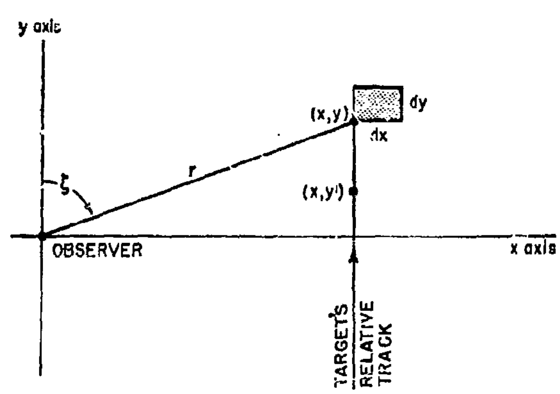


FIGURE 7. Target detection at given relative position.

increasing y. *How many targets are detected on the average per unit time in the small region of area $dxdy$ of Figure 7?* The number will be proportional to N and to $dxdy$, and may accordingly be represented by $N\rho(x,y)\,dxdy$, where $\rho(x,y)$, which may be described as the rate of first contacts at the point (x,y) per unit area and per unit density of targets, is obtained by the argument which follows.

The number of targets entering $dxdy$ in unit time is $Nwdx$. A given target's probability of being detected therein is the product of the probability that it fail to be detected before entering this region times the probability that, when not previously detected, it be detected while crossing $dxdy$, i.e., during the time $dt = dy/w$. The former probability is $e^{-F(x,y)}$ in virtue of equation (16), where $F(x,y)$ is given in the case of glimpses by equation (20), with i summed over all values for which $y_i < y$ (with sufficient accuracy we may write $y_i = y - iwT$ and sum for $i = 1, 2, \cdots \infty$); and in the case of continuous looking, by equation (21) with $y' = -\infty$ and $y'' = y$. The latter probability is given by $g(r)dy/wT$ (intermittent glimpses, one every T units of time, dy/wT being the probability that a glimpse occur while target is in $dxdy$), or by $\gamma(r)dy/w$ (continuous looking). Thus the probability is

$$\frac{e^{-F(x,y)}g(r)dy}{wT} \quad \text{or} \quad \frac{e^{-F(x,y)}\gamma(r)dy}{w},$$

according to whether glimpsing or continuous looking is used. To obtain the mean number of detections per unit time in $dxdy$, these expressions are multiplied by the number of targets exposed to such detection, $Nwdx$. Hence the answer to the question in italics above is supplied by the following expressions for $\rho(x,y)$:

$$\rho(x,y) = \frac{e^{-F(x,y)}g(r)}{T}$$
$$F(x,y) = -\sum_{i=1}^{\infty} \log\left[1 - g\left(\sqrt{x^2 + (y - iwT)^2}\right)\right] \tag{30}$$

for intermittent glimpsing, and

$$\rho(x,y) = e^{-F(x,y)}\gamma(r)$$
$$F(x,y) = \frac{1}{w}\int_{-\infty}^{y}\gamma\left(\sqrt{x^2 + y^2}\right) dy \tag{31}$$

for continuous looking.

It is seen by carrying out the differentiation that in the case of equation (31),

$$\rho(x,y) = w\frac{\partial}{\partial y}\left[1 - e^{-F(x,y)}\right]. \tag{32}$$

In the case of (30), the corresponding formula is

$$\rho(x,y) = w\Delta_y[1 - e^{-F(x,y)}], \tag{33}$$

where the operation Δ_y applied to a function denotes the result of the following process: first, replace y in the function by $y + \Delta y$, $(\Delta_y = wT)$; second, subtract the original value of the function from the new; third, divide by Δy.

If A is a plane region moving with the observer over the ocean, the average number Q_A of targets detected per unit time within A is (by addition of averages)

$$Q_A = N \iint_A \rho(x,y)dxdy. \tag{34}$$

In particular, when A embraces the whole plane, equations (32) and (33) lead from equation (34) to the previously obtained expression $N_0 = NwW$ of (28) by straightforward calculation. When $A = A_R$ is a circle of radius R centered on the observer, (34) expressed in polar coordinates (r,ζ) (ζ = angle from positive y axis to vector $\mathbf{r}$ drawn from observer to target) becomes:

$$Q(R) = N \iint_{A_R} \rho(x,y)dxdy$$
$$= N \int_0^R dr \int_0^{2\pi} r\rho(r \sin \zeta, r \cos \zeta)\, d\zeta.$$

Now the number of targets detected per unit time at a distance (true range) from the observer between r and $r + dr$ is of the form $N\rho(r)dr$ (being proportional to both N and dr), and since its integral from O to R must, for every value of R, be equal to $Q(R)$, it follows (by equating the two integral expressions for $Q(R)$ and differentiating through with respect to R, etc.) that

$$\rho(r) = \int_0^{2\pi} r\rho(r \sin \zeta, r \cos \zeta)\, d\zeta. \tag{35}$$

$\rho(r)dr$ may be described as the rate of detection in the range interval $(r, r + dr)$ at unit target density.

If, now, a target is known to be detected but at unknown range, the probability that the range of detection has been between r and $r + dr$ is $\rho(r)dr/wW$. For this target may be thought of as chosen at random from the set of all the NwW detected targets, of which there are $N\rho(r)dr$ detected at range between r and $r + dr$. Hence the probability that the target be detected at such a range is the quotient

$$\frac{N\rho(r)dr}{NwW} = \frac{\rho(r)dr}{wW}.$$

The function $\rho(r)$ (or the equivalent functions $N\rho(r)$ or $\rho(r)/wW$) expresses the *distribution in (true) range*, and the graphs of these functions against r are called *range curves*. They fall considerably for small values of r, since relatively few targets come close to the observer by chance, and of these a still smaller number are apt to survive undetected up to a close proximity of the observer. Figure 8 shows a typical range curve (actually, for the inverse cube law); as the situation approaches the definite range law, the curve humps up indefinitely about the value $r = R$ of the definite range, and falls to the axis of abscissas elsewhere.

The mean value of the range of detection is given by

$$\bar{r} = \frac{1}{wW}\int_0^\infty r\rho(r)dr = \frac{1}{wW}\int_0^{2\pi} r^2\rho(r \sin \zeta, r \cos \zeta)\, d\zeta \tag{36}$$

in all cases.

In the case of the simplified inverse cube law, we have equation (23), in which we set $\omega' = \pi/2 - \zeta$ and $\omega'' = \pi/2$; we obtain

$$F[C_y] = \frac{m}{x^2}(1 + \cos \zeta) = \frac{m}{r^2}\frac{1 + \cos \zeta}{\sin^2 \zeta},$$
$$= \frac{m}{2r^2}\csc^2 \frac{\zeta}{2}.$$

Thus

$$\rho(x,y) = w \exp\left(-\frac{m}{2r^2}\csc^2\frac{\zeta}{2}\right)\frac{m}{r^3}.$$

Hence[b]

$$\rho(r) = \frac{wm}{r^2}\int_0^{2\pi} \exp\left(-\frac{m}{2r^2}\csc^2\frac{\zeta}{2}\right)d\zeta, \tag{37}$$
$$= 2wm\pi\frac{1}{r^2}\left[1 - \operatorname{erf}\left(\frac{1}{r}\sqrt{\frac{m}{2}}\right)\right].$$

[b]This depends on the evaluation of the integral

$$\phi(\lambda) = \int_0^{\pi/2} e^{-\lambda \csc^2 \theta}\, d\theta.$$

This is done by the following device: Differentiating with respect to λ,

$$\phi'(\lambda) = -\int_0^{\pi/2} e^{-\lambda \csc^2 \theta} \csc^2\theta\, d\theta$$
$$= e^{-\lambda} \int_\infty^0 e^{-\lambda \cot^2 \theta}\, d \cot \theta$$
$$= -\frac{1}{\sqrt{\lambda}} e^{-\lambda} \frac{\sqrt{\pi}}{2},$$

as is seen on changing to the variable of integration $x = \sqrt{\lambda} \cot \theta$, and the use of the formula $\int_0^\infty e^{-x^2} dx = \sqrt{\pi}/2$. Integrating $\phi'(\lambda)$ with respect to λ, observing that $\phi(0) = \pi/2$, we obtain

$$\phi(\lambda) = \frac{\pi}{2} - \sqrt{\pi}\int_0^\lambda e^{-\lambda}\frac{d\lambda}{2\sqrt{\lambda}} = \frac{\pi}{2}[1 - \operatorname{erf}\sqrt{\lambda}], \tag{38}$$

as it appears on changing to the variable of integration $\mu = \sqrt{\lambda}$.

where "erf X" (the "error function" or "probability integral") is defined as

$$\text{erf } X = \frac{2}{\sqrt{\pi}}\int_0^x e^{-x^2}dx.$$

By expressing m in terms of the search width W by means of equation (29), equation (37) is reduced to

$$\rho(r) = \frac{mW^2}{4r^2}\left[1 - \text{erf}\left(\frac{W}{4\sqrt{\pi}\, r}\right)\right]. \qquad (39)$$

This is the function actually graphed in Figure 8.

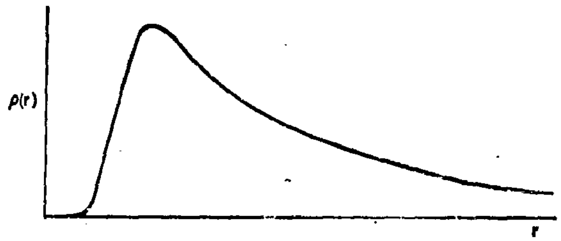


FIGURE 8. True range curve.

For values of r not over about 15 miles, it is in reasonable rough agreement with operational data (visual); but farther out it has too high an ordinate.

2.6 RANDOM SEARCH

In the last two sections both observer and target were on straight courses at constant speeds; this represents the extreme of simplicity of paths. At the other extreme is the case where both are moving in complicated paths over the ocean and at speeds which may vary in the course of time, a case which is called that of *random search*. If the position of the target is in the area of interest A (which may be many hundred square miles) in which the observer is moving, and if the observer is without preknowledge indicating that the target is more likely to be in one part of A than in another, a good approximation to the probability p that the observer make a contact is given on the basis of the following three assumptions:

1. The target's position is uniformly distributed in A.

2. The observer's path is random in A in the sense that it can be thought of as having its different (not too near) portions placed independently of one another in A.

3. On any portion of the path which is small relatively to the total length of path but decidedly larger than the range of possible detection, the observer always detects the target within the lateral range $W/2$ on either side of the path and never beyond.

These assumptions lead to the *formula of random search*

$$p = 1 - e^{-WL/A}, \qquad (40)$$

where A = area in square miles, W = effective search width in miles, L = total length of observer's path in A in miles.

To prove this, suppose that the observer's path L is divided into n equal portions of length L/n. If n is large enough so that most of the pieces are randomly related to any particular one, the chance of failing to detect during the whole path L is the product of the chances that detection fail during motion along each piece. If, further, L/n is such that most of the pieces of this length are practically straight and considerably longer than the range of detection, then in virtue of (3) the latter chance of detection is the probability that the target be in the area swept (whose value is WL/n square miles), and this probability is WL/nA [assumption (1)]. Hence the chance that along all of L there be no detection is $(1 - WL/nA)^n$, and hence

$$p = 1 - \left(1 - \frac{WL}{nA}\right)^n$$
$$= 1 - e^{-WL/A} \text{ for large } n.$$

This reasoning assumes, of course, that a large n having these properties exists. This is essentially assumption (2).

If the exponential in equation (40) is replaced by the first two terms in its power series expansion, the equation is replaced by $p = WL/A$; this corresponds to the probability in the case that L consists of a

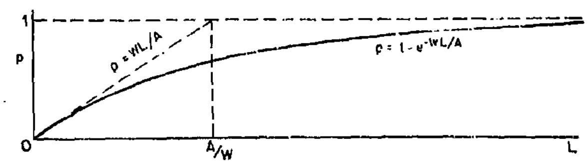


FIGURE 9. Detection with random search.

single straight line, or a path so little bent that there is practically no overlapping of swept regions: The total area swept is WL and the chance of the target being in it is WL/A. The departure from this simple value represents the effect of random overlapping of swept areas.

Figure 9 shows the way in which the probability increases with the length of observer's path L. For smaller values of L it is closely approximated by its tangent $p = WL/A$. For much larger values, it approaches unity, exhibiting a "saturation" or "diminishing returns" effect.

2.7 PARALLEL SWEEPS

Search by parallel sweeps is a method frequently employed, and many apparently more complicated schemes turn out to be equivalent to parallel sweeps, either exactly or with sufficient approximation for practical purposes. A target is at rest on the ocean in an unknown position, all equal areas having the same chance of containing it; it is decided to search along a large ("infinite") number of parallel lines on the ocean, their common distance apart, or *sweep spacing*, being S miles; what is the probability $P(S)$ of detection? Or again, the target's speed and direction are known, but the position is uniformly distributed as above; it is possible to search in equally spaced parallel paths *relative to the target*, i.e., in the plane moving with the target's motion and in which it appears to be a fixed point, as in the first case. It is immaterial whether all the parallel paths are traversed by the same observer or by different observers having similar observing characteristics.

In Figure 10 the parallel paths are shown referred to a system of rectangular coordinates; the axis of ordinates is along one of the paths and the target's (unknown) position is at (x,y), and $0 \leqq x < S$. It is observed that this inequality, expressing the fact that the target is in the strip immediately to the right of the axis of ordinates, is a consequence of the method of choice of the axes, and implies no restriction in the position of the target.

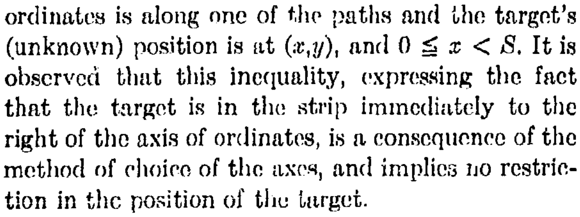


FIGURE 10. Parallel sweeps

The first step in calculating $P(S)$ is to write down the lateral ranges of the target from the various observer paths. For paths at or to the left of the axis of ordinates, the lateral ranges are

$$x, \quad x + S, \quad x + 2S, \quad x + 3S, \quad \cdots$$

For those to the right,

$$S - x, \quad 2S - x, \quad 3S - x, \quad \cdots$$

All these cases may be combined into the absolute value formula:

$$\textit{lateral range} = | x - nS | \tag{41}$$

where

$$0 \leqq x < S \quad \text{and} \quad n = 0, \pm 1, \pm 2, \pm 3, \quad \cdots.$$

Equation (25) is now applied to find the probability $p_n = p$ (nth lateral range) of detection by the nth sweep when the target's position is given as (x,y)

$$p_n = 1 - e^{-F(|x-nS|)},$$

where the potential F is given by the appropriate formula. The probability of no detection by the nth sweep is $1 - p_n$; that of no detection by any sweep is the (infinite) product $\Pi(1 - p_n)$ for all values of n ($< 0, = 0, > 0$); and the probability that at least one sweep detect a target given at (x,y) is

$$P(x,S) = 1 - \prod_{n=-\infty}^{+\infty} e^{-F(|x-nS|)},$$

or, finally,

$$\begin{aligned} P(x,S) &= 1 - e^{-\Phi(x,S)} \\ \Phi(x,S) &= \sum_{n=-\infty}^{+\infty} F(|x - nS|), \quad 0 \leqq x < S. \end{aligned} \tag{42}$$

This is essentially a repetition of the argument proving the additivity of the potentials Section 2.3.

It remains to find the probability of detection $P(S)$ when the target's position (the value of x) is not given, but has a uniform distribution between 0 and S. An easy probability argument shows that $P(S)$ is the average of $P(x, S)$ over all values of x in this interval:

$$\begin{aligned} P(S) &= \frac{1}{S}\int_0^S (1 - e^{-\Phi(x,S)})\, dx \\ \Phi(x,S) &= \sum_{n=-\infty}^{+\infty} F(|x - nS|). \end{aligned} \tag{43}$$

This gives the general solution of the problem.

The *effective visibility* E is defined as half that sweep spacing for which the probability of detection by parallel sweeps is one half. In other words, E is determined as the solution of the equation

$$P(2E) = \frac{1}{2}.$$

Three cases are of particular interest. The first is that of continuous looking on the assumption of a definite range law. Detection will surely occur if, and only if, the target happens to be within the definite range R of either of the two adjacent sweeps. The chance for this is $2R/S = W/S$ when $S > 2R = W$, and unity when $S \leq W$. It is easy to see that the effective visibility $E = W$.

The second case is that of the inverse cube law (which will be taken here in its simplest form). We obtain $\Phi(x,S)$ with the aid of equation (26)

$$\begin{aligned}\Phi(x,S) &= 2m \sum_{n=-\infty}^{+\infty} \frac{1}{(x-nS)^2} \\ &= 2m \frac{\pi^2}{S^2} \csc^2 \frac{\pi x}{S},\end{aligned} \tag{44}$$

the latter equality resulting from a well-known formula of analysis (obtained, e.g., from the expansion of the sine in an infinite product by taking logarithms and then differentiating twice). Inserting this expression into equation (43), we must find

$$P(S) = \frac{1}{S}\int_0^S \left[1 - \exp\left(-\frac{2m\pi^2}{S^2}\csc^2\frac{\pi x}{S}\right)\right]dx.$$

This is found by means of equation (38), on setting $\theta = \pi x/S$ and $\lambda = 2m\pi^2/S^2$. The result, which can be transformed by means of equation (29), is

$$P(S) = \operatorname{erf}\frac{\pi\sqrt{2m}}{S} = \operatorname{erf}\left(\frac{\sqrt{\pi}}{2}\frac{W}{S}\right). \tag{45}$$

We are now in a position to express m and W in terms of the effective visibility E. To find E we solve

$$P(2E) = \operatorname{erf}\frac{\pi\sqrt{2m}}{2E} = \operatorname{erf}\frac{\sqrt{\pi}\,W}{4E} = \frac{1}{2}.$$

The tables of the probability integral show that erf $0.477 = 0.5$; hence

$$\pi\frac{\sqrt{2m}}{2E} = \frac{\sqrt{\pi}\,W}{4E} = 0.477,$$

i.e.,

$$m = 0.0116E^2, \quad W = 1.076E. \tag{46}$$

These values substituted into equation (45) give

$$P(S) = \operatorname{erf}\left(0.954\frac{E}{S}\right). \tag{47}$$

A third case is useful to consider, although strictly speaking it is not one of parallel sweeps but of uniform random search. It may be described as the situation which arises when the searcher attempts to cover the whole area uniformly by a path or paths which place about the same length of track in each strip but which operate within a given strip in the

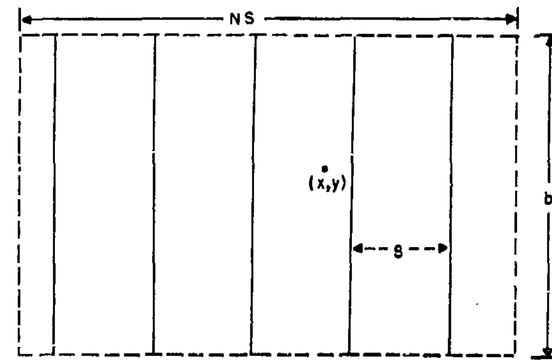


FIGURE 11. A rectangle of random sweeps.

manner of the searcher of Section 2.6. Let all the strips be cut by two horizontal lines a distance of b miles apart and suppose that the search is for a target inside the large rectangle bounded by these lines and two vertical lines NS miles apart, as shown in Figure 11. The area is $A = NSb$ square miles. Assume that the total length of track is equal to that of all included parallel sweeps, $L = Nb$, then apply equation (40); we obtain

$$P(S) = 1 - e^{-W/S}. \tag{48}$$

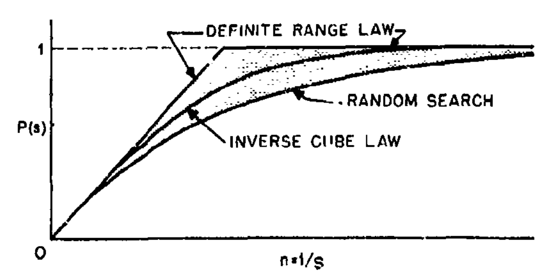


FIGURE 12. Probabilities with parallel sweeps.

It is independent of N and b.

These three cases may be represented by means of a common diagram (Figure 12) by plotting

$P = P(1/n)$ where $n = 1/S$ is the *sweep density*, or number of sweeps per mile. At one extreme is the case of the definite range law, at the other the case of random search. All actual situations can be regarded as leading to intermediate curves, i.e., lying in the shaded region. The inverse cube law is close to a middle case, a circumstance which indicates its frequent empirical use, even in cases where the special assumptions upon which its derivation was based are largely rejected.

2.8 FORESTALLING

When the observer is using two different means of detection simultaneously and independently (i.e., when neither interferes with or aids the other), it is sometimes necessary to know the probability of making a first detection by a particular one of the two means. Since the second means of detection can deprive the first of a chance of detecting (by detecting the target first) this probability may be lower than if the second means had not been present: We say that the second can *forestall* the first. For example, when both radar and visual detection are possible, in gathering data bearing on the effectiveness of the radar, the possibility of visual forestalling of the radar must be taken into account.

Again, when the target is itself capable of detecting the observer, and if it is important to detect the target before it can detect the observer (as when the former is a surfaced submarine which can submerge if it detects the observer first and so deprive it of its chance of detecting), it is important to find the probability that the observer detect the target first, before it has been forestalled by the target's detection.

Just as the chance of detection is mathematically equivalent to that of hitting a target continuously exposed to our fire (intensity varying in general with the time), so the question of forestalling is mathematically identical with that of hitting the target before it hits us, in the case where it is an enemy continuously firing back.

It will be sufficient to consider the case of continuous looking with the instantaneous probability $\gamma_t dt$ for the first means of detection without forestalling (Section 2.2), the probability $p(t)$ of detection when there is no forestalling being given by equation (6). Let $\gamma_t' dt$ and $p'(t)$ be the corresponding quantities for the second means of detection, or for the target's detection of the observer in the second example above.

If $P(t)$ is the required probability of first detection during the interval of time from 0 to t by the first means, we consider the value of $P(t + dt)$. It is the probability of an event which can succeed in either of the following mutually exclusive ways: either by having the required detection between 0 and t, or by having neither means detect during this period, and having a detection by the first means between t and $t + dt$. This leads to the equation

$$P(t + dt) = P(t) + [1 - p(t)]\,[1 - p'(t)]\gamma_t\, dt,$$

whence a differential equation is obtained, the solution of which is

$$P(t) = \int_0^t \gamma_t \exp\left[-\int_0^t (\gamma_t + \gamma_t')dt\right]dt. \quad (49)$$

Precisely the same reasoning leads to the expression

$$P'(t) = \int_0^t \gamma_t' \exp\left[-\int_0^t (\gamma_t + \gamma_t')dt\right]dt$$

for the probability of a first detection by the second means in the time interval $0,t$.

Note that the sum $P(t) + P'(t)$ is the probability of a first detection either by the first means or by the second, in other words, the probability of a detection by *some* means between 0 and t. The expression obtained by adding the above equations and carrying out one integration is

$$P(t) + P'(t) = 1 - \exp\left[-\int_0^t (\gamma_t + \gamma_t')dt\right],$$

which is simply the expression (6) with γ_t replaced by $\gamma_t + \gamma_t'$, the latter being the instantaneous probability when both means of detection act in conjunction (additivity of potentials).

When a large number of independent trials of the detection experiment are made under identical conditions and all cases which have resulted in a first detection by the first means are sorted out and the precise epochs of this detection are averaged, the result will (statistically) be equal to

$$\bar{t} = \frac{\int_0^\infty t\gamma_t \exp\left[-\int_0^t (\gamma_t + \gamma_t')dt\right]dt}{\int_0^\infty \gamma_t \exp\left[-\int_0^t (\gamma_t + \gamma_t')dt\right]dt}. \quad (50)$$

The denominator is proportional to the total number of first detections by first means, and the result of dividing $\gamma_t \exp\left[-\int_0^t (\gamma_t + \gamma_t')dt\right] dt$ by the denominator is the proportion of such detections between t and $t + dt$. Thus the expression in equation (50) represents the expected value $\bar{t}$ of t, the epoch of detection by the first means.

As a first application we consider the case of constant instantaneous probabilities of detection, $\gamma_t = \gamma$, $\gamma_t' = \gamma'$. Equations (48) and (49) reduce to

$$P(t) = \frac{\gamma}{\gamma + \gamma'}\left[1 - e^{-(\gamma + \gamma')t}\right],$$
$$\bar{t} = \frac{1}{\gamma + \gamma'}.$$

It is thus seen that the proportion of the total number of first contacts by the first means (as $t \to \infty$) is $\gamma/(\gamma + \gamma')$, and, correspondingly, by the second, $\gamma'/(\gamma + \gamma')$. And the mean time elapsed to the former is the same as for the latter, i.e., $1/(\gamma + \gamma')$; this is different from the mean time $1/\gamma$ when no forestalling had been possible.

As a second application we consider the straight track case of Section 2.4, and assume that the observer is an aircraft and the target a surfaced submarine. If the observer sights the wake of the submarine, his ability to detect may reasonably be taken as the inverse cube law of equation (10), and if the submarine sights the horizontal surfaces of the aircraft's wing, the same law (with k replaced by a different constant k') can reasonably be assumed for the submarine's detection of the aircraft. If it is assumed that the submarine dives as soon as it detects the aircraft, what is the probability that the aircraft detect the submarine, as a function of lateral range x? By how much is its effective search width decreased by this new possibility?

Equation (10) under the circumstances of Section 2.4 leads to

$$\gamma_t = \frac{kh}{(x^2 + w^2t^2)^{\frac{3}{2}}}, \quad \gamma_t' = \frac{k'h}{(x^2 + w^2t^2)^{\frac{3}{2}}};$$

only in the present case the time interval is from $-\infty$ to t instead of from 0 to t. With these changes equation (49) leads to the following expression for the probability of sighting the submarine before the time t:

$$P(x,t) = m\int_{-\infty}^{wt} \exp\left[-\frac{m + m'}{x^2}\left(1 + \frac{y}{\sqrt{x^2 + y^2}}\right)\right]\frac{dy}{(x^2 + y^2)^{\frac{3}{2}}}$$

where $m = kh/w$ and $m' = k'h/w$. The integral can be evaluated explicitly when it is noted that the integrand is proportional to the derivative of the exponential expression, i.e.,

$$\frac{1}{(x^2 + y^2)^{\frac{3}{2}}} \exp\left[-\frac{m + m'}{x^2}\left(1 + \frac{y}{\sqrt{x^2 + y^2}}\right)\right]$$
$$= -\frac{1}{m + m'}\frac{d}{dy}\exp\left[-\frac{m + m'}{x^2}\left(1 + \frac{y}{\sqrt{x^2 + y^2}}\right)\right].$$

The result, on setting $t = +\infty$, gives the following probability of sighting the submarine some time on its whole straight course.

$$P(x,\infty) = \frac{m}{m + m'}\left[1 - e^{-2(m + m')/x^2}\right].$$

To find the value of the search width (which will be denoted by W^*), this expression must be used in the place of (26) in equation (27). The answer is obtained from (29) by replacing m by $m + m'$ and then multiplying the result by $m/(m + m')$; it is

$$W^* = 2\sqrt{2\pi m}\sqrt{\frac{m}{m + m'}} = W\sqrt{\frac{m}{m + m'}}.$$

Thus the effect of forestalling is to multiply W by a factor less than unity of $\sqrt{m/(m + m')}$. And the probability of detection even when the target is flown over ($x = 0$) is $P(0,\infty) = m/(m + m')$ instead of unity, as it would have been in the absence of forestalling.

For a definite range law, that means of detection which has the greater range will always forestall the other. (Of course this is strictly true only in the case of *continuous looking*.)

2.9 CONCLUSION—OPERATIONAL DISTRIBUTIONS

Returning to first principles, as set forth in Sections 2.1 and 2.2, it has been laid down as basic that detection, even when possible, is an *uncertain* event; and the whole subsequent course of development of this chapter has been toward the calculation of *probabilities* of detection. But an essential restriction has been imposed in all these calculations: The one source of uncertainty which has been considered is the human fallibility of the observer, and the sudden uncontrollable fluctuations in the physical state of affairs, but not in the random element introduced by unknown, long-term variations in the underlying physical conditions (conditions which are expressible as parameters). Thus, as we have said in Section 2.1(ii), under given meteorological conditions of visibility V the observer will have a definite chance $\gamma(r)dt$ of sighting a target of given size A and background contrast C; and subsequent deductions have been made on the assumption that while the range r may vary in a given manner in the course of time, the parameters V, A, and C *all remain fixed*. The distributions calculated on this assumption can be expected to agree with the distributions found empirically when the results of a large number of experiments are obtained, all of which are performed under the same conditions of visibility and size and contrast of the target, geometrical quantities like r alone being allowed to vary. But as soon as operational results are compiled which refer to cases in which V, A, and C vary from incident to incident, an altogether different situation is present: The cause of the uncertainty of the event of detection is twofold, being dependent not only on the human fallibility of the observer and short-term fluctuations, but on the more or less unknown and heterogeneous nature of the underlying physical conditions. And it is important to realize that in many cases this second factor may outweigh the first. When this is judged to be the case, it may well be expedient to employ a highly simplified law of detection, such as the definite range law, and then seek to explain the distributions found in the operational data simply by averaging the calculated results of such laws over different possible values of the parameters. Thus if the definite range law is assumed, mathematical expressions deduced from it will involve this range R; then it may be considered that in the operational incidents different values of R are present; by choosing appropriate frequencies for the different values of R and combining or averaging the theoretical results over such distributions of R, a good agreement may often be found with the observations.

It must be emphasized that equations such as (1), (3), (5), and (6) are true only when the first cause of uncertainty alone is present, and when the underlying physical conditions remain constant (and are known to be of constant, though not necessarily of known values) throughout the course of the looking. Thus in proving (1), the probability of detection for one glimpse was g, of not detecting, $1 - g$; now precisely at the point where it was asserted that the probability of failure to detect at each and every one of the first n glimpses is $(1 - g)^n$, the assumption that the n different events are independent was made. This is justified only in two cases: first, when the only uncertainty is in the observer's chance performance so that his different opportunities (glimpses) are regarded as repeated independent trials (as in successive tosses of a coin); second, when there are indeed changes in physical conditions, but of such a rapidly fluctuating character that if no detection is known to occur at one glimpse, no inference can be drawn regarding the physical conditions pertaining to any other glimpse. But if, for example, the visibility V is not fully known, the fact that earlier glimpses have failed to detect may lead to the presumption that V is less than might otherwise have been supposed, and hence that later chances of detection are less: the expression $(1 - g)^n$ is false.

The method of procedure is clear. The first step is to carry out the calculations as described in the previous sections of this chapter, assuming fixed conditions (such as V,A,C). The second step is to average the results obtained for the probabilities (e.g., over the possible values of V,A,C, with appropriate weighting). Only the final result can reasonably be expected to furnish the probabilities which accord with the operational data. What is true of probabilities is also true of mean or expected values defined by them.

This will be illustrated by many practical examples, particularly in Chapters 4, 5, and 6. But three simple cases can be mentioned here. Firstly, suppose that the lateral range curve (Section 2.4) involves a parameter λ referring to an unknown factor in the underlying physical conditions. Its equation is $p = p(x,\lambda)$. Once the distribution of frequencies with which the different values of λ occur in an operational situation has been estimated, the operational lateral range curve $p = p_{op}(x)$ (i.e., the one furnished by a histogram of the observed data) is found by averaging $p(x,\lambda)$ over the values of λ on the basis of this frequency. Thus it might be reasonable in some cases to assume that the values of λ are normally distributed about a known mean l with a known standard deviation σ. Accordingly,

$$p_{op}(x) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} p(x,\lambda)\, e^{-(\lambda-l)^2/2\sigma^2}\, d\lambda.$$

Thus if $p(x,\lambda)$ results from a definite range law of range $R = \lambda$, so that

$$p(x, \lambda) = 1, \text{ when } x < \lambda,$$

$$p(x, \lambda) = 0, \text{ when } x > \lambda,$$

the equation becomes

$$p_{op}(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_x^{+\infty} e^{-(\lambda - l)^2/2\sigma^2}\, d\lambda$$
$$= \frac{1}{2}\left(1 - \operatorname{erf} \frac{x - l}{\sigma\sqrt{2}}\right),$$

the graph of which is shown in Figure 13.

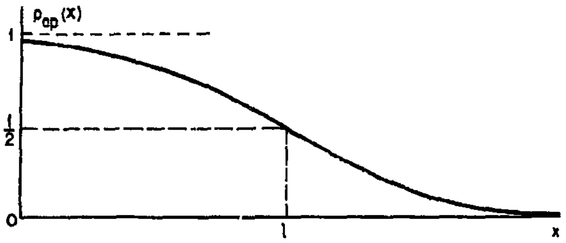


FIGURE 13. Lateral range curve based on a normal distribution of definite ranges.

It is noted that when $x = 0$, $p_{op}(0)$ is slightly less than unity, whereas it should exactly equal unity. This is because the normal distribution of definite ranges allows a (slight) chance of negative ranges, a physical absurdity. It would have been more realistic to have assumed a skew nonnegative distribution (e.g., Pearson's Type III distribution, $k\lambda^{m-1} e^{-\beta\lambda}$).

As a second example, consider the search for a fixed target by two parallel sweeps at distance S apart. If the underlying conditions are the same during the two sweeps, and if $p(x)$ is the lateral range probability, the chance of detection of a target between the paths and x miles from one of them is shown by the usual reasoning to be

$$P(x,S) = 1 - [1 - p(x)]\,[1 - p(S - x)]$$
$$= p(x) + p(S - x) - p(x)p(S - x).$$

If $p(x) = p(x,\lambda)$, a weighted averaging process must be performed in order to get the operational probability $P_{op}(x,S)$ from $P(x,S,\lambda)$ given by the above equation. And of course if x is determined at random between 0 and S, a second averaging must be done to get $P_{op}(S)$, the operational probability of detecting the target given only to be somewhere between the sweeps and with a given distribution of physical conditions. The order in which these two averagings are done is immaterial. A corresponding treatment is given in the case of infinitely many parallel sweeps. It may be remarked that the operational effective visibility E_{op}, which is defined by the equation (see Section 2.7):

$$P_{op}(2E_{op}) = \int P(2E_{op},\lambda) f(\lambda) d\lambda = \frac{1}{2}$$

is quite different from the average $\bar{E}$:

$$\bar{E} = \int E(\lambda) f(\lambda) d\lambda,$$

of the effective visibility defined under fixed conditions corresponding to a particular value of λ. Here $f(\lambda)$ is the assumed frequency with which the values of λ are taken to be distributed under the operational conditions in question.

As a third example, suppose that a radar set is chosen at random from a lot, only the fraction ϵ of which are in good adjustment, the remaining $1 - \epsilon$ not in a condition to make any detections possible. The radar lateral range curve $p(x)$ for a radar set in good adjustment and, e.g., mounted on an aircraft, must be multiplied by ϵ to obtain the operational curve that will be obtained when many observations are made with the aid of many sets chosen in this way. When a set or similar observing instrumentality or setup is not giving the results which could be expected of it, it is often said to be working at an *efficiency* less than 100 per cent. In the above case, a natural definition of efficiency is 100ϵ. In more complicated cases, the concept, while useful as a general concept, may not be convenient to define in all precision.

In conclusion, the following principle is laid down:

If the object of the calculation of probabilities, averages, and similar statistical detection quantities is to coordinate and explain the data of the operations of the past, then the heterogeneity of conditions (dispersion of slowly varying parameters) is placed at the apex of the discussion, the influence of "subjective" probabilities and short-term fluctuations (the main subject of this chapter) usually playing a secondary role.

If on the other hand the object of the calculation is to obtain contemplated performance data for the design of search plans to be used in the future, and, as is generally the case, when the conditions (slowly varying parameters) are known, then the probabilities originating from subjective and rapidly fluctuating sources occupy the center of the stage; any study of the heterogeneity of conditions is made only in order to check the sensitivity of the search plan to accidental imperfections in the knowledge of the conditions.

All this will be made clear on the basis of examples in the succeeding chapters.

Chapter 3

THE DISTRIBUTION OF SEARCHING EFFORT

3.1 THE GENERAL QUESTION

In an important class of problems of naval search, the target has an unknown position but a known distribution: while we do not know where it is, we do know the probability that it is in this place rather than in that. If, as is usual, the total available searching effort (number of hours the observer can devote to the search) is limited, how should the searching be done? How much time should the observer spend searching in this place and how much in that?

In other cases, there is no doubt as to *where* to search but there may be a question as to *when*. If the target is only temporarily present in the region where the search can be made, all hours of the twenty-four being equally likely, should the limited searching effort be spread out evenly during the twenty-four hours, or should a more intensive search be conducted during only part of the time? This question is of particular interest if the search is done by aircraft, when only a limited number of aircraft-hours are available, and if search in daylight when vision and radar can be used simultaneously is more effective than search at night when only radar is available. How should the search be divided between day and night?

There are similar questions concerning the optimum distribution of scanning effort: An observer on an aircraft is searching for a surface target; how much of his time should be devoted to looking straight ahead, how much in looking abeam, and how much on the intermediate bearings?

There is a close mathematical analogy between some of these problems and certain questions of gunnery and bombing in which the distribution of targets is known, and the optimum distribution of firing is required.

It has been seen in Chapter 2 that when a region is searched the chance of detection depends not only on the law of detection (lateral range curve, search width, etc.), but on the method of search. At one extreme, such a highly systematic method as that of parallel sweeps can be used; at the other extreme, random search can be employed, leading to the equation (40) of Chapter 2,

$$p = 1 - e^{-WL/A}, \qquad (1)$$

where

A = area searched (square miles),
W = effective search width (miles),
L = length of observer's path inside A (miles),
p = probability of detection of a target given to be in A and uniformly distributed therein.

Throughout this chapter, the problems just mentioned will be treated on the basis of equation (1). The reason for the assumption of random search is twofold. On the one hand it is realistic, since in any protracted search, however systematic in intent, navigational errors and other irregularities and uncertainties are pretty sure to impart to the search a character of random; hence (1) is a proper estimate, on the conservative side, of the practical results achieved. On the other hand, the assumption is convenient and leads to usefully simple results; this is partly because equation (1) requires nothing concerning the particular detection law (other than the value of W) to be assumed, and partly because of the usable nature of (1) itself.

3.2 ALTERNATIVE REGIONS OF SEARCH

Let A_1 and A_2 be two areas of the ocean (either separate or having a common boundary). The target to be found is either in A_1 or in A_2, with the respective probabilities p_1 and p_2 of so being, so that relations

$$p_1 + p_2 = 1, \quad p_1 > 0, \quad p_2 > 0$$

hold; the target being stationary, p_1 and p_2 do not change with the time. Let the target be uniformly distributed in whichever of A_1 or A_2 it lies. Finally, let the total length of track of the observer (or observers) be L. How must L be distributed between A_1 and A_2 if the chance of detection is to be greatest? In other words, if $L = L_1 + L_2$, L_1 being the length of the observer's track in A_1, L_2 that in A_2, what relation must exist between L_1 and L_2 for the optimum search? If W is the search width, the probability p of detection is given by

$$p = p_1 (1 - e^{-WL_1/A_1}) + p_2 (1 - e^{-WL_2/A_2}), \quad (2)$$

as results from a simple probability argument based on (1).

Mathematically, the problem is to find the values of L_1 and L_2 which maximize equation (2), subject to the conditions

$$L_1 + L_2 = L, \quad L_1 \geqq 0, \quad L_2 \geqq 0. \tag{3}$$

It is convenient to proceed graphically. Setting

$$L_1 = x, \quad L_2 = L - x,$$
$$y_1 = p_1 e^{-Wx/A_1}, \quad y_2 = p_2 e^{-W(L-x)/A_2},$$

we have $$p = 1 - (y_1 + y_2),$$

so that, for the optimum search, x must be determined so as to *maximize* p, i.e., to *minimize*

$$y = y_1 + y_2,$$

subject to the restriction that $0 \leqq x \leqq L$. Figure 1 shows the graph of y against x in a typical case; the ordinate is obtained by adding the ordinates of the

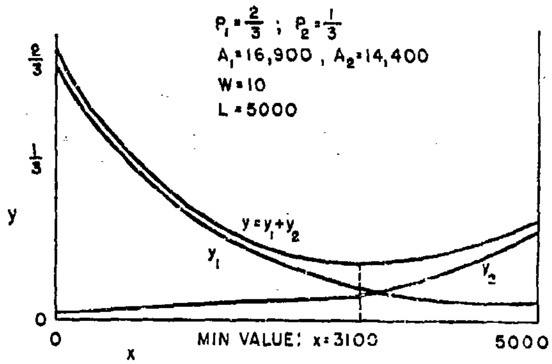


FIGURE 1. Minimizing the sum.

graphs of y_1 and y_2, also shown in the figure; the latter are simple exponential curves. It is seen, either by differentiation or by an obvious graphical argument, that for y to have a minimum at a point x in the interval $(0, L)$ it is necessary and sufficient that the inclinations (viz., absolute values of the slopes) of the tangents to the y_1 and the y_2 curves be equal and opposite (they are always opposite in the present case). Now the inclination of the y_1 is always down and that of the y_2 always up. The former inclination decreases with increasing x from its maximum at $x = 0$, the latter increases to its maximum at $x = L$. Thus there are three mutually exclusive possibilities.

Case 1. The inclination of y_1 at $x = 0$ is less than (or at most equal to) that of y_2 at this point, i.e.,

$$\frac{p_1}{A_1} \leqq \frac{p_2}{A_2} e^{-WL/A_2}. \tag{4}$$

Then the inclination of y_1 will continue to be still less than that of y_2 throughout the interval: No internal minimum of y exists, but since y_2 increases faster than y_1 decreases as x moves from 0 to L, *the minimum occurs at* $x = 0$ (i.e., $L_1 = 0$).

Case 2. The inclination of y_2 at $x = L$ is less than (or equal to) that of y_1 at this point, i.e.,

$$\frac{p_2}{A_2} \leqq \frac{p_1}{A_1} e^{-WL/A_1}. \tag{5}$$

Then, by a similar argument, *the minimum occurs at* $x = L$.

Case 3. Neither one of the above cases occurs:

$$\frac{p_1}{A_1} > \frac{p_2}{A_2} e^{-WL/A_2} \quad \text{and} \quad \frac{p_2}{A_2} > \frac{p_1}{A_1} e^{-WL/A_1}. \tag{6}$$

Then the inclination of y_1 at $x = 0$ is greater than that of y_2 at this point, but this advantage is steadily diminished as x increases, and is reversed when $x = L$, at which point the inclination of y_2 exceeds that of y_1. Hence *there is just one point* $x = x_0$ *in the interval* $(0, L)$ *which makes* y *a minimum.* This value of x is found by solving the equation $dy/dx = 0$; a convenient form of the answer will be given below.

In order to grasp the meaning of the situation more easily, the following terms are introduced:

$$\begin{aligned} \rho_1 &= p_1/A_1, & \rho_2 &= p_2/A_2 \\ \phi_1 &= WL_1/A_1, & \phi_2 &= WL_2/A_2 \\ A &= A_1 + A_2, & \Phi &= WL. \end{aligned} \tag{7}$$

On account of the uniformity of the distribution of the target in whichever region it is, the probability that it lie in a subregion of unit area within A_1 is the product of the probability p_1 that it be in A_1 by the probability that (in this case) it be in said subregion, $1/A_1$: Thus the probability in question is $p_1/A_1 = \rho_1$. Hence ρ_1 is the *probability density* for the first region; ρ_2 has the corresponding meaning for A_2. Furthermore, the length of observer's path L_1 in A_1 is a measure of the searching effort devoted to A_1. But WL_1 is an equally good measure: it is the area swept (some of it multiply) in A_1. Thus the expression $WL_1/A_1 = \phi_1$ is the *density of searching effort* in the first region, and ϕ_2 is that in the second. Φ is the total available searching effort, and we have

$$A_1\phi_1 + A_2\phi_2 = \Phi, \quad \phi_1 \geqq 0, \quad \phi_2 \geqq 0. \tag{8}$$

In the third case, corresponding with equation (6), the solution of $dy/dx = 0$ for x gives a result which can be simplified by first replacing x and $L - x$ by L_1 and L_2, and then replacing ratios by the quantities introduced in equation (7). The following equations result from this process; they answer the question concerning the optimum distribution of searching effort in the present case.

$$\phi_1 = \log \rho_1 - \frac{1}{A}(A_1 \log \rho_1 + A_2 \log \rho_2) + \frac{\Phi}{A}$$
$$\phi_2 = \log \rho_2 - \frac{1}{A}(A_1 \log \rho_1 + A_2 \log \rho_2) + \frac{\Phi}{A}. \qquad (9)$$

On the basis of equations (4), (5), and (6), expressed in the terms of equations (7), and (9), everything may be summed up as follows:

When the target's probability density ρ_1 in the first region is not only less than its density in the second, but is so much less that it remains less when the second is multiplied by the factor $e^{-\Phi/A_2}$ (< 1), i.e.,

$$\rho_1 \leqq \rho_2 e^{-\Phi/A_2}, \qquad (10)$$

then no searching whatsoever should be done in the first region, A_1, and the whole effort should be devoted to searching the second, A_2. Similarly, if

$$\rho_2 \leqq \rho_1 e^{-\Phi/A_1}, \qquad (11)$$

no searching should be done in A_2. When, on the other hand,

$$\rho_1 > \rho_2 e^{-\Phi/A_2} \quad \text{and} \quad \rho_2 > \rho_1 e^{-\Phi/A_1}, \qquad (12)$$

the searching effort should be distributed in accordance with equation (9). If k denotes the exponential of the common value added to $\log \rho_1$ or to $\log \rho_2$ in (9), these equations become

$$\phi_1 = \log k\rho_1, \quad \phi_2 = \log k\rho_2. \qquad (13)$$

Thus *the optimum densities of searching effort* [in case of equation (12)] *equal the logarithms of quantities proportional to the respective probability densities.*

It is interesting to see how the situation develops as the total available searching effort Φ is progressively increased. To have a definite case, suppose that $\rho_1 > \rho_2$ and $A_1 < A_2$. When Φ is very small, equation (11) holds: all searching must be done in A_1, none in A_2. When Φ increases sufficiently, (12) becomes valid, and remains so for all further increase in Φ. Then the searching has to be distributed between A_1 and A_2 according to the logarithmic law enunciated in equations (9) or (13). This leads to a curious conclusion as Φ becomes extremely large. For equation (9) shows that

$$\lim_{\Phi \to \infty} \frac{\phi_1}{\phi_2} = 1.$$

Thus for very large Φ, ϕ_1 is about equal to ϕ_2, i.e., $WL_1/A_1 = WL_2/A_2$, which means in the present case where $A_1 < A_2$ that $L_1 < L_2$. The first region, which for very small Φ should take up all the searching effort, should for very large Φ actually have a shorter length of observer's track than the second region, a phenomenon of *reversal for large* Φ.

Suppose that after completing the optimum search with the expenditure of the total effort $\Phi = WL$, with no resulting detection, a further amount of effort $\Phi' = WL'$ becomes available. What is the optimum manner of expending this additional effort?

Assume that in the first part of the search, the third case was presented, so that Φ was distributed in accordance with equation (9). Since, as we are assuming, the target is not detected, the situation at the end of the search is similar to that at the beginning, except that the probabilities p_1 and p_2 have to be replaced by different values p_1' and p_2', the values of which are computed by means of Bayes' theorem (see Section 1.5, in particular footnote c). In this application, the a priori probabilities of A_1 and A_2 containing the target are p_1 and p_2, the a posteriori probabilities are p_1' and p_2'. The [illegible]ductive probabilities," i.e., those of not find[illegible] target in A_1 or in A_2 when the Φ search is done as assumed in (9) are the values which accrue to the quantities e^{-WL_1/A_1} and e^{-WL_2/A_2} when L_1 and L_2 are given by (9) in conjunction with (7). Thus the first productive probability is, by definition, the chance that if the target is actually in A_1 the searching effort $\phi_1 = WL_1/A_1$ devoted to this region shall fail to reveal it; by the formula of random search, this is e^{-WL_1/A_1}, and similarly for the second region. Hence by Bayes' formula,

$$p_i' = \frac{p_i e^{-WL_i/A_i}}{p_1 e^{-WL_1/A_1} + p_2 e^{-WL_2/A_2}}, \quad i = 1 \text{ or } 2;$$

or, using equations analogous to (7),

$$\rho_i' = \frac{\rho_i e^{-\phi_i}}{A_1 \rho_1 e^{-\phi_1} + A_2 \rho_2 e^{-\phi_2}}, \quad i = 1 \text{ or } 2.$$

To find the optimum distribution of Φ', observe first that on account of (9), the relation $\rho_1 e^{-\phi_1} = \rho_2 e^{-\phi_2}$ holds; hence we are in the presence of the third case [the primed analogue of (12)]. Hence ϕ_1' and ϕ_2' are given by equations like (9) (with appropriate primes). An obvious algebraic simplification, using the original (9), shows that

$$\phi_1' = \phi_2' = \frac{\Phi'}{A}.$$

Now the total density of searching effort devoted to the ith region (i = 1 or 2) is simply $\phi_i + \phi_i'$. This reduces with the aid of (9) to

$$\log \rho_i - \frac{1}{A}(A_1 \log \rho_1 + A_2 \log \rho_2) + \frac{\Phi + \Phi'}{A}.$$

But this is exactly what (9) would have given if we had known in advance that the total amount of searching effort would be $\Phi + \Phi'$ rather than Φ. In other words:

A well-planned search can not be improved by a redistribution of search made at an intermediate stage of the operation in an attempt to make use of the fact that up to that time the target had not yet been observed.

Of course as soon as the target *is* observed, an improvement can be made: Discontinue the search.

This theorem has been proved here only in the case where equation (12) is valid. Other cases are treated in a similar manner, with similar results.

In the case of equation (12), formula (9) gives only the magnitude of effort to be devoted to A_1 and to A_2; it does not tell how the search should be conducted *in time*. Suppose that at the end of t hours the amount of search effort is $\Phi(t) = ct$, where $\Phi = \Phi(T) = cT$, T being the total time available and c a constant of proportionality. Then in order to find the target as soon as possible, we must proceed as follows: Whatever the value of t (> 0), the search effort ct must be used so as to maximize the chance of detection up to that time, in accordance with the formulas developed above. Thus, if $\rho_1 > \rho_2$, we must make $\phi_1 = ct/A_1$, $\phi_2 = 0$, as t goes from zero to $A_2(\log \rho_2 - \log \rho_1)/C$, i.e., when (12) comes into effect. From then on, ϕ_1 and ϕ_2 must be taken from equation (9) in which Φ is replaced by ct.

An obvious extension of the problem treated in this section is to the case of n regions $A_1, \cdots, A_n$. But, while this presents no difficulty, it is much more worth while to treat a perfectly general case which will yield the n region case by specialization. This will be the object of the succeeding section.

3.3 THE GENERAL CASE—TARGETS CONTINUOUSLY DISTRIBUTED

The target is stationary and is contained in a known region A of the ocean (assumed to be a plane in which a cartesian system of coordinates (x,y) is established). The probability before the search that the target be in the infinitesimal region $dxdy$ is $p(x,y)dxdy$, the probability density $p(x,y)$ being known and satisfying the conditions of continuity, of having in A the positive minimum.

$$\min_A p(x,y) = p_0 > 0,$$

and of satisfying the obvious equation.

$$\iint_A p(x,y)dxdy = 1.$$

The total available searching effort as represented either by the total observer's track L or by $\Phi = WL$ is given. How shall the searching be distributed throughout A in order that the chance of finding the target be a maximum?

The idea of a distribution of searching effort in the present case of a continuously varying distribution involves a less precise conception than in the case of Section 3.2. It is necessary to arrive at the notion of density of searching effort $\phi(x,y)$ as a function of position (x,y) in A. This is accomplished as follows:

Consider a subregion B within A. Let L_B be the total length of observer's track in B (composed, perhaps, of many pieces). The quantity $\phi_B = WL_B/B$ obviously corresponds to B in the same way in which ϕ_1 of (7) does to A_1 in Section 3.2. But consider what happens as B shrinks up to fixed point (x,y) contained therein. If B is large and of not too irregular a shape [if it is square or circular with (x,y) at its center], the ratio L_B/B behaves as the ratio of the number of molecules of an inhomogeneous body in a volume to the volume itself as it shrinks: After changing slowly, it settles down to a quasilimiting value, remains at this value for a long time, only to depart radically from it as the area (or volume) falls below a critically small size. The quasilimit $l(x,y)$ of L_B/B (corresponding to the "mean number of molecules per unit volume," proportional to the density of the body) may be called the density of observer track at (x,y). The product $\phi(x,y) = Wl(x,y) =$ quasilimit of WL_B/B shall be called the density of searching effort or *search density* at (x,y). An obvious construction shows that

$$\iint_A l(x,y)dxdy = L,$$

and consequently that

$$\iint_A \phi(x,y)dxdy = WL = \Phi. \tag{14}$$

Suppose that the search density $\phi(x,y)$ is a given function. What is the probability $P[\phi]$ of detecting the target? The classical reasoning of the integral calculus (subdivision of the region A; approximation to $P[\phi]$ by a sum obtained by total and compound probability; and application of the formula (1) in the limit) furnishes the following answer:

$$P[\phi] = \iint_A p(x,y)(1 - e^{-\phi(x,y)})dxdy. \tag{15}$$

We are therefore in the presence of a problem of the calculus of variations: *Among all functions* $\phi(x,y)$ *which satisfy* (14) *together with*

$$\phi(x,y) \geqq 0, \tag{16}$$

find that one which gives to $P[\phi]$, *defined by* (15), *the greatest value.*

This is not a "regular" problem of the calculus of variations, because of the one-sided condition (16), i.e., because this condition is an inequality rather than an equality. And indeed the example of Section 3.2 prepares us for the possibility that under certain conditions the maximum cannot be found by simply equating a combination of differentials (variations) to zero.

As a preliminary step, we shall solve the regular problem [illegible]lar to the above but with the condition (16) omitted. The use of the Lagrange multiplier λ in the variation of (14) and (15) leads to

$$\delta \iint_A p(x,y)(1 - e^{-\phi(x,y)})dxdy - \lambda\delta \iint_A \phi(x,y)dxdy = 0,$$

$$\iint_A [p(x,y)\, e^{-\phi(x,y)} - \lambda]\, \delta\phi(x,y)dxdy = 0,$$

from which is derived

$$p(x,y)\, e^{-\phi(x,y)} = \lambda,$$
$$\phi(x,y) = \log p(x,y) - \log \lambda. \tag{17}$$

To determine λ, introduce this in (14) and solve for $\log \lambda$;

$$\log \lambda = \frac{1}{A}\iint_A \log p(x,y)dxdy - \frac{\Phi}{A}; \tag{18}$$

whereupon the following solution of the problem is obtained from (17):

$$\phi(x,y) = \log p(x,y) - \frac{1}{A}\iint_A \log p(x,y)dxdy + \frac{\Phi}{A}. \tag{19}$$

Returning to the original problem with (16) in force, it is seen that equation (19) automatically gives the solution whenever Φ is large enough to make the right-hand member of (19) nonnegative for all (x,y) of A, i.e., whenever (p_0 being the minimum of $p(x,y)$ in A)

$$\log p_0 - \frac{1}{A}\iint_A \log p(x,y)dxdy + \frac{\Phi}{A} \geqq 0. \tag{20}$$

In case this inequality is not realized, we shall allow the previous formulas to suggest a $\phi(x,y)$ which is written down ad hoc, and then show a posteriori that it is in effect the solution of the problem.

Introduce the variable q ranging in the interval

$$p_0 \leqq q \leqq p_1 = \max_A p(x,y),$$

and denote by A_q and $\bar{A}_q$ the two parts of A (i.e., $A = A_q + \bar{A}_q$) defined as follows:

$$\begin{aligned} p(x,y) &\geqq q \quad \text{for all } (x,y) \text{ in } A_q, \\ p(x,y) &< q \quad \text{for all } (x,y) \text{ in } \bar{A}_q. \end{aligned} \tag{21}$$

Consider the quantity

$$\iint_{A_q} [\log p(x,y) - \log q]dxdy$$
$$= -A_q \log q + \iint_{A_q} \log p(x,y)dxdy.$$

Geometrically, it represents the volume of that part of the solid under the surface $z = \log p(x,y)$ [plotted in the cartesian coordinates (x,y,z)] cut off by the plane $z = \log q$ (and above this plane). As q *decreases* continuously from p_1 to p_0, this volume *increases* continuously from 0 to

$$-A \log p_0 + \iint_A \log p(x,y)dxdy.$$

Hence the expression

$$A_q \log q - \iint_{A_q} \log p(x,y)dxdy + \Phi$$

decreases continuously from Φ to the value

$$A \log p_0 - \iint_A \log p(x,y)dxdy + \Phi$$

which, on account of the assumed invalidity of (20), is negative. There exists, therefore, a unique value $q = b$ for which the expression is zero, i.e., for which

$$\log b - \frac{1}{A_b}\iint_{A_b} \log p(x,y)dxdy + \frac{\Phi}{A_b} = 0. \quad (22)$$

We now define $\phi(x,y)$ as follows:

$$\phi(x,y) = \log p(x,y) - \frac{1}{A_b}\iint \log p(x,y)dxdy + \frac{\Phi}{A_b}$$

$$= \log \frac{p(x,y)}{b} \quad (23)$$

when (x,y) is in A_b;

$$\phi(x,y) = 0$$

when (x,y) is in $\bar{A}_b$.

To show that this $\phi(x,y)$ is the solution of the problem, it is observed, firstly, that it satisfies (16) in virtue of (22) and (21); and, secondly, that (by direct integration) it satisfies (14). It remains to show that of all the functions satisfying these two conditions, it renders $P[\phi]$ a maximum, for which we will regard it as sufficient to show that every arbitrarily small change of $\phi(x,y)$ through values satisfying (16) and (14) decreases $P[\phi]$—or leaves it stationary.

First, consider variations which leave the function zero in $\bar{A}_b$, i.e., which correspond to a rearrangement of values inside $\bar{A}_b$. To find the maximum, repeat the preceding calculation starting from the equation $\delta P - \lambda\delta\Phi = 0$ with A_b replacing A; this leads to (19) with A replaced by A_b, i.e., to (23).

Second, consider variations which transfer some of the searching effort in A_b to $\bar{A}_b$. They are obtained as follows: Let $\psi(x,y)$ be any function continuous in A and satisfying the conditions

$$\psi(x,y) \leqq 0 \text{ for } (x,y) \text{ in } A_b,$$

$$\psi(x,y) \geqq 0 \text{ for } (x,y) \text{ in } \bar{A}_b,$$

$$\phi(x,y) + \psi(x,y) \geqq 0 \text{ for } (x,y) \text{ in } A_b,$$

$$\iint_A \psi(x,y)dxdy = 0.$$

Evidently if $0 \leqq \xi \leqq 1$, $\xi\psi(x,y)$ will have these same general properties as $\psi(x,y)$. Then $\phi(x,y) + \xi\psi(x,y)$ represents the result of *varying* $\phi(x,y)$ in the manner described. It remains to show that if

$$P(\xi) = \iint_A p(x,y)\,[1 - e^{-(\phi + \xi\psi)}]dxdy$$

then $P'(\xi) \leqq 0$. We compute as follows:

$$P'(0) = \iint_A p(x,y)\, e^{-\phi(x,y)}\, \psi(x,y)dxdy$$

$$= \iint_{A_b} p(x,y)\, e^{-\phi(x,y)}\, \psi(x,y)dxdy$$

$$+ \iint_{\bar{A}_b} p(x,y)\, e^{-\phi(x,y)}\, \psi(x,y)dxdy.$$

From (23),

$$p(x,y)\, e^{-\phi(x,y)} = b \text{ when } (x,y) \text{ is in } A_b,$$

$$p(x,y)\, e^{-\phi(x,y)} < b \text{ when } (x,y) \text{ is in } \bar{A}_b,$$

hence

$$P'(0) = b\iint_{A_b} \psi(x,y)dxdy + \iint_{\bar{A}_b} p(x,y)\psi(x,y)dxdy,$$

$$\leqq b\left[\iint_{A_b} \psi(x,y)dxdy + \iint_{\bar{A}_b} \psi(x,y)dxdy\right],$$

$$= b\iint_A \psi(x,y)dxdy = 0.$$

This completes the proof that (23) gives the solution required.

It is possible to put this result in a geometrical form. In the rectangular coordinates of the variables x,y,z, plot the surface $z = \log p(x,y)$ over the region A in the xy plane. Cut this surface by the horizontal plane $z = \log b$, choosing the constant b so that the volume cut off (above the plane and below the surface) shall equal Φ, and denote the orthogonal projection upon the xy plane of the portion of the surface above the plane by A_b. This construction is the geometrical counterpart of equation (22) defining A_b. The search density $\Phi(x,y)$ is then taken as zero outside A_b (i.e., wherever the surface is below the plane) and equal to the length cut from a vertical line by the surface and the plane, at each point (x,y) of A_b (through which point the vertical line is drawn). This is a consequence of $\phi(x,y) = \log p(x,y) - \log b$, which is contained in (23).

This construction shows how to lay on additional searching in case an additional search effort $\Delta\Phi$ becomes available after Φ has been used up fruitlessly: Lower the horizontal plane so that the additional volume between it and the surface is $\Delta\Phi$; then search throughout the (generally larger) region A_b with a search density equal to the length of the vertical segment between the new and the old horizontal planes, or, in new parts of A_b, between the new plane and the surface. The resulting total layout of search density is obviously that which corresponds to the total available effort $\Phi + \Delta\Phi$. To prove that the knowledge, that the first part of the search (with Φ) has failed to detect the target, leads by Bayes' theorem to the above as the optimum procedure, one employs the method of reasoning illustrated in the corresponding question in Section 3.2.

These considerations show how to carry out a search in time as the available effort $\Phi = \Phi(t)$ increases progressively, the purpose being to detect as early as possible: Lower the horizontal plane $z = \log b$ at such a rate that the volume between it and the surface $z = \log p(x,y)$ constantly equals $\Phi(t)$, and constantly add search density by amounts equal to the additional lengths of vertical segments described above.

It may be noted that a very slight change in wording of certain of the proofs in this section shows that the solutions here given continue to be valid when the condition $p_0 > 0$ and p_1 finite are both dropped; and also, when the region A is infinite in area. Actually, $p(x,y)$ need not even be continuous. And of course the distribution can be in a line or in space rather than in the xy plane, with appropriate rewording. Finally, this problem specializes and becomes like the one of Section 3.2, as is seen on setting $A = A_1 + A_2$ and taking $p(x,y)$ equal to p_1/A_1 or p_2/A_2 according as (x,y) is in A_1 or in A_2. Then equation (23) will reduce to such equations as (9), etc. Similarly, equation (23) yields the solution of the n region case spoken of in the last paragraph of Section 3.2.

Corollary: If in all the previous discussion and equations, $p(x,y)$ is interpreted not as the probability density of the distribution of an individual target in A, but as the expected number of targets per unit area, i.e., $p(x,y)dxdy$ is now the mean or expected number in the infinitesimal $dxdy$ region at (x,y) where there is an indefinite number or swarm of targets in A (which may be an infinite region), and where the quantity to be maximized is not the probability of detecting *the* target but the expected number detected—then all the conclusions, equations, and geometrical constructions remain valid.

This can of course be proved by paralleling the whole of the previous discussion. But it is simpler to observe that the probability of an arbitrarily chosen individual target of the swarm being in $dxdy$ is *proportional* to $p(x,y)$, and then to show that the constant of proportionality cancels out of the crucial equations (22) and (23) and makes no difference in A_b, properly defined.

3.4 AN APPLICATION

Let an approximately stationary target be placed on the ocean according to a bivariate circular normal distribution centered at the origin of the coordinate system, and of standard deviation σ; thus

$$p(x,y) = \frac{1}{2\pi\sigma^2} e^{-r^2/2\sigma^2}, \quad r^2 = x^2 + y^2. \tag{24}$$

The region A of Section 3.3 is now the whole plane and $p_0 = 0$; but, as has been remarked, the solution given in 3.3 is valid for such a case, and the optimum distribution of the available searching effort Φ is given by (23) together with (22).

Let a be the distance r at which $p(x,y)$ assumes the critical value b. Then A_b is the region for which $r \leqq a$, and evidently

$$A_b = \pi a^2, \quad b = \frac{1}{2\pi\sigma^2} e^{-a^2/2\sigma^2}. \tag{25}$$

Substituting these expressions into equation (22) and eliminating b, we derive the expression for a,

$$a^4 = \frac{4\sigma^2\Phi}{\pi}. \tag{26}$$

Thus a is proportional to $\sqrt{\sigma}$; also to $\sqrt[4]{\Phi}$.

Outside the circle centered at the origin and of radius a no searching should be done. Inside the circle the searching effort should be distributed according to the formula (23) which in this case reduces to

$$\phi(x,y) = \frac{a^2 - r^2}{2\sigma^2}. \tag{27}$$

The graph of the equation $z = \phi(x,y)$ in the space of the coordinates (x,y,z) is a paraboloid of vertex at $(0,0,a^2/2\sigma^2)$, axis coincident with the z axis, cutting the xy plane in the above circle of radius a, and is the xy plane itself outside this circle.

To consider a concrete case, let $\sigma = 100$ miles (which means that there is half a chance that the target is within a circle of 118-mile radius). Assume that there are five 130-knot aircraft, each available for five hours of search, so that $L = 3{,}250$ miles. Assume a search width on the present type of target of $W = 5$ miles. Thus $\Phi = WL = 16{,}250$ square miles; by (26) $a = 121$ miles. Only within a circle of this radius should searching take place, and there the path length per square mile should depend on the distance r from the center according to the formula

$$\frac{\phi}{W} = \frac{121^2 - r^2}{100{,}000}.$$

3.5 DISTRIBUTION OF EFFORT IN TIME

In certain tactical situations involving a planned search there is a question not of the distribution of searching effort in space but in time. In order to illustrate such cases, a typical but simple problem will now be considered.

A certain relatively narrow region A of the ocean has to be crossed by very fast enemy surface units. It takes each one a definite time T to cross A, T being of the order of an hour or two. We wish to detect such units by means of aircraft patrolling the region A: Certain features of the tactical situations require them to stay within A at all times. We have at our disposal a fair number of aircraft of the same type, each capable of a definite number of flying hours during the twenty-four, 6 to 12, for example. Thus the total length of track of all aircraft during twenty-four hours has a fixed value of L miles; this is a measure of the total available searching effort. During the twelve daylight hours, radar search can be supplemented by visual, the combined power of detection being expressed by the value W_1 of the search width for each separate aircraft. During the twelve hours of darkness, radar is the sole means of detection, and the search width falls to a value W_2,

$$W_1 > W_2. \tag{28}$$

One final assumption is fundamental: The enemy, not being aware of our search of the region A, is as likely to cross this region at any one time as at any other.

It may be assumed that for the best search the number of aircraft patrolling during any one hour of the twelve daylight hours should be the same as during any other. For if during a particular hour there are fewer patrols than during another, the loss of chance of detection during the former is not quite compensated by the additional chance during the latter, on account of the tendency of overlapping (saturation effect), which is always present, but increases with increasing number of patrols in a given region. This situation could of course be given a precise mathematical formulation and proof, with which, however, we will dispense. In conclusion, we may say that in any scheme of search to be considered there exists a constant number n_1 of planes airborne during the day; and, by corresponding reasoning, a constant number, n_2, during the night. If v is the common aircraft speed, the total daytime and nighttime length of track flown is $12vn_1$ and $12vn_2$ miles respectively. Thus the following equation, expressing the limited total searching effort, must hold.

$$n_1 + n_2 = \frac{L}{12v}. \tag{29}$$

Beyond this, n_1 and n_2 are at our disposal; they characterize the distribution of searching effort between day and night.

If the enemy unit crosses A during daylight, the total length of track flown while he is in A will be Tvn_1 (there are n_1 planes in the air, each of speed v, during the time T of passage of the enemy). Hence by the formula of random search[a] the probability of detecting such a target is

$$1 - e^{-Tvn_1W_1/A},$$

while the probability of detecting a target passing at night is

$$1 - e^{-Tvn_2W_2/A}.$$

Since the chance of the enemy's crossing A by day or by night are the same, the sum of one-half of each of the above expressions gives the required probability of detection p (assuming for simplicity, and with satisfactory approximation, that the chance of a passage of A partly at night and partly in daytime is of negligible probability):

$$p = 1 - \frac{1}{2}\left(e^{-Tvn_1W_1/A} + e^{-Tvn_2W_2/A}\right). \tag{30}$$

Mathematically, our problem is to choose n_1 and n_2 subject to (29) so as to maximize p. The details of the work are altogether similar to those of Section 3.2, except that n_1 and n_2 are the variables in the present case. The results are as follows:

Case 1. L is so small that

$$W_1 e^{-TLW_1/12A} \geqq W_2, \tag{31}$$

[a] It is assumed here, as throughout this chapter, that the flights are at random. When the *direction* of the enemy target across A is known, a more efficient disposition of aircraft tracks is in the form of a crossover barrier patrol (Chapter 7); the formulas, *but not the ideas and essential results*, will then be changed.

then

$$n_1 = L/12v, \quad n_2 = 0, \tag{32}$$

and all the flying must be done in daytime.

Case 2. L is large enough to make

$$W_1 e^{-TLW_1/12A} < W_2, \tag{33}$$

then both day and night searching must be done according to the equations [analogous to (9)]

$$n_i = \frac{\log\left(\frac{TvW_i}{A}\right)}{\left(\frac{TvW_i}{A}\right)} + \frac{\frac{1}{W_i}}{\frac{1}{W_1}+\frac{1}{W_2}}\left\{\frac{L}{12v} - \left[\frac{\log\left(\frac{TvW_1}{A}\right)}{\left(\frac{TvW_1}{A}\right)} + \frac{\log\left(\frac{TvW_2}{A}\right)}{\left(\frac{TvW_2}{A}\right)}\right]\right\} \tag{34}$$

where $i = 1$ and 2.

As long as L is only moderately greater than the critical size expressed in equations (31) and (33), i.e., for which (31) becomes an equality, more searching should be devoted to day than to night ($n_1 > n_2$). But in the limit, for increasing L, (34) shows that n_1/n_2 approaches W_2/W_1, which is less than unity; thus again we have the phenomenon of *reversal* signalized in Section 3.2: If a very large amount of searching effort is available, *more* searching should be done during the unfavorable period than during the favorable one.

3.6 OPTIMUM SCANNING

Returning to the framework of ideas and notation of Chapter 2, suppose that the observer and target are on straight courses at fixed speeds, so that relative to the observer the target is moving in a straight line at the speed of w knots, his lateral range being x miles (see Chapter 1, Figure 5). It is convenient here to regard the target as moving down the line parallel to the axis of ordinates cutting the axis of abscissas at the point of abscissa $x > 0$. His position at the epoch t is at the point of coordinates (x,y), where x remains constant and $y = -wt$ (the negative sign, because of his downward motion: y decreases as t increases). The observer remains at the origin.

Instead of taking the instantaneous probability of detection $\gamma_t dt$ for granted, as we largely did in Chapter 2, we here propose to inquire into it more deeply, and in particular to examine the effect of varying the method of directing the line of sight (or the radar or sonar beam) from position to position over the field of view. We propose, in other words, to examine the effect of different scanning procedures upon γ_t, and through it, upon the search width W. And we will say that *the scanning method is optimum if it renders W a maximum.*

For any method of detection (visual, radar, sonar, etc.), there exists a quantity λ defined as follows: Let the relative positions of the target and observer remain virtually unchanged during a short interval of time dt, during which the observer directs his axis of vision (or the radar or sonar beam axis) straight at the target. Assuming no previous detection, the probability of detection during dt is λdt, which can be called the instantaneous line-of-sight detection probability. We shall assume that λ depends only on range, $\lambda = \lambda(r)$. We shall assume furthermore that the probability of detection during dt is insignificantly changed by a slight change in the position of the target out of the line of sight, e.g., by an angle less than ϵ radians, but that it falls virtually to zero at greater angles. The reader will appreciate that this assumption is not unrealistic in the important case of a target close to the threshold of visibility (or with narrow radar or sonar lobes).

For any method of scanning, there exists a function $f(z)$ (where z is an angle in radians measured from the positive axis of ordinates in the clockwise sense), defined as follows: $f(z)dz$ is the length of time out of a complete scanning cycle during which the axis of vision is between the angles z and $z + dz$. If the total time of one scanning cycle is T, obviously

$$\int_0^{2\pi} f(z)dz = T; \quad \text{also} \quad f(z) \geqq 0. \tag{35}$$

It is now possible to obtain γ_t, the instantaneous detection probability density resulting both from the instrument of detection and its use (method of scan). Let the angle ζ (from the positive axis of ordinates to the target, Figure 7 of Chapter 2) and the range r remain practically constant during one scanning cycle (slow relative motion or fast scan). Then at each cycle the length of time during which the target is within the angle ϵ of the visual axis is

$$\int_{\zeta-\epsilon}^{\zeta+\epsilon} f(\zeta)d\zeta = 2\epsilon f(\zeta) \text{ to terms of higher order in } \epsilon.$$

Hence the probability of detection is

$$g = 2\epsilon f(\zeta)\lambda(r).$$

This is the one-glimpse probability of detection; indeed, the idea of scanning automatically commits us to the notion of detection by discrete glimpses rather than by continuous looking. But in view of our assumption of a scanning which is fast with respect to the relative motions, it is a legitimate approximation to pass to the latter viewpoint, and to convert g into γ_l by division by T (compare with the similar reasoning in Section 4.4, equation (7) in particular). This we shall write

$$\gamma_l = \gamma(r,\zeta) = \frac{2\epsilon}{T} f(\zeta)\lambda(r). \tag{36}$$

It is to be noted that whereas in the greater part of Chapter 2, $\gamma_l = \gamma(r)$, a function of range alone, the very nature of the present considerations focuses attention upon a γ_l which depends both on target range and bearing. Nevertheless, the relevant reasoning and formulas of Section 2.4 are applicable, and we have as the expression for the search width

$$W = \int_{-\infty}^{\infty}\left[1 - \exp\left(-\frac{2\epsilon}{Tw}\int_{-\infty}^{\infty} f(\zeta)\lambda(r)dy\right)\right]dx. \tag{37}$$

The mathematical nature of the problem is now clear. The quantities w, ϵ, $\lambda(r)$ are fixed by the conditions, whereas T and $f(\zeta)$ are at our disposal, subject only to the conditions (35) and that T must be small in comparison with the time required for an appreciable change in relative position of target and observer. And *we have to find that function $f(\zeta)$ which makes W a maximum.*

By an easy argument of symmetry, the optimum $f(\zeta)$ is symmetrical about the axis of ordinates: $f(-\zeta) = f(\zeta)$. Assuming this, the integrand in (37) becomes an even function of x, so that W is twice the integral from 0 to ∞. Similarly, the integral in (35) can be replaced by twice its value from 0 to π. Writing for convenience

$$\phi(\zeta) = \frac{2}{T} f(\zeta), \quad \Lambda(r) = \frac{\epsilon}{w}\lambda(r)$$

(35) and (37) are further simplified, and our problem is reduced to the following:

Find that function $\phi(\zeta)$, $0 \leqq \zeta \leqq \pi$, which, among all functions satisfying

$$\int_0^{\pi} \phi(\zeta)d\zeta = 1, \quad \phi(\zeta) \geqq 0, \tag{38}$$

makes the expression

$$\frac{W}{2} = \int_0^{\infty}\left[1 - \exp\left(\int_{-\infty}^{\infty} \phi(\zeta)\Lambda(r)dy\right)\right]dx \tag{39}$$

a maximum. This is an "irregular" problem in the calculus of variations, not to be solved simply by equating certain variations to zero.

Introducing polar coordinates (r,ζ), (39) becomes

$$\frac{W}{2} = \int_0^{\infty}\left[1 - \exp\left(-x\int_0^{\pi}\phi(\zeta)\Lambda(x\csc\zeta)\csc^2\zeta d\zeta\right)\right]dx. \tag{40}$$

The integral with respect to x will increase when its (nonnegative) integrand increases; and such an increase will take place when the integral in the exponent, i.e.,

$$U = \int_0^{\pi} \phi(\zeta)\Lambda(x\csc\zeta)\csc^2\zeta d\zeta$$

is increased. The evident difficulty is that U involves x as well as $\phi(\zeta)$. If we fixed the value of x we could try to find the function $\phi(\zeta)$ maximizing U subject to (38); but the resulting $\phi(\zeta)$ could be expected to be one function for one value of x, and a different one for another x, and consequently useless as far as maximizing $W/2$ is concerned. The remarkable fact is that in a broad class of cases, including all those important in practice, this turns out not to be the case. Indeed we have the theorem:

If $\lambda(r) = \lambda_1(r)/r^2$, where $\lambda_1(r)$ decreases with increasing r, the optimum scan consists in fixing the line of sight directly along the axis of abscissas, dividing the time equally between right and left. When most of the relative motion is due to the observer (e.g., a searching aircraft), this means that *all scanning should be done abeam.*

To prove this theorem, we note that $\Lambda(r) = \Lambda_1(r)/r^2$, where $\Lambda_1(r)$ decreases as r increases. Hence

$$U = \frac{1}{x^2}\int_0^{\pi} \phi(\zeta)\,\Lambda_1(x\csc\zeta)\,d\zeta,$$

and our problem is to choose $\phi(\zeta)$ subject to (38) which maximizes this integral. As ζ goes from 0 to π, $x\csc\zeta$ decreases from $+\infty$ to a minimum of x at $\zeta = \pi/2$, and then increases to $+\infty$ again. Hence Λ_1 has its *maximum* when $\zeta = \pi/2$, and this is true for any $x > 0$. The graph of $\Lambda_1(x\csc\zeta)$ against ζ is shown in Figure 2, which also shows that of $\phi(\zeta)$ [with shading under it; the shaded area must be unity by virtue of (38)].

Geometrically, the problem is to deform the $\phi(\zeta)$ curve (always maintaining it above the ζ axis and bounding unit area) so that the area under the product ordinate curve $\phi(\zeta)\Lambda_1(x \csc \zeta)$ shall be a maximum. Obviously, the more the $\phi(\zeta)$ is peaked about the mid-point $\zeta = \pi/2$ the larger the area under the

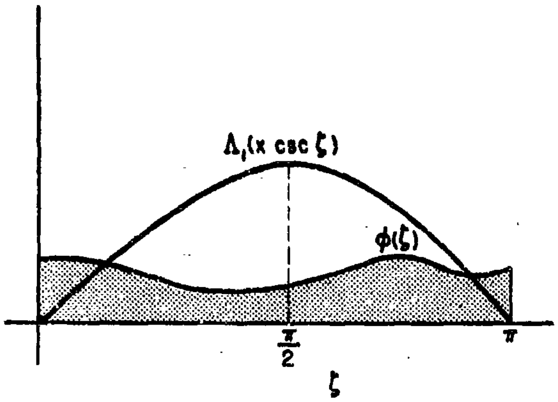


FIGURE 2. Graphical representation of the scanning problem.

product curve. In other words, the more of the time the axis of vision is directed at the angle $\zeta = \pi/2$, the greater will be the value of the search width W. This proves the above theorem.

The case $\lambda(r) = \lambda_1(r)/r^2$, $\lambda_1(r)$ decreasing with increasing r occurs when $\lambda(r)$ is an inverse nth power with n greater than 2, in particular, with the inverse cube power law; also, when there is an exponential attenuation factor multiplying an inverse square or higher power. In fact, it occurs in the majority of cases studied in Chapters 4, 5, and 6. And the theorem in the case of a definite range law is trivially true.

It would be misleading to conclude that scanning should always be confined to the beam. In most cases it is imperative to detect the target *early*, for example, when it is a surfaced submarine which may submerge if not detected and attacked before it sees us, or in the case where it may be expected to attack us as soon as it sees us.

3.7 OPTIMUM DESTRUCTIVENESS

There is a close analogy both in thought and in mathematics between the problem of the optimum distribution of searching effort and the optimum distribution of gunfire, bombs, or other types of destructive missiles which are projected into an area. To bring this out, it will suffice to consider a question which leads to results completely parallel to those of Section 3.2.

Two areas A_1 and A_2 contain targets of the nature of factories and other buildings important for the enemy's economy. High-level night bombing is contemplated, and it is assumed that such bombing is sufficiently accurate to place the bombs in A_1 or in A_2, but not to hit particular targets in these areas except by the operation of chance. Let the total vulnerable area of targets in A_1 and A_2 be B_1 and B_2 respectively. Assume, furthermore, that the vulnerability of all targets is the same, and that the result of n random hits on the target area (B_1 or B_2) is to reduce its effectiveness (value to the enemy) by a proportion $1 - k^n$ (where $k < 1$). This means that if the B_i targets had a value to the enemy (productivity, rate of casualties it could inflict on us, etc.) expressible by a number E, the value after n hits on A_i would be reduced to k^nE, the reduction in value being $E - k^nE$, and hence the proportional reduction in value being the ratio of this quantity to E. Assume, finally, that the value to the enemy of a set of targets is proportional to its total area.

If N similar bombs are at our disposal and if it is as easy to drop bombs in A_1 as in A_2, the decision as to how to divide the bombs between A_1 and A_2 is properly made on the basis of maximizing the expected damage to be inflicted.

Let N_1 bombs be dropped in A_1 and N_2 in A_2,

$$N_1 + N_2 = N, \quad N_1 \geqq 0, \quad N_2 \geqq 0. \tag{41}$$

The probability that one bomb dropped in A_1 will hit the target is B_1/A_1. If N_1 are dropped in A_1, the probability that R_1 of them will hit the target is (by the binomial distribution) equal to

$$\frac{N_1!}{R_1!\,(N_1 - R_1)!}\left(\frac{B_1}{A_1}\right)^{R_1}\left(1 - \frac{B_1}{A_1}\right)^{N_1 - R_1}.$$

The proportional damage done by the R_1 hits being $1 - k^{R_1}$, the *expected damage* produced by N_1 bombs dropped in A_1 is given by

$$\sum_{R_1=0}^{N_1}\frac{N_1!}{R_1!\,(N_1 - R_1)!}\left(\frac{B_1}{A_1}\right)^{R_1}\left(1 - \frac{B_1}{A_1}\right)^{N_1 - R_1}\left(1 - k^{R_1}\right) = 1 - \left(1 - \frac{B_1\,(1-k)}{A_1}\right)^{N_1},$$

the last equation being established by the algebra of the binomial theorem. It is convenient to write (for $i = 1$ and 2)

$$\mu_i = -\log\left(1 - \frac{B_i(1-k)}{A_i}\right), \tag{42}$$

$$= \frac{B_i(1-k)}{A_i} \text{ approximately, when } B_i/A_i \text{ is small.}$$

With this expression, the *expected* (proportional) *damage* is

$$1 - e^{-\mu_i N_i}.$$

Since, as we are assuming, the value to the enemy is proportional to the target area, the expected loss of value to him when N_1 and N_2 bombs are dropped in A_1 and A_2 respectively is proportional to

$$V = A_1(1 - e^{-\mu_1 N_1}) + A_2(1 - e^{-\mu_2 N_2}). \tag{43}$$

This is the quantity which has to be maximized by a choice of N_1 and N_2 satisfying conditions (41). Obviously the problem is mathematically identical with that of Section 3.2. Accordingly, it suffices to enumerate the results.

Case 1: $\mu_2 A_2 e^{-\mu_2 N} \geqq \mu_1 A_1$.

All N bombs should be dropped in A_2.

Case 2: $\mu_1 A_1 e^{-\mu_1 N} \geqq \mu_2 A_2$.

All N bombs should be dropped in A_1.

Case 3:

$$\mu_2 A_2 e^{-\mu_2 N} < \mu_1 A_1$$

and

$$\mu_1 A_1 e^{-\mu_1 N} < \mu_2 A_2.$$

Then some bombs must be dropped in A_1 and some in A_2 according to the formulas (for $i = 1$ and 2)

$$N_i = \frac{\log \mu_i A_i}{\mu_i} + \frac{\frac{1}{\mu_i}\left[N - \frac{\log \mu_1 A_1}{\mu_1} - \frac{\log \mu_2 A_2}{\mu_2}\right]}{\frac{1}{\mu_1} + \frac{1}{\mu_2}}. \tag{44}$$

When N is very large, $N_1/N_2 = \mu_2/\mu_1$, which according to (42) reduces to $(B_2/A_2)/(B_1/A_1)$, so that the numbers dropped are *inversely* proportional to the probabilities of hitting the targets.

Chapter 4

VISUAL DETECTION

4.1 INTRODUCTION

This chapter deals with the problem of visual search under the general conditions of illumination which obtain during the daylight hours. The theoretical studies described are based on laboratory tests, operational data, and service trials. They attempt to answer two of the questions proposed in Chapter 2. These are, first: What is the maximum range within which a given target *can* be seen, and second: What is the chance that the target *will* be seen while it is at any given range? The answers require some knowledge of the construction and performance of the eye considered as a detecting instrument. In what follows in this chapter, therefore, the eye is studied first as a detecting instrument and then as applied to specific operational problems.

4.2 THE EYE AS A DETECTING INSTRUMENT—GENERAL DESCRIPTION

In general construction, the eye is very similar to a camera. The transparent front surface or *cornea* and the *crystalline lens* together constitute a compound lens which forms on the *retina*, at the back wall of the eye, an image of any given object in front of the eye. Between the *cornea* and the *crystalline lens* there is a small aperture known as the *pupil*. This aperture is variable in size over a limited range and determines the quantity of light which enters the eye.

The *retina* corresponds to the sensitized plate or film in the camera. It contains two different types of sensitive elements or receptors known as *rods* and *cones*. The rods serve for night vision and are incapable of distinguishing color. The cones are responsible for vision in daylight and for all color vision. The central part of the retina, through which the visual axis passes, is known as the *fovea*. The visual axis makes a small angle with the optic axis of the compound lens system. The diameter of the *fovea* subtends an angle of between 1 degree and 2 degrees at the effective center of the lens (actually the nodal point of the compound lens). The fovea is the region of most distinct daylight vision. It contains cones only, the average angular distance between centers of adjacent cones being about 0.5 minute of arc. As the angular distance from the axis increases beyond the edge of the fovea (i.e., as the parafoveal region is entered), the number of cones in unit area decreases, at first rapidly and then more slowly while the number of rods in unit area gradually increases out to about 18 degrees and then decreases. In daylight, therefore, a given target can be seen most easily by looking straight at it while at night a better view is obtained by looking about 6 degrees off from the most direct line of sight.

Both rods and cones are capable of adjusting themselves for the general level of illumination to which they are exposed. This adjustment is known as *adaptation*. For both rods and cones adaptation is much more rapid to an increase in illumination than to a decrease. Furthermore, both light and dark adaptation are more rapid for cones than for rods. Dark adaptation for the cones takes several minutes, whereas for the rods this time is of the order of half an hour.

Unlike the radar which scans continuously, the eye moves in jumps while searching and is capable of vision only during periods of little or no motion. These periods are known as *fixations* and last for about 0.25 second. In a given fixation or group of fixations, a target at extreme range can be seen in daylight only on the fovea so that the visual axis must be well within 1 degree of the line joining the target and the eye. As the range decreases, regions outside the fovea become capable of detecting the target, at first those near the fovea and then those farther out. Hence targets at less than extreme range can be seen not only on the fovea but off fovea as well, i.e., "out of the corner of the eye."

4.3 THE EYE AS A DETECTING INSTRUMENT—DETAILED PERFORMANCE

The characteristics of the target and its background, which determine whether or not the target can be seen, are:

1. Brightness of the background.
2. Brightness of the target.
3. Color of the background.

4. Color of the target.
5. Size of the target.
6. Range or distance of the target.
7. Shape of the target.

The background brightness, by reason of light adaptation, determines that differential sensitivity with which the eye discriminates as to differences in brightness between the object and its immediate surroundings. When the background is nonuniform, the sensitivity is set by an *effective background*. Possibly because of the fact that light adaptation is more rapid than dark, the contributions of the lighter portions of the field to the state of adaptation is larger than that of the darker portions. For a given target at a given range under daylight illumination, it is the magnitude of the difference between target and immediate background brightness, expressed in units of effective background brightness, which determines whether or not the target can be seen. This quantity is defined, for purposes of this chapter, as the *brightness contrast*. Under conditions of daylight illumination, i.e., for all illuminations greater than that of early twilight, a given contrast will cause the same visual response regardless of the magnitudes of the various brightnesses which go to make up this contrast. As the illumination is decreased below that of early twilight, the various brightnesses involved enter explicitly. Since this chapter is concerned with daylight illumination only, this further complication is not considered. It is to be remembered that the definition of brightness contrast given above for purposes of this chapter is not the usual one found in the literature.

It has long been believed that in comparison with brightness contrast, color is of little importance in determining whether or not a given target can be seen. Recent investigations[1] have supported this belief and have shown that any effects due to color can be ignored in most operational problems of visual search without thereby introducing any appreciable errors. With this as justification, all effects due to color are ignored in this chapter.

The size of the target and its range combine to determine the solid angle which the target subtends at the eye and hence the size of the image on the retina. Hence the three characteristics of the target and its background upon which the discrimination of the eye depends under daylight illumination are:

1. Contrast of the target against its background.
2. Solid angle subtended by the target.
3. Shape of the target.

The two sets of measurements employed to determine the effects of these three variables are those of K. J. W. Craik[2,3] and some unpublished measurements made in collaboration with Selig Hecht and Simon Shlaer at the Laboratory of Biophysics, Columbia University.[4] The Columbia experiments were designed primarily to determine the effect of object shape. While they are more detailed than those of Craik, they are not as extensive in retinal areas investigated. The results of both sets of measurements constitute the primary laboratory data for this chapter. Craik's measurements are employed to determine the effect of solid angle, while those made at Columbia serve to determine the effect of target shape and as a check on the Craik experiments for those retinal regions for which both measurements are available.

For any given target, the quantity measured in both the Craik and the Columbia experiments was the just-perceptible or *threshold contrast*. As is the case with any measurement, this quantity is not determined exactly but within some tolerance or uncertainty. In the Craik experiments, that contrast above which the target could always be seen and that below which it could never be seen were measured and the average of these two was taken as the threshold contrast. In the Columbia experiments a frequency method was employed to determine the probability of sighting the target as a function of contrast. That contrast for which the sighting probability was 57 per cent was taken as the threshold contrast. The results of the two sets of measurements are in excellent agreement so that the two experiments evidently measure the same quantity.

From the results of the Craik experiments with circular targets it was found that the threshold contrast C_t can be represented as a function of the solid angle ω subtended by the target at the eye, by the following equation:

$$C_t = a + \frac{b}{\omega}, \tag{1}$$

where a and b are constants for any one retinal region. Instead of using the solid angle ω, it is often more convenient to employ the square of the *visual angle* α, i.e., *the angle subtended at the eye by the diameter of the equivalent circle*. The quantities a and b have different values at different angular distances from the center of the fovea. If θ is this angular distance in degrees, from center of the equivalent circle to center of fovea; α, the visual angle in minutes; and

C_t the threshold contrast in per cent, Craik's data can be represented by the following equation:

$$C_t = 1.75\theta^{\frac{1}{2}} + \frac{19\theta}{\alpha^2}. \qquad (2)$$

The value of θ which must be employed to obtain the foveal data from equation (2) is 0.8 degree and the threshold contrast is constant between $\theta = 0$ degrees and $\theta = 0.8$ degree. It is to be pointed out that equation (2) is purely empirical. It represents the Craik experiments with fair accuracy. Since the number of measurements made by Craik is relatively small, the experimental error is fairly high: Hence slight modifications of equation (2) are to be expected when the more extensive Columbia experiments are complete. The reader will note that in the present chapter α and θ are used in an altogether different meaning from elsewhere in this book.

The targets employed in the Columbia experiments were all rectangles and the quotient q, the ratio of length to width, was taken as a measure of the asymmetry. This quotient q is known as the *asymmetry factor*. The results of these experiments for the different background brightnesses and for the various retinal regions investigated all show the same general trends. As the shape of a small target is changed, keeping its angular area constant, the threshold contrast remains constant until the asymmetry factor reaches a value such that the angular length of the target is about 3 minutes of arc. As the asymmetry factor is increased beyond this point, the threshold contrast gradually increases. As the shape of a large target is changed, keeping its angular area constant, the threshold contrast gradually decreases until the asymmetry factor is such that the angular width of the target is about 2 minutes of arc. Beyond this point, the threshold contrast again increases. The greatest effect of asymmetry is observed when the angular size of the target is about 10 square minutes. As the asymmetry factor is changed from 2 to 200, the threshold contrast increases by a factor of 4. Between 2 and 100 the factor is 3.

Two further fundamental facts have been established by the Columbia experiments: If the position of the target is known so that no search is required, the probability of sighting the target is independent of the time of exposure provided this time exceeds that required for a single fixation; and the "glimpse" probability g that a target be sighted at a single fixation obeys the following law:

The probability g is a function of the ratio C/C_t = (target's apparent contrast)/(target's threshold contrast alone),

$$g = f\left(\frac{C}{C_t}\right) \qquad (3)$$

and only through C_t do the target's apparent size α and the off axis angle θ intervene. The quantity g is presented in Figure 1 as a function of C/C_t. This is the experimental curve for the function f.

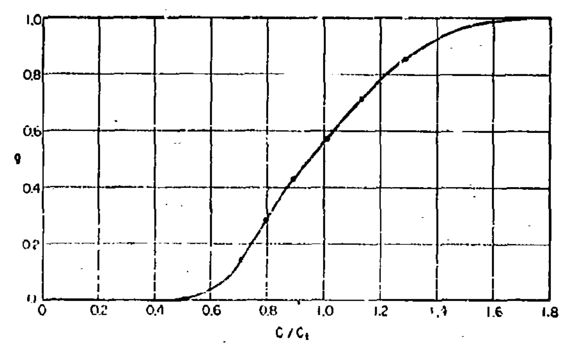


FIGURE 1. "Glimpse" probability of detection as a function of apparent contrast/threshold contrast—experimental data.

4.4 CONTACT PROBABILITY AND SCANNING—GENERALITIES

Let a given target be viewed under given conditions: θ, and its apparent size (visual angle) and contrast, α and C, are therefore given. The latter determine a *threshold* value θ_0 of the angle from the line of sight: mathematically, by the law expressed in (2), i.e., θ_0 is defined by the equation

$$C = 1.75\theta_0^{\frac{1}{2}} + \frac{19\theta_0}{\alpha^2}.$$

On the other hand, θ and α determine a threshold contrast C_t by the same law (2):

$$C_t = 1.75\theta^{\frac{1}{2}} + \frac{19\theta}{\alpha^2}.$$

On combining these equations with (3), the following is derived,

$$g = f\left(\frac{1.75\theta_0^{\frac{1}{2}} + \dfrac{19\theta_0}{\alpha^2}}{1.75\theta^{\frac{1}{2}} + \dfrac{19\theta}{\alpha^2}}\right). \qquad (4)$$

The philosophy of this method of obtaining the expression (4) for g may be described as follows: The triplet of values C, θ, and α determines the required probability g. To find the unknown g, we contemplate successively the triplets C_t, θ, and α and C, θ_0, and α, each determining the standard threshold probability 0.57 (by determining θ_0 from C and α, and C_t from θ and α, by (2) to give this threshold probability). Because of the fundamental equation (3), this effectively furnishes the required g in terms of C, θ, α.

Now consider what happens when the conditions of the last paragraph are varied as follows: The apparent size and contrast, α and C, remaining fixed, the angle θ is varied. Since θ_0 is a function of α and C, it will stay fixed. Thus the letters θ_0 and α in (4) represent unchanging quantities, θ alone varying. Consequently (4) *expresses the manner of dependence of the one-fixation chance of sighting g upon the angle θ off the line of sight.*

For targets of very small apparent size ($\alpha << 10$), (4) reduces to $g = f(\theta_0/\theta)$, a graph of which is shown

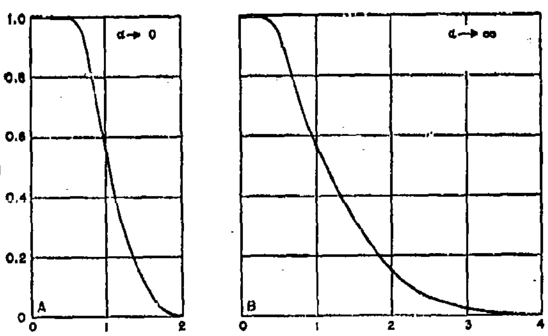


FIGURE 2. "Glimpse" probability of detection as a function of off-axis angle/threshold off-axis angle, for targets of (A) very small, and (B) very large apparent size. θ_0/θ is the x axis.

in Figure 2A. For targets of very large apparent size ($\alpha >> 10$), (4) becomes $g = f\left(\sqrt{\theta_0/\theta}\right)$, plotted in Figure 2B.

Having considered the dependence of g upon θ, it is now necessary to evaluate the one-fixation probability $\bar{g}$ *when the value of θ is unknown but distributed at random with a known frequency*, C and α still being given constants. The circumstances which impose this necessity are the use of the eye to *scan* a given location within which the target is supposed to be, but in an unknown position. Two cases are important in practice:

1. *Linear Scan.* The target is located on a line L and is uniformly distributed thereon; the pertaining quantities C and α are assumed to be independent of the position on L. The eye fixates once upon a random point on L. If the angle subtended at the eye by L is Θ degrees, it is easily shown that when Θ is large $\bar{g}$ is given with sufficient approximation by the formula

$$\bar{g} = \frac{2}{\Theta}\int_0^\infty g\,d\theta. \tag{5}$$

For the chance that the point of fixation O be close to the ends of L can be neglected, so that it can be assumed that L extends a considerable distance on either side of O. The probability that the target be between θ and $\theta + d\theta$ degrees away from O is $d\theta/\Theta$. Hence the chance of sighting is the integral of $g\,d\theta/\Theta$ over the whole angular range Θ. But since g falls to zero at appreciable angles from O [as can easily be shown on the basis of equation (4) and because of what has been said concerning the distance of the extremities of L from O], the limits of integration can be taken as $-\infty$ and $+\infty$ [g being defined as zero for values of θ beyond those contemplated in (4)]. By symmetry of the integrand, twice the integral between the limits 0 and $+\infty$ can be used; hence the validity of equation (5).

In combination with (4), (5) becomes

$$\bar{g} = \frac{2}{\Theta}\int_0^\infty f\left(\frac{1.75\theta_0^{\frac{1}{2}} + \dfrac{19\theta_0}{\alpha^2}}{1.75\theta^{\frac{1}{2}} + \dfrac{19\theta}{\alpha^2}}\right)d\theta$$

$$= \frac{2\theta_0}{\Theta}\int_0^\infty f\left(\frac{1.75\theta_0^{\frac{1}{2}} + \dfrac{19\theta_0}{\alpha^2}}{1.75\theta_0^{\frac{1}{2}}\lambda^{\frac{1}{2}} + \dfrac{19\theta_0\lambda}{\alpha^2}}\right)d\lambda$$

$$= \frac{\theta_0}{\Theta}\Lambda\,(\theta_0,\alpha),$$

where the new variable of integration $\lambda = \theta/\theta_0$. The coefficient $\Lambda\,(\theta_0,\alpha)$ has in the extreme cases $\alpha << 10$ and $\alpha >> 10$, the values

$$2\int_0^\infty f\left(\frac{1}{\lambda}\right)d\lambda = 2.16 \quad \text{and} \quad 2\int_0^\infty f\left(\frac{1}{\sqrt{\lambda}}\right)d\lambda = 2.56$$

respectively, and intermediate values of α are found by graphical integration to lead to intermediate values of $\Lambda\,(\theta_0,\alpha)$. To the degree of approximation which is permissible, it can therefore be assumed to be independent of θ_0 and α. We shall assume henceforth that $\Lambda = \Lambda(\theta_0,\alpha) = 2.36$. This leads to the basic proposition:

The probability g is proportional to the threshold angular distance θ_0 of the target from the visual axis, and is inversely proportional to the angle Θ subtended by the linear locus L of target positions:

$$g = \frac{2.36\theta_0}{\Theta}. \qquad (6)$$

It is now expedient to modify the point of view, and instead of regarding the act of sighting as proceeding by a rapid succession of fixations (glimpses), to envisage it as progressing continuously in time (continuous looking). If dt is an interval of time which is short compared with the time taken for the observer and target to change relative positions by an appreciable amount, and short also in the sense that the chance of a detection during it is small, but large in comparison to the time of a single fixation (actually, one quarter of a second), it becomes legitimate to consider γdt, the probability of detection between the epochs t and $t + dt$. Since when $dt = \frac{1}{4}$ second, $\gamma dt = \bar{g}$, we have $\gamma = 4\,\bar{g}$, i.e.,

$$\begin{aligned}\gamma &= \frac{9.44\theta_0}{\Theta} \text{ (time measured in seconds)} \\ &= 9.44 \times \frac{60^2\theta_0}{\Theta} \text{ (time measured in hours).}\end{aligned} \qquad (7)$$

According to equation (4) of Chapter 2, the mean "pickup time" (time required to detect) is $1/\gamma$. This provides a method of measurement of γ. Craik[3] performed a series of laboratory experiments measuring mean pickup time with line scanning, with $\Theta = 45$ degrees. The results are shown in Figure 3, where γ is plotted against θ_0. The approximate linearity is evident, in accordance with (7). But the coefficient in (7) has to be multiplied by about 0.13, a discrepancy which could be accounted for on the assumption that the fixations in Craik's experiments, instead of being arranged at random, occur effectively in groups of 7 or 8. We will return to this coefficient in a later place in connection with the practical applications.

2. *Area Scan.* The target is located in a region of a plane, subtending the solid angle Ω square degrees. The distribution is uniform over the area, and the quantities C and α are independent of position Reasoning precisely similar to that employed earlier (with double integration replacing single integration) establishes the formula for the one-fixation probability:

$$\bar{g} = \frac{2\pi}{\Omega}\int_0^\infty g\theta d\theta. \qquad (8)$$

Introducing (4) with the variable of integration $\lambda = \theta/\theta_0$ leads to

$$\begin{aligned}\bar{g} &= \frac{\theta_0^2}{\Omega} M(\theta_0,\alpha) \\ M(\theta_0,\alpha) &= 2\pi\int_0^\infty f\left(\frac{1.75\theta_0^{\frac{1}{2}} + \frac{19\theta_0}{\alpha^2}}{1.75\theta_0^{\frac{1}{2}}\lambda^{\frac{1}{2}} + \frac{19\theta_0\lambda}{\alpha^2}}\right)\lambda d\lambda.\end{aligned} \qquad (9)$$

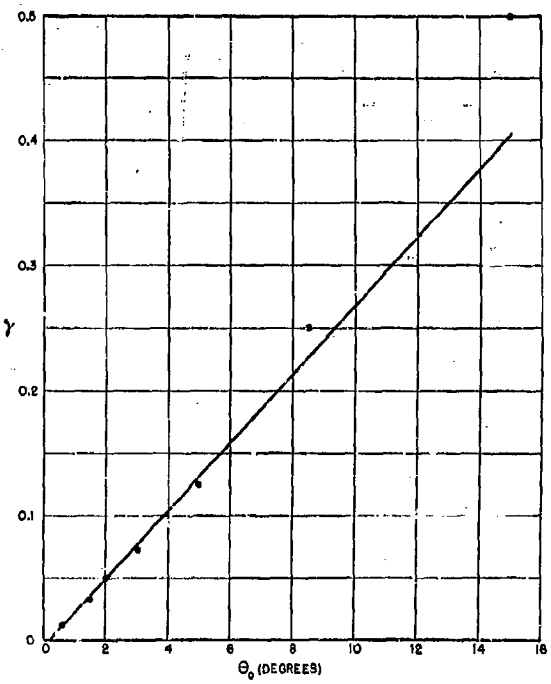


FIGURE 3. Instantaneous probability of detection as a function of threshold off-axis angle—experimental data by Craik.[3]

The extreme values of this quantity occur when $\alpha << 10$ and $\alpha >> 10$; they are respectively

$$2\pi\int_0^\infty f\left(\frac{1}{\lambda}\right)\lambda d\lambda = 4.16 \quad \text{and} \quad 2\pi\int_0^\infty f\left(\frac{1}{\sqrt{\lambda}}\right)\lambda d\lambda = 6.38.$$

Correspondingly, γ is given by

$$\gamma = \frac{\theta_0^2}{\Omega} 4M(\theta_0,\alpha) \quad \text{(time in seconds).} \qquad (10)$$

We shall return to these expressions later.

4.5 TARGET AND BACKGROUND—GENERAL

In the preceding sections, the eye has been considered as a detecting instrument operating on targets of *given apparent size and contrast* (and *shape*, in so far as this is relevant), and the laws governing the probabilities of detection have been set forth. Before these results can be applied to an actual case (viz., a naval operation), it is necessary to find how the circumstances of the case determine these variables and hence, indirectly, the probabilities of detection. Now in any actual case, certain intrinsic characteristics of the target may be regarded as known. These are its geometrical shape and dimensions and its diffuse reflecting power. Except within small angular distances from the sun, the latter determines the intrinsic brightness of the target in units of sky brightness. The intrinsic characteristics of the immediate and general background and the relationships between all intrinsic and apparent quantities are determined by the circumstances of the case. The study of these various dependences and relations is the subject of the two following sections.

4.6 DEPENDENCE OF APPARENT CONTRAST ON ATMOSPHERIC CONDITIONS

The presence of *haze* in the atmosphere alters the pattern of light received by the eye from the various points of the target and background. It acts in two ways:

1. The haze *removes* some of the light by absorption and scattering *out* of the line of sight.
2. The haze *adds* some light reaching the eye from the direction in which the observer is looking by scattering *into* the line of sight.

In order to work out the general equations, let

B = the apparent brightness of the target,
B_0 = the intrinsic brightness of the target,
B^1 = the apparent brightness of the background,
B_b = the intrinsic brightness of the background,
B_s = the brightness of the sky,
β = the atmospheric scattering coefficient,
V = the meteorological visibility,
R = the target range.

The apparent brightness of the target is

$$B = B_0 e^{-\beta R} + B_s (1 - e^{-\beta R}),$$

where the first term represents the light reaching the eye directly from the target, and the second the light scattered into the eye by the haze between the target and the eye. Similarly the apparent brightness of the background immediately surrounding the target is given by

$$B^1 = B_b e^{-\beta R} + B_s (1 - e^{-\beta R}).$$

The difference in brightness between object and immediate background is, therefore,

$$B - B^1 = (B_0 - B_b) e^{-\beta R}.$$

The apparent contrast is the ratio of this brightness difference and the effective background brightness in accordance with the discussion of brightness contrast given in Section 4.3.

There are two main cases which arise in actual operations, one in which the immediate background and the effective background are the same and one in which they are different. The first case is exemplified by search, from land or from a surface ship, for a surface ship silhouetted against the sky. The second case is exemplified by search from the air for a target on the sea. In this case, the line of sight frequently approaches the horizon so that the adaptation of the eye is determined partly by the sea brightness and partly by that of the sky. Because of the fact that light adaptation is rapid compared to dark adaptation, the bright sky is responsible, almost entirely, for setting the level of response of the eye. Hence in both cases, the sky brightness is a reasonable approximation to the effective background brightness and hence

$$C = C_0 e^{-\beta R}, \tag{11}$$

where

$$C_0 = \frac{B_0 - B_b}{B_s}$$

is the intrinsic contrast as defined earlier for purposes of this chapter. It is to be remembered that within 10 or 15 degrees of the sun, the level of response of the eye is altered by the sun's glare so that the effective background brightness is somewhat greater than the average sky brightness B_s. This effect is neglected in the present theory but must be considered in any exact treatment.

The quantity usually quoted as a measure of atmospheric conditions is not the scattering coefficient β but the *meteorological visibility* V. It is desirable, therefore, to express equation (11) in terms of the meteorological visibility rather than in terms of the scattering coefficient. The meteorological visibility is usually defined loosely as the maximum distance at which large targets such as mountains or high coast lines can be seen against the sky. Merton[5] gives 78.3 per cent as the intrinsic contrast of such targets and 2.5 per cent as the threshold contrast. Substituting $C = 2.5$; $C_0 = 78.3$ and $R = V$ in equation (11), $\beta = 3.44/V$. From this it follows that

$$C = C_0 \epsilon^{-3.44\, R/V}. \tag{12}$$

Although the threshold contrast measured in the laboratory is often less than 2.5 per cent, an examination of operational estimates of the meteorological visibility indicates that 2.5 per cent is a good approximation to the practical operation figure.

4.7 INTRINSIC BRIGHTNESS OF THE SEA

The light flux reaching the eye from the sea consists of two parts, that reflected specularly and that reflected diffusely. If the sea is perfectly calm, these two combine to produce an intrinsic sea brightness, 4 per cent that of the sky immediately below the observer and increasing gradually to 100 per cent that of the sky at the horizon. Measurements of the sky/sea brightness ratio[6] indicate that even in mild seas the mirror surface characteristic of a perfectly calm sea is so broken as to cause wide departures from calm sea conditions except within relatively few degrees of a low-lying sun or moon. For sea states other than calm, the intrinsic sky/sea brightness ratio is fairly constant from the horizon to about 45 degrees below and has a value of about 2. For angles greater than 45 degrees below the horizon, the intrinsic brightness of the sea gradually decreases from about 50 per cent at 45 degrees to near 4 per cent directly below the observer.

4.8 MAXIMUM SIGHTING RANGE

With the information in earlier sections concerning the characteristics of target and background on which visual detection depends and the influence of the operational situation on these characteristics, it is possible to obtain answers to the two questions proposed in the introduction. The first of these is: What is the maximum range within which a given target can be seen? *The quantity which is here defined as the maximum range is that range at which the target contrast reaches the foveal threshold.* As was the case with the threshold contrast discussed in Section 4.3, the maximum range is not determined exactly but within some tolerance or uncertainty, so that some targets viewed foveally will be seen beyond this range while others within this range will be missed altogether. We will see, later in this section, the magnitude of this uncertainty, i.e., the spread of ranges over which detection foveally is not a certainty. It is the purpose of this section to determine the maximum sighting range in terms of those quantities which occur in the operational situation. We have given a target of actual area A_0, apparent area A; intrinsic contrast C_0 and asymmetry factor q viewed through an atmosphere in which the meteorological visibility is V. In terms of the apparent area A and the range R, the visual angle α, defined in Section 4.3 is given by

$$\alpha = 0.64 \frac{\sqrt{A}}{R}, \tag{13}$$

where A is in square feet, R in nautical miles and α in minutes of arc.

Consider first the case in which the target is circular and the meteorological visibility is unlimited. Let R_0 be the maximum sighting range under conditions of unlimited visibility. Substituting equation (13) in equation (2) with $\theta_0 = 0.8$, its value for foveal vision, as discussed in Section 4.3, is

$$R_0 = 0.164\,[(C_0 - 1.57)A]^{\frac{1}{2}}. \tag{14}$$

In terms of the actual area of the target A_0, there are two cases which occur frequently in operational situations, one in which the target is viewed at approximately normal incidence, and the other in which the target, flat on the surface of the sea, is viewed at an angle. For the first case, the apparent and actual areas of the target are equal, so that

$$R_0 = 0.164\,[(C_0 - 1.57)A_0]^{\frac{1}{2}}. \tag{15}$$

If h is the altitude of the observer in feet, then for the second case $A = 1.64\,(10^{-4})A_0 h/R_0$. Substituting in equation (14) we have for the second case

$$R_0 = 1.64\,(10^{-2})\,[(C_0 - 1.57)A_0 h]^{\frac{1}{3}}. \tag{16}$$

Consider next a circular target viewed through an atmosphere having a meteorological visibility V and let R_m be the maximum sighting range. The contrast is given by equation (12). Making the indicated substitutions in equation (2) leads to transcendental equations for case I and case II. These can be solved for the meteorological visibility V in terms of the maximum sighting range R_m. These equations are, for case I,

$$V = \frac{1.49 R_m}{\log_{10}\left[\dfrac{C_0 A_0}{36.9\, R_m^2 + 1.57\, A_0}\right]}, \qquad (17)$$

and for case II,

$$V = \frac{1.49\, R_m}{\log_{10}\left[\dfrac{C_0 A_0 h}{2.26\,(10^5)\, R_m^3 + 1.57\, A_0 h}\right]},$$

which can be rearranged to give

$$\frac{V}{h^{1/3}} = \frac{1.49(R_m/h^{1/3})}{\log_{10}\left[\dfrac{C_0 A_0}{2.26\,(10^6)\,(R_m^3/h) + 1.57\, A_0}\right]}. \qquad (18)$$

The effect of target asymmetry reaches its maximum in the following operational situation: A target of the case II type having a ratio of apparent length to apparent width of 100, observed at a range such that the solid angle subtended by the target is 4.5 square minutes under conditions of unlimited meteorological visibility. Under these conditions the maximum range turns out to be half what it would be for a perfectly symmetrical target (disk) of the same intrinsic contrast and solid angle subtended. Except in such rare cases as the extreme just described, the effect of target asymmetry on maximum range is small. This is true not only for the maximum range, but for all similar quantities (e.g., θ_0). In the interest of simplicity, therefore, and with only slight loss of accuracy, *the effect of target asymmetry shall be neglected henceforth.*

We are now in a position to investigate the uncertainty in the range at which a target, viewed foveally, can be seen. This is done by computing the ratio of apparent to threshold contrasts for each of a numbe of ranges and looking up the sighting probabilities corresponding to these ratios from Figure 1. [Essentially, the use of equation (3).] Sighting probabilities, determined in this way, are presented in Figure 4 as functions of the ratio of range to maximum range as defined earlier in this section. The curves A, C are for small high-contrast targets seen under conditions of unlimited meteorological visibility while the curves B, D are for large targets seen through dense haze. From these curves it would appear that, as measurable quantities involving human behavior go, the maximum sighting range is a quantity which

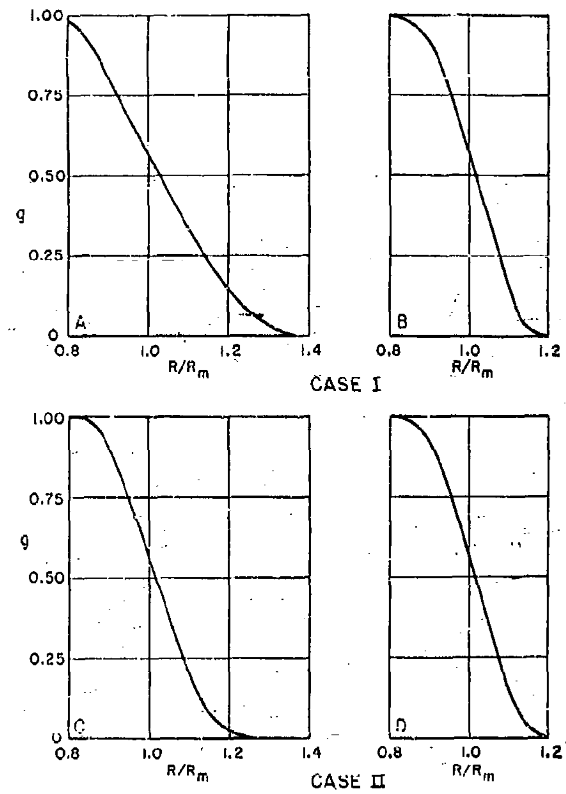


FIGURE 4. "Glimpse" probability of detection as a function of target range/maximum range.

can be determined to within a relatively small uncertainty.

In order to compute the maximum sighting range under some specified set of conditions, using equation (17) or equation (18) as the case may be, it is necessary to know the area and intrinsic contrast of the target. These constants have not yet been determined for any large number of targets. However, there is one large class of targets for which these constants are known reasonably well, i.e., surfaced ships as viewed from an aircraft. Here the target seen first in the vast majority of the cases is the wake. As will be seen in a later section, the intrinsic contrast of a wake is about 50 per cent. Three general classes of ships are considered here, surfaced submarines, ships having medium-sized wakes, and those having large wakes. The second class includes destroyers, destroyer escorts, and medium-sized merchant ships, while the third class includes large combatants and high-speed ocean liners. The wake areas for the three classes are given approximately as 1.3×10^4, 1×10^5, and 2×10^6 square feet, respectively. In Figure 5, $R_m/h^{\frac{1}{3}}$ is presented for each of these cases

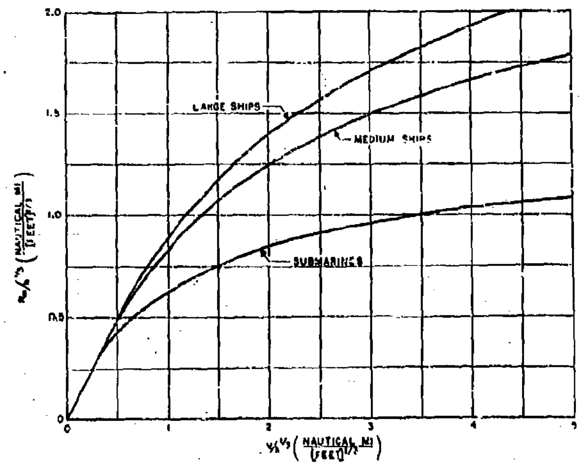


FIGURE 5. Variation of maximum range with visibility for three types of target (naked eye).

as a function of $V/h^{\frac{1}{3}}$ for naked eye search. In Figure 6, the same quantity is presented for search with U. S. Navy standard 7x50 binoculars. The only effect

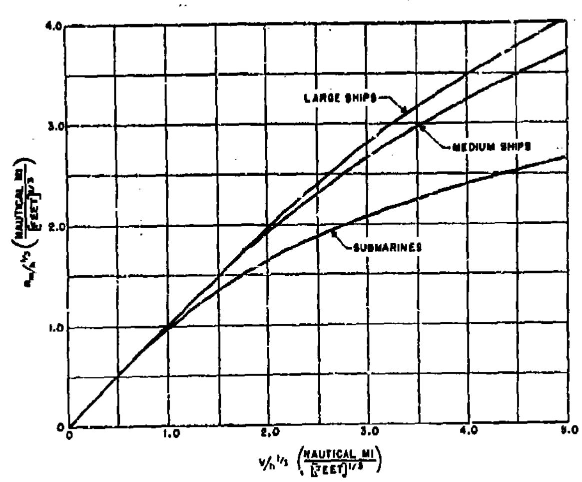


FIGURE 6. Variation of maximum range with visibility (7x50 binoculars).

of the binoculars which has been considered is that of its magnifying power on the apparent target area. Because of the neglect of the slight reduction in apparent contrast (beyond that due to haze) which results from the fact that the binoculars are not perfectly transparent, the results presented in Figure 6 are slightly optimistic as regards maximum sighting range.

4.6 VISUAL PERCEPTION ANGLE—PROBABILITY OF DETECTION

As observed in Section 4.4, the fundamental element γ, in any study of the probability of sighting can be expressed *in terms of θ_0, which can be considered as the visual perception angle*. The equations relating γ and θ_0 are equation (7) and equation (10) for linear scan and for area scan respectively. Just as was the case with the maximum sighting range discussed in the last section, so the visual perception angle θ_0 can be obtained from equation (2) in terms of the variables which occur in the operational situation, by making the proper substitutions for visual angle α and contrast C. This leads, in the case of a target viewed normally, to an equation for θ_0 involving A, C_0, V, and R, and of a target on the sea surface, to an equation involving A, C_0, V, h, and R. If, instead of the variables listed above, the variables R/R_m, R_m/V, and C_0 are employed, the number of variables is reduced, in case I from 4 to 3, and in case II from 5 to 3. These changes of variables will now be made. Substituting in equation (2) for C from equation (12) and for α from equation (13) leads to,

for Case I,

$$C_0 e^{-3.44 R/V} = 1.75\theta_0^{\frac{1}{2}} + \frac{46.4\theta_0 R^2}{A_0}. \quad (19)$$

If $R = R_m$, equation (19) becomes

$$C_0 e^{-3.44 R_m/V} = 1.565 + \frac{37.12 R_m^2}{A_0}. \quad (20)$$

Eliminating A_0 between equations (19) and (20) gives

$$\frac{C_0 e^{-3.44(R/R_m)(R_m/V)} - 1.75\theta_0^{\frac{1}{2}}}{C_0 e^{-3.44 R_m/V} - 1.565} = 1.25\left(\frac{R}{R_m}\right)^2 \theta_0. \quad (21)$$

For Case II, a similar procedure leads to

$$\frac{C_0 e^{-3.44(R/R_m)(R_m/V)} - 1.75\theta_0^{\frac{1}{2}}}{C_0 e^{-3.44(R_m/V)} - 1.565} = 1.25\left(\frac{R}{R_m}\right)^3 \theta_0. \quad (22)$$

From these equations it follows that

$$\theta_0 = F\left(\sqrt{\frac{G}{F}+1}-1\right)^2, \tag{23}$$

where

$$F = \frac{0.49\left(\frac{R_m}{R}\right)^{2n}}{(C_0e^{-3.44(R_m/V)}-1.565)^2}$$

and

$$G = \frac{0.8\ C_0e^{-3.44(R/R_m)(R_m/V)}\ (R_m/R)^n}{C_0e^{-3.44\ R_m/V}-1.565}.$$

The constant n has the value 2 for Case I and 3 for Case II.

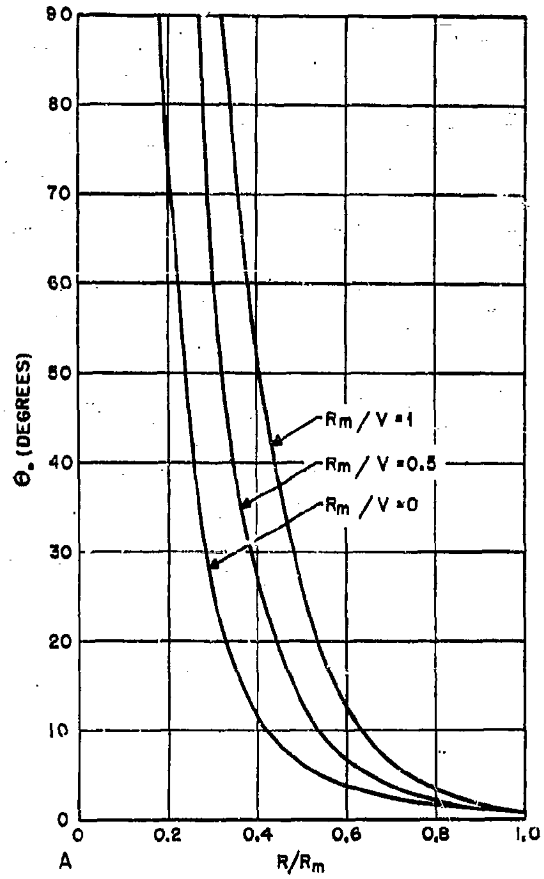


FIGURE 7A. Threshold off-axis angle as a function of range/maximum range. (Case II: C_0 = 50 per cent.)

As an example, values of θ_0 for wake-type targets are presented in Figure 7A as functions of R/R_m. It is to be remembered that a wake is a Case II target having an intrinsic contrast of 50 per cent. The three curves shown are for three different values of R_m/V, 0, 0.5, and 1. The value $R_m/V = 0$ corresponds to unlimited meteorological visibility (no haze) while $R_m/V = 1$ corresponds to haze so dense that the maximum range is determined by the haze alone.

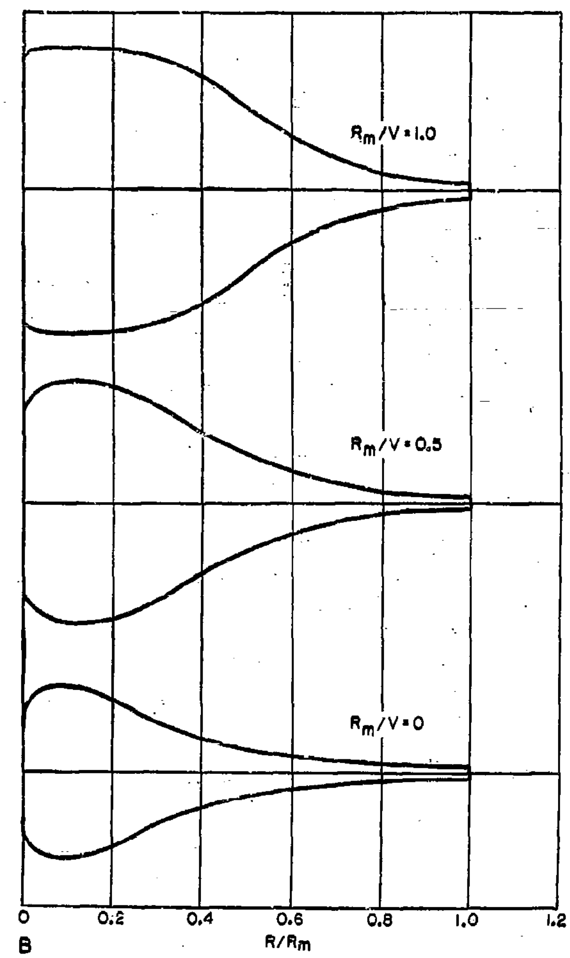


FIGURE 7B. Contours of 57 per cent detection probability.

The curves shown in Figure 7A have all been terminated at 90 degrees. This is about the maximum value of θ_0 for which vision is possible.

It is interesting to note that equation (23), as a polar equation expressing the functional relationship between θ_0 and R/R_m, describes a surface of revolution which is generated by rotation about the line of sight. Within this surface the contrast of the given target is above threshold and along the surface the probability of detection in a single fixation is constant and equal to 57 per cent. This constant probability surface of revolution can be thought of as a detection lobe analogous to the equal power surface which describes the radar detection lobe. Visual scanning, therefore, can be thought of as an attempt to so align the detection lobe as to cause the target to fall within it so that detection can take place. As examples, central sections through lobes corresponding to the three curves of Figure 7A are presented in Figure 7B.

If the type of scan employed is known so that it can be classified as line scan or area scan, then the instantaneous probability of detection γ can be obtained from the visual perception angle θ_0 by means of equation (7) or equation (10), as the case may be, with one reservation: The constant in either equation must be determined from the circumstances of the case, i.e., the number of lookouts employed and the fraction of the time during which each is searching effectively. In some cases it is possible to determine, from operational or test data, an overall constant for the particular search situation.

Thus the second fundamental question finds itself answered: *The probability that the target will be seen* has been obtained.

4.10 SUBMARINES AND SURFACE SHIPS AS SIGHTED FROM AIRCRAFT

With the material presented in earlier sections of this chapter and that in other pertinent chapters of this book, we are now prepared to analyze some actual operational situations. In this section, the sightings of submarines and surface ships from aircraft are examined to illustrate the application of the methods so far developed and to check some of the conclusions of earlier sections of this chapter. The emphasis, in this section, is on submarine sightings, since this is the situation for which the most complete operational information is available.

The sightings of submarines and surface ships from aircraft come under the Case II classification of Sections 4.8 and 4.9, since it is the wake which is sighted first in the vast majority of cases. In order to determine the maximum sighting range R_m and the visual perception angle θ_0, it is necessary to know two constants characteristic of the particular target and one characteristic of the surroundings alone. These are the intrinsic contrast C_0, the area A_0, and the meteorological visibility V. An estimate of the last is usually given in any operational sighting report.

The wake of a ship is an excellent diffuse reflector, sending back practically all the light which falls on it so that its brightness approaches that of the sky from which it is illuminated. From Section 4.7, it is clear that the sea brightness is half that of the sky unless the sea state is glassy or the line of sight is within about 10 degrees of the sun. From Section 4.6 it is seen that the effective background brightness is that of the sky, unless the line of sight comes within about 10 degrees of the sun. Hence from the definition of intrinsic contrast given in Section 4.3, C_0 is about 50 per cent except under the rare conditions of sea and sun mentioned above. These rare conditions are neglected here.

There are two methods of obtaining the target area A_0 both of which are capable of reasonably high accuracy when applied to data taken under the ideal conditions which usually obtain during service trials, and usable when applied to operational data. In the first method, the wake area is obtained from measurements of photographs taken from the air. Several submarine photographs, taken during operational sorties, were measured and the values ranged from 7,000 to 15,000 square feet, with an average of 11,000 square feet.

In the second method the area is obtained by means of equation (18) from measurements of the maximum range and the meteorological visibility. In service trials these two measurements can be made with fair precision. Under operational conditions, the ranges are distributed between zero and maximum and the meteorological visibility is only estimated, so that the required quantities must be obtained indirectly. The uncertainty concerning the meteorological visibility is overcome by considering only those sighting reports in which the effect of atmospheric haze can be neglected, i.e., those in which the meteorological visibility was estimated to be unlimited. Under these conditions equation (18) reduces to equation (16). To obtain R_0 a histogram was plotted with $R/h^{\frac{1}{3}}$ as abscissa and the number in a specified interval of the abscissa as ordinate. The maximum value which $R/h^{\frac{1}{3}}$ can have for ordinates greater than 0 is $R_0/h^{\frac{1}{3}}$. The value of $R_0/h^{\frac{1}{3}}$ obtained from the histogram of submarine sightings is 1.4, which, when substituted in equation (16), gives $A_0 = 1.3(10^4)$ square feet, the value quoted in Section 4.8. This value is believed to be more accurate than the $1.1(10^4)$ square feet obtained by the other method. The photographic method was applied to ships' wakes to obtain the values of 10^5 and $2(10^6)$ square feet for medium and large ships respectively, as quoted in Section 4.8. No great accuracy is claimed for these last two values of target area.

In order to obtain γ, the instantaneous probability of detection, it is necessary to know first the type of scan employed and hence the functional relationship between γ and the visual perception angle θ_0, and second the constant of proportionality which relates γ to the required function of θ_0. In order to determine the best method of scan, a little more information is required concerning the operational situation. An examination of the submarine sighting data shows that the altitude of flight is always small compared to the maximum range. If both are expressed in the same units, the average ratio is 0.06. Under these conditions, if fixation is at a point on the ocean distant R_m from the aircraft, then any target between this point and that directly under the aircraft is well within the visual perception angle θ_0. Indeed this is still the case if the altitude is 0.10 times the maximum range or 600 feet per nautical mile of maximum range.

The greatest chance of sighting a given target is obtained if the aircraft is made to do the scanning. This scanning is done by simply keeping the fixations, on each side of the aircraft, within a small angular distance from a point on the surface of the ocean, directly abeam and distant R_m from the aircraft. With this scheme, every target must pass through the visual perception angle θ_0 and if it stays within this angle for more than one-quarter of a second, detection by the eye is practically certain, given, of course, complete attention: missing of any targets under these conditions is due to the observer's lack of attention.

The system just outlined has been found the best possible for hunting friendly targets such as life rafts. For enemy targets, however, it is desirable to sacrifice some chance of sighting the target in order to increase the range at which it is most likely to be seen. A first approximation to the best compromise for enemy targets is a uniform scan along a line on the ocean surface distant R_m from the aircraft. It is recommended that scanning be confined to the front 180 degrees of azimuth only. The chance of sighting the target under these conditions is the same for surfaced targets as if the scanning were uniformly distributed over the entire 360 degrees, for, whereas the scanning azimuth is halved, the effort over that azimuth is doubled. Employing the first 180 degrees only, therefore, results in no loss of targets, and has the advantage of early target detection.

At the altitudes usually employed, less than 600 feet per nautical mile of maximum range, the angular departure of any target inside the maximum range circle from the scanning line is so small compared to the visual perception angle θ_0 that no great error is made by considering all targets as being situated on the scanning line. Hence the problem under consideration is one of pure line scan with γ proportional to θ_0.

To obtain the constant of proportionality between γ and θ_0 it is necessary to resort to the operational data. The quantity which can be obtained from these data is the sweep width W, defined in Chapter 2 as

$$W = 2\int_0^\infty p_x dx, \qquad (24)$$

where p_x is the probability of detecting a target distant x from the aircraft track. To do this the 529 submarine sightings on which reports were available were divided into groups representing ranges of R_m. For each group a lateral range distribution curve was plotted and normalized to unity in the center, (i.e., at $x = 0$). This normalization to unity in the center is equivalent to saying that it is virtually impossible to fly directly over a target as large as a surfaced submarine without seeing it—in spite of such rare occurrences as that mentioned in Section 2.2. There is good evidence in support of this view.[7] The area under the lateral range distribution curve so normalized is $\int_0^\infty p_x dx$ which determines W. The values of W obtained from the operational data are presented

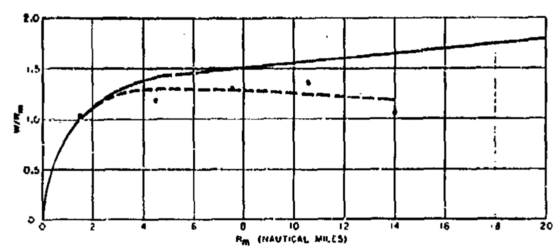


FIGURE 8. Path swept/maximum range as a function of maximum range—comparison of computed (solid line) and operational (dotted line) results (surfaced submarines).

in Figure 8 as a function of R_m. These values are in units of R_m and the dots are the operational points.

In order to compute W, it is to be recalled that, from Chapter 2,

$$p_x = 1 - e^{-\int_0^\infty \gamma dt}, \tag{25}$$

and equation (7) $\gamma = k\theta_0$. The constant k for the operational situation under consideration is as yet unknown since it depends upon the number of lookouts, the fraction of lookout time spent in search, and the degree to which the search approaches uniformity over the scanning line considered. If y is the distance, in nautical miles, of the target from the point of nearest approach to the aircraft and v is the aircraft velocity in knots, $dt = 3{,}600\, dy/v$ in seconds, so that

$$\int_0^\infty \gamma dt = k \int_0^\infty \theta_0 dt = \frac{3{,}600}{v} k \int_0^\infty \theta_0 dy. \tag{26}$$

Expressing dy in units of R_m,

$$\int_0^\infty \gamma dt = \frac{3{,}600\, R_m}{v} k \int_0^\infty \theta_0 d\frac{y}{R_m}. \tag{27}$$

The integration indicated in equation (27) has been carried out graphically for various values of x/R_m. These are presented in Figure 9 as functions of x/R_m. The two curves shown are for $R_m/V = 0.5$ and $R_m/V = 1$, respectively. The values of θ_0 employed for the integrations were obtained from Figure 7A. In the operational data $R_m/V = 1$ for values of $R_m \leqq 5$ nautical miles and 0.5 for $R_m \geqq 10$ nautical miles. The aircraft velocity for the 529 incidents averaged 135 knots.

To obtain k, the following procedure was employed: For $R_m = 5$, $R_m/V = 1$, and a given value of k, $\int_0^\infty \gamma dt$ was computed from equation (27) and the integral given in Figure 9 for each of a number of values of x/R_m. For each of these, p_x was computed from equation (25) and a lateral range curve was plotted. From the lateral range curve a value of W/R_m was obtained by graphical integration. This procedure was repeated for a number of values of k until one was found which fitted the operational data of Figure 8. The value of k obtained in this way is 1.75×10^{-3}.

Various lateral range curves, computed using $k = 1.75 \times 10^{-3}$ and $v = 135$ knots, are presented in Figure 10. The quantity W/R_m obtained from these curves is presented in Figure 8 as a function of R_m. An examination of Figure 8 shows good agreement between the computed and the operational points at low values of R_m and a gradual separation of the

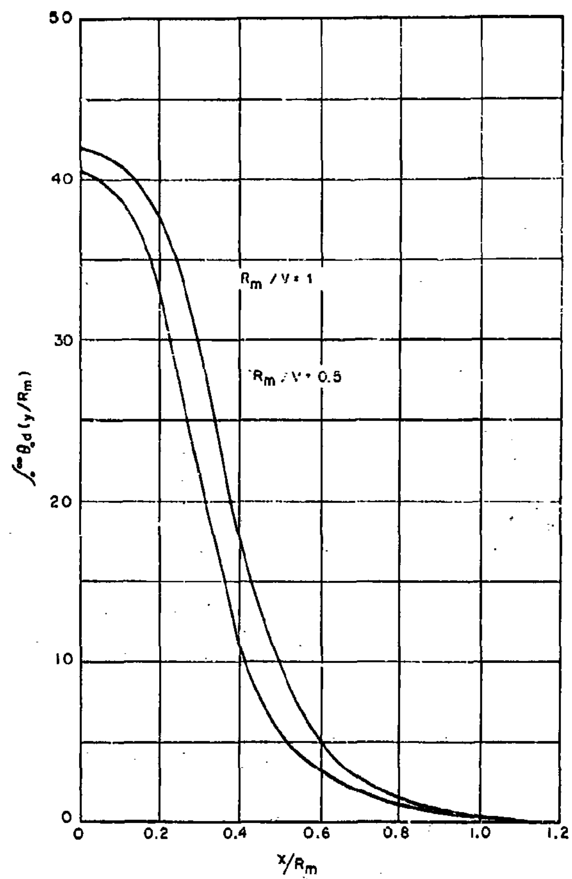


FIGURE 9. Integral of equation (27) as a function of lateral range/maximum range.

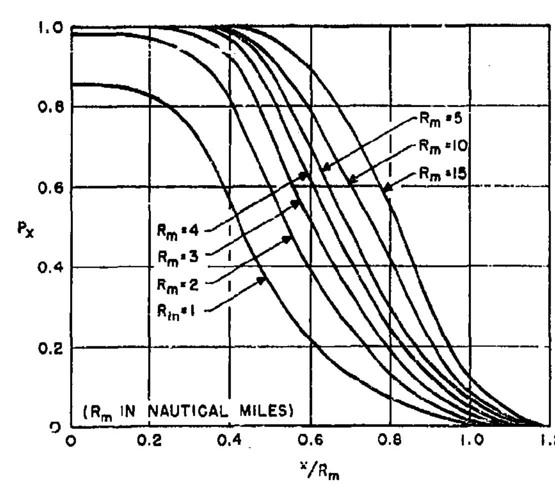


FIGURE 10. Lateral range distribution as a function of lateral range/maximum range, assuming $k = 1.75 \times 10^{-3}$, $v = 135$ knots (surfaced submarine).

two as R_m increases. Now the theoretical curve is for ships which remain surfaced, while the experimental curve is for submarines which can evade detection by diving. This effect has been investigated[7] and it has been shown that the effect of this evasion on the path swept W is negligible for values of $R_m \leqq 5$ nautical miles and gradually increases to about 35 per cent for $R_m = 10$. This tendency is apparent in Figure 8.

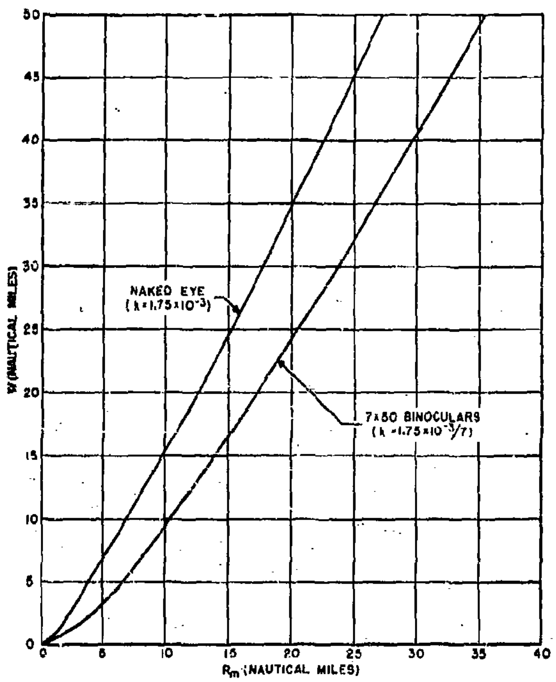


FIGURE 11. Path swept as a function of maximum range, assuming v = 150 knots.

In Figure 11, W for 50 per cent intrinsic contrast targets is presented as a function of R_m for naked-eye search. The more modern figure of 150 knots has been taken as the aircraft velocity. To obtain W in any particular case, R_m is first obtained from Figure 5 and then W for the given R_m is obtained from Figure 11. It is to be *remembered that altitudes greater than 600 feet per nautical mile of maximum range are not recommended.*

The greater values of R_m found in Figure 6 for 7x50 binocular search might, at first sight, lead to the belief that W should be greater for binocular than for naked-eye search. It is to be remembered, however, that inside a binocular field of magnifying power 7, the eye must search a scanning line 7 times as long. Hence for the same search effort, k is 1/7 as great. The quantity W, computed using $k = 1.75 \times 10^{-3}/7$, is also presented in Figure 11 as a function of R_m. An examination of the search situation, using Figures 5, 6, and 11, shows that binoculars, even under the best of conditions, are not as effective as the naked eye if the meteorological visibility is less than 10 nautical miles.

As examples of the results to be obtained from Figures 5 and 11, values of W for various altitudes and meteorological visibilities are presented in Table 1 for the three targets considered in Figure 5. To compute such tables the following procedure is em-

TABLE 1. Sweep width W in nautical miles for naked-eye search.

Altitude (feet)	Meteorological Visibility (nautical miles) 1	3	5	10	15	20	30	40	50
				Submarines or small merchant vessels					
500	0.9	3.2	4.3	7.5	8.6	9.6	11	12	13
1,000	0.9	3.7	5.3	8.5	11	12	14	15	16
2,000	...	3.7	5.9	9.6	12	14	17	18	19
3,000	...	...	6.4	11	13	15	18	20	21
5,000	...	...	...	12	15	17	20	22	25
				Medium-sized ships					
500	0.9	3.7	5.3	12	15	17	20	22	25
1,000	0.9	3.7	6.4	13	16	19	25	27	29
2,000	...	3.7	7.0	14	18	22	28	31	34
3,000	...	...	...	14	19	24	30	34	37
5,000	...	...	...	14	21	26	33	37	42
7,000	...	...	...	...	21	27	34	40	45
				Large combatants and high-speed liners					
500	0.9	3.7	6.4	13	16	19	25	27	29
1,000	0.9	3.7	6.4	14	18	22	28	31	34
2,000	...	3.7	7.0	15	19	25	31	36	37
3,000	...	...	7.0	15	21	26	34	40	43
5,000	...	...	...	15	22	28	36	43	48
7,000	...	...	...	...	22	29	38	45	50
10,000	...	...	...	...	...	30	40	47	53

ployed. For each pair of values of V and h, $V/h^{\frac{1}{2}}$ is computed and the corresponding value of $R_m/h^{\frac{1}{2}}$ is obtained from Figure 5, from which R_m is computed. Knowing R_m, W is obtained from Figure 11.

The value of k obtained from the operational data provides means of determining the average effectiveness of search from aircraft. The value of k obtained from the Craik scanning experiments described in Section 4.4 was 2.73×10^{-2}. Since this is for scan over 45 degrees instead of 180 degrees it must be divided by four for comparison with the operational value of 1.75×10^{-3}. This division gives 6.8×10^{-3}, about four times the operational value. It is clear, therefore, that the entire crew of an aircraft is about equivalent to one-quarter of an ideal lookout carrying out search under laboratory conditions. A number of reasons for this discrepancy can be enumerated. These are soiled or scratched windows, fatigue, interruption from search caused by other duties, inability to search along an accurately determined best scanning line, and nonuniform angular distribution of scanning effort. Soiled or scratched windows tend to reduce the apparent contrast and hence both R_m and θ_0. The need for keeping the windows clean and clear can not be overemphasized. On long sorties, fatigue is inevitable. However, its effects can be minimized by relieving monotony. This can be done by making frequent exchanges in station for the various lookouts. The best scanning line, (i.e., the locus of points on the ocean surface distant R_m from the aircraft) is usually 3 or 4 degrees below the horizon. A rough and ready rule for finding this locus is to extend the fist at arm's length and look about two or three fingers below the horizon. There is good reason to believe that the scanning effort, instead of being uniformly distributed over the forward 180 degrees is heavily weighted in the front 45 or 90 degrees. One suggestion for overcoming this tendency is to assign all lookouts, other than pilot and copilot, to the two 45 degree sectors just forward of the beams, leaving the forward 90 degrees to the pilot and copilot.

Chapter 5

RADAR DETECTION

5.1 INTRODUCTION

The immense value and versatility of the radar-echo principle in its military applications were amply demonstrated during World War II. The ability of radar to provide precise values of the range and bearing of objects on or above the surface of the sea, under all conditions of visibility, and frequently at distances considerably beyond the range of the human eye, assured its constant naval use as an instrument of search and early warning. In addition, it readily proved its usefulness in a number of related applicabilities. These included fire control, identification, altimetry, and aid to bombing, interception, fighter direction, station keeping, and navigation.

Nevertheless, airborne and shipborne search radar, used both offensively and defensively to gain contact with enemy forces, and to locate missing friendly units, represented perhaps the most widespread and successful of these military applications, and may be expected to continue to do so in the immediate future. It is with this aspect of radar that the present chapter deals. Emphasis is placed on basic search considerations; only those technical questions which have bearing on this subject are discussed.

5.2 MODERN SEARCH RADAR CHARACTERISTICS

The type of radar currently most useful in sea search is the airborne microwave (wavelength $\lambda \leqq$ 10 cm) search radar. A brief outline of the principles of operation of such equipment will serve in a general way to illustrate those of similar sets, including shipborne search gear.

High-frequency radio energy generated in the transmitter of such a radar is led through a wave guide (a resonant copper pipe of rectangular cross section) to a funnel-shaped horn, or to a dipole radiator, located at the focus of a paraboloidal metal reflector. The energy is re-radiated from this reflector in a lobe-shaped pattern, as indicated cross sectionally in Figure 1. The beam width—conventionally defined as the angle θ between half-power directions—is determined by the size of the paraboloid relative to λ. In practice subsidiary back and side lobes (not shown in the figure) are produced in addition to the main lobe; but these can be minimized by proper antenna design—i.e., by modification of the reflector shape, addition of parasitic radiators, etc.

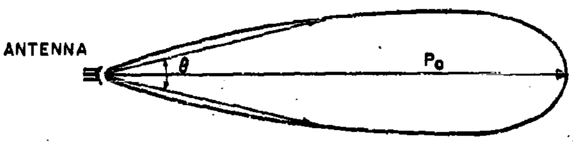


Figure 1. Idealized pattern of a microwave search radar. Length of radius vector is proportional to power radiated in that direction.

Actually, in most search equipment, the paraboloidal reflector is truncated or otherwise modified in such a way as to produce a narrow antenna pattern in the horizontal plane (sometimes as narrow as 1 degree) and a relatively broad pattern (usually about 10 degrees) in a vertical plane. This provides extended coverage in altitude and lessened sensitivity to antenna tilt, features particularly desirable for airborne early warning and shipborne aircraft warning radar.

In present radar sets, high-frequency energy is never emitted continuously; instead, it is transmitted in successive high-power pulses of very short duration. In order to decrease the minimum radar range, and to improve range resolution, as well as to permit the use of higher peak transmitted power, it is desirable to reduce the duration of such pulses to a minimum. However, against these considerations must be balanced the fact that a very sharp pulse wave form can be reproduced accurately only by a broad-band receiver; and an increase in bandwidth results in an increase of receiver noise level relative to echo strength, with a corresponding reduction in maximum radar range. Pulse durations used in practice are generally between 0.1 and 2 microseconds, with 1 microsecond a common value. The rate at which such pulses are repeated is usually between 400 and 1,600 pulses per second in current search equipment. Such repetition rates allow a sufficient time interval for energy in one transmitted pulse to travel out to a distant target and be reflected back to the radar before the next pulse is emitted. In search for very far distant targets, lower pulse repetition rates may be necessary.

Signals reflected from a target are picked up by the same antenna used to transmit them; they are then amplified in the radar receiver and presented as a response on an indicator. The kind of presentation most frequently employed in modern search radars is that of the *plan position indicator* (PPI scope). In this type of "intensity modulated" indication, response is obtained on the face of a cathode-ray tube by means of variations in the intensity of a radially sweeping electron beam. At the instant an energy pulse is transmitted from the radar antenna, this beam begins to sweep out at a uniform rate from the center of the indicator. When the reflected pulse from a target returns to the radar, the amplified energy is used to increase the volume of the electron stream impinging on the scope. This causes a bright spot, known as a "blip," to appear on the fluorescent screen at a radial distance proportional to the elapsed time, and therefore to the range of the target. The persistence of the screen is sufficiently great that such an intensified spot usually remains visible for several seconds. When the electron beam has swept to the outer edge of the scope, corresponding to maximum range obtainable on the particular range scale employed, it reverts almost instantaneously to the center of the scope, and another pulse is emitted. The angular position (azimuth) of the sweep trace on the scope is determined by the direction in which the antenna is momentarily pointing.

FIGURE 2. Plan position indicator (PPI) presentation. Plane bearing radar was at an altitude of 20,000 feet, directly over Boston. Cape Cod may be clearly discerned near the center of the photograph.

In order to obtain area coverage, the radar antenna is rotated about a vertical axis. Various radars permit manual control, sector scan, or 360 degree scan. The last is most often employed in airborne search, the antenna being rotated at rates usually between 5 and 24 revolutions per minute (12 rpm is a common value). The use of this type of scan in conjunction with PPI indication, in which the radial sweep trace of the electron beam is made to rotate in synchronism with the antenna, results in continuous presentation of a plan map of the region surrounding the radar, as illustrated in Figure 2. Maximum scanning rate is limited by the necessity of receiving several successive pulses to produce a noticeable blip (estimates of this minimum number of pulses vary from 4 to 10, depending on the type of screen used). It is thus related to the antenna beam width and the pulse repetition rate, which in turn depends on pulse duration and the maximum average power the transmitter tubes can handle; scanning rate is therefore related indirectly to target resolution. In present search equipment, between 10 and 100 pulses reach the target per scan, so that the maximum possible scanning rate is not generally attained. It may be noted that with rapidly scanning, narrow-beam antennas some "scanning loss" may be experienced, owing to a turning of the antenna during the finite echoing time. The effect is usually small in practice.

Other types of antenna scan and visual presentation than those described above are occasionally employed in search equipment. Narrow-beam antennas, for instance, may be scanned helically, or rocked; some of the newer airborne sets have their antennas gyro-stabilized in order to prevent distortion of the PPI map due to antenna tilt. An alternative method of intensity modulated presentation sometimes encountered is that of the B scope, in which range is measured along a vertical and bearing along a horizontal axis; this necessarily results in some distortion, but provides greater resolution of nearby targets. It usually is combined with an antenna scan of 180 degrees or less in the forward direction.

5.3 VISUAL AND RADAR SEARCH

Having outlined the basic operative features of current search radar equipment, we shall find it instructive at this point to compare the process of visual search, dealt with extensively in the previous chapter, with that of radar search.

It will be recalled that the eye, when searching systematically, tends to look in one direction for a short period of time (of the order of one second), during which several fixations occur. It then skips to a new line of sight, frequently differing in direction from the previous line by as much as 10 degrees. Since, during any single fixation, the eye can resolve distant objects only within an arc of about 1 degree, the distant coverage pattern for visual search tends to be ragged; there is a considerable probability that small objects at long ranges will be passed over. At shorter ranges, on the other hand, there is a broad lobe of peripheral vision in which prominent objects off the direct line of sight are readily detected (see Figure 7B, Chapter 4).

In contrast, radar scans continuously, without gaps in its coverage, and does so over a considerable range, at a rate of scanning usually considerably greater than that of the eye. A minimum radar range limitation is imposed in airborne search by the shape of the vertical antenna pattern and by the antenna tilt setting. Furthermore, there will in general be a near-by "sea return" area, in which transmitted energy is reflected back to the radar from waves, with the result that an irregular bright patch is presented in the center of the PPI scope (see Figures 2 and 3). The extent of this patch increases with the

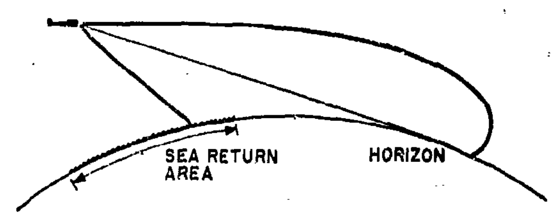


FIGURE 3. Schematic illustration of vertical antenna pattern and sea return area for airborne early warning radar.

altitude of the antenna and the roughness of the sea. Targets within the sea return area cannot readily be detected in most cases, although special circuits in newer sets offer some improvement in discrimination. Maximum radar range is generally limited (approximately) by the horizon, although under certain conditions of energy propagation considerably greater ranges may be obtained. This effect will be discussed in Section 5.4.

The types of search coverage obtained by the eye and radar are illustrated qualitatively, by means of contours of constant detection probability, in Figures 4 and 5. It should be understood that only the general shape of these patterns, presented for comparison, is significant. These diagrams serve to illustrate the fact that radar and visual search, as regards coverage, are to a large extent complementary. This is true in many other respects, too. While the eye, for instance, is not directly capable of exact range determination, the accuracy of electronic timing circuits makes radar well adapted for this purpose; on

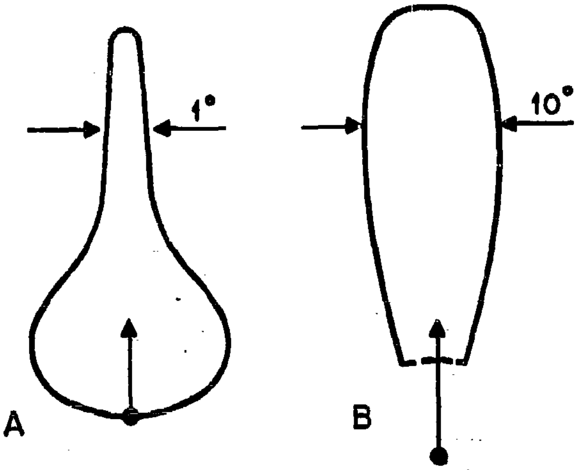


FIGURE 4. Schematic horizontal coverage patterns for (A) the eye, and (B) radar. Contours of constant probability of detection are shown.

the other hand, the eye possesses obvious superiority in target resolution. We may also note that, since radar wavelengths are large in comparison with the dimensions of smoke, dust, and water particles in the atmosphere, radar radiation (except that of extremely high frequency) is only slightly affected by

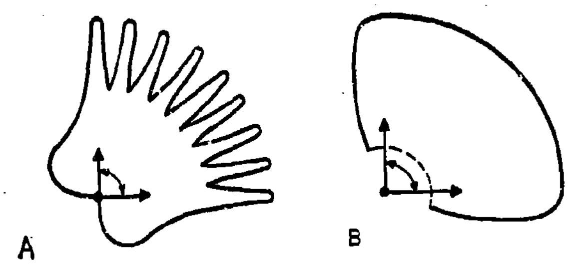


FIGURE 5. Coverage obtained in scanning through 90° for (A) the eye, and (B) radar.

such obstacles, whereas visible radiation is strongly attenuated or reflected by them.

In the visual case, fixation on a target has been regarded as making detection certain; the element of chance is introduced by the uncertainty of obtaining a fixation. It is an important differentiating characteristic of search radar that, even with a well-tuned set scanning uniformly over a (small) target within easy range (a case analogous to assured fixation) blips may not be returned on every scan. This effect, largely a consequence of varying target aspect, will be discussed at greater length in Section 5.5. The point of chief importance to the present discussion is that in both cases detection is a matter of uncertainty. Clearly, the probability methods developed in Chapter 2 in the treatment of the general theory of detection will have applications in the solving of radar detection problems. We shall investigate the mathematical formulation of some of these in Section 5.6.

5.4 PROPAGATION

Radar energy is propagated in free space according to an inverse square variation with distance. That is, if P_t is the power transmitted in a pulse, G_t the power gain of the antenna for transmission (a function of its shape and size), and P_0' the power density at distance r along the axis of the antenna,

$$P_0' = \frac{P_t G_t}{4\pi r^2}. \tag{1}$$

When the energy strikes a target at range r, a certain portion of it is scattered, that is, diffusely reflected (re-radiated). The scattering ability of the target is measured by a quantity σ, the "effective radar cross section," defined as the ratio of total power reflected from the target to incident power density impinging from the direction of the radar. The amount of power re-radiated from the target is therefore $P_0'\sigma$. This power is likewise attenuated according to the inverse square law; its density at the radar antenna is

$$P_r' = \frac{P_0'\sigma}{4\pi r^2} = \frac{P_t G_t \sigma}{(4\pi)^2 r^4}. \tag{2}$$

The total power received from the reflected pulse is the product of P_r' and G_r, the gain of the antenna for reception, i.e.,

$$P_r = P_r' G_r = \frac{P_t G_t G_r \sigma}{(4\pi)^2 r^4}. \tag{3}$$

In case the antenna is a paraboloid of aperture area A, it can be shown that the above expression reduces to

$$P_r = \frac{P_t A^2 \sigma}{4\pi r^4 \lambda^2}. \tag{4}$$

If P_{min} is the minimum value of reflected power that can be amplified by the radar receiver to furnish a recognizable blip, we find, solving equation (4) for r (with $P_r = P_{min}$), that the maximum range of detection of a target of effective cross section σ is

$$r_{max} = \sqrt[4]{\frac{G_t G_r \sigma}{(4\pi)^2} \cdot \frac{P_t}{P_{min}}}. \qquad (5)$$

Although these formulas have been developed for the case of propagation in free space, it is found that they may be employed to a reasonable degree of approximation under many conditions encountered in practice. For example, in the case of sea search by airborne radar at common altitudes (such that the effects of interference with power reflected from the sea are of secondary importance), the last formula provides at least a qualitative indication of the influence of various factors on maximum radar range. In particular, the relative insensitivity of r_{max} to large variations in transmitted power is shown: owing to the fourth root relationship, an increase in transmitted power of sixteen fold is necessary to double the maximum range. (In this connection, it might be pointed out that in Figure 1 of Section 5.2 of this chapter the reflected signal has sufficient strength to give good *range* over a wider beam than the power pattern would indicate. Thus, at the half-power

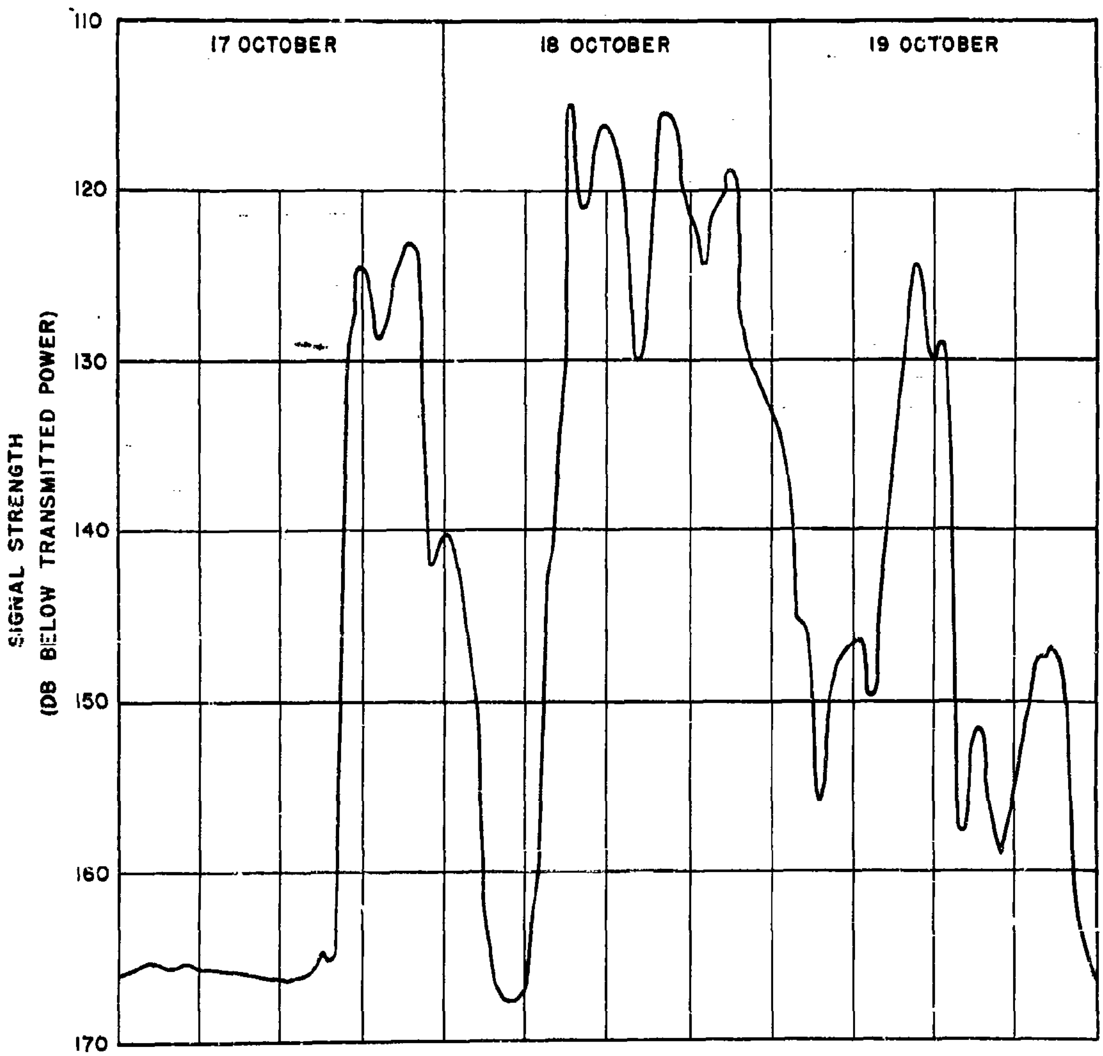


FIGURE 6. Variation of echo amplitude with constant radar performance. Target 30 miles from radar, with path of pulses passing over sea.

points, maximum range is reduced by a factor of only $1/\sqrt[4]{2} = 0.84$.)

Near the surface of the sea, a different type of propagation is commonly observed. Energy that strikes the surface at a small angle is reflected, as from a good conductor, undergoing a 180 degree change of phase. Over a calm ocean, the alternate reinforcements and cancellations of radar energy following the two possible paths, direct and reflected, cause the antenna pattern to become stratified. With longer wave radars, the null regions thus produced cause a pronounced recurrent fading of echoes from airborne targets. But for the microwave sets in increasingly common use in search, the stratified lobes are so closely spaced (spacing less than 1 meter) that such fading ceases to be a problem. Under these conditions of transmission near a calm sea, the echo power received from a target also near the sea can be shown to vary according to an inverse eighth power of range, rather than an inverse fourth power. Near the surface of a rough sea, some intermediate law may, in effect, be more closely followed.

As previously mentioned, maximum radar range, except in the case of very small targets, usually is not limited as a result of attenuation suffered by the radiated energy (for, although this attenuation is, as we have seen, considerable, correspondingly large amounts of power can be transmitted in the radar pulses), but is limited by the horizon. Actually, owing to refraction in the earth's atmosphere, a portion of the energy is bent around somewhat beyond the horizon. It has been found that this effect can be taken into account approximately by computing maximum range as horizon range for a fictitious earth of radius four-thirds that of the real one. A convenient formula based on this assumption is $R = 1.25 (\sqrt{h_r} + \sqrt{h_t})$, where R is the maximum radar range in nautical miles, h_r is the radar altitude in feet, and h_t is the target altitude in feet.

Under certain meteorological conditions, generally associated with inversion of the normal temperature or humidity gradients, abnormally long radar ranges may be observed. The effect, particularly noticeable if both radar and target are close to the surface of the sea, is much the same as if a portion of the radiated energy were trapped beneath the inversion level. Apparently, trapping of some energy in a surface duct does not in general interfere with propagation at higher altitudes, since abnormally long detection ranges near the surface are frequently observed to be accompanied by good ranges from surface to air, or vice versa. In extreme cases, the existence of anomalous conditions, both of abnormal and of subnormal propagation, may lead to very pronounced and rather rapid variations in echo strength, as indicated, for example, in Figure 6. (Note that on one of the days covered in this chart a variation in received power of 53 db, or 200,000-fold, in a period of 2½ hours is recorded.) In certain geographical regions, notably, the eastern seaboard of the United States, conditions of "anomalous" propagation are more or less prevalent. In most localities, however, large or rapid fluctuations in propagation conditions are not generally to be expected.

5.5 BLIP-SCAN RATIO

It was mentioned in Section 5.3 that when a radar scans across a target, particularly near the limit of its range capabilities for that target, it is the general experience that a blip is not presented on the radar indicator on each scan. In order to characterize analytically the behavior of a given radar with respect to a specified target, it is convenient to introduce the concept of "blip-scan ratio." This ratio, which we shall denote by $\psi(r)$, is defined as the proportion of scans, upon a target at range r, during which a recognizable signal is presented on the PPI scope. It therefore represents the probability that a single scan will produce an effective blip, i.e., a blip which is actually recognized by an operator who is focusing his attention on the part of the scope where it appears: a "recognizable" blip may actually fail to be recognized by an operator whose attention is lagging, or is directed to another part of the scope where objects of interest are seen; this effect of operator fallibility will be considered later. In other words, we are separating the study of the uncertainties (probability) of radar detection into two parts: the question of the probability ψ of detection of the blip when the operator is concentrating on the part of the scope where it occurs, and the matter of how likely he is to be so concentrating. While this separation is somewhat unrealistic (an intense blip is apt to be seen out of the corner of the eye, a faint but "recognizable" one is not), it affords a convenient simplification and will be made the basis of the present treatment.

Clearly, the value of the blip-scan ratio is dependent on a number of long-term variables that may in an approximate treatment be regarded as constant during a particular search, or at least during long parts thereof. These include type of target, conditions of propagation, sea state, direction of search with respect to that of the wind, radar altitude, antenna tilt, level of operator and set performance, and radar characteristics such as wavelength, scanning rate, etc. In addition, two basic short-term variables are involved—range and target aspect. As we shall see, the aspects of naval targets vary characteristically in short-time cycles; it is this fact, represented by appropriate mathematical assumptions, that permits the treatment of ψ as a specific function of range only. Before formulating these assumptions, however, let us investigate briefly the general subject of radar echo fluctuations.

The components of radar targets which are most effective in reflecting energy are flat surfaces (normal to the axis of the radar beam) and internal rectangular corners. Metallic conductors are always more effective than nonconducting materials. Since a flat surface reflects radiation specularly, its orientation must be within a few degrees of normality to the beam direction to return an appreciable signal; otherwise, most of the incident energy is shunted off into space. The corner reflector, on the other hand, has the property for microwave radar energy, as for light, of reflecting radiation along the direction of incidence, over a wide range of angular aspect. It is thus relatively insensitive to momentary aspect and movement.

We may accordingly in a general way distinguish two types of naval targets. The first type, which we shall call Class A, represented by large vessels, including warships and merchant ships, and in most cases by beam-aspect surfaced submarines, is characterized by the prevalence of flat reflecting surfaces and rectangular corners and brackets. Energy reflections from the many components of such targets reinforce or cancel in accordance with their various momentary phase relationships. When micro-wavelengths are employed, very slight target movements cause radical alterations in these phase relationships, with consequent rapid variations in echo strength. In the case of naval targets, such movements may be represented by ship roll and pitch, or even by physical distortion of the vessel in a seaway, bending of the masts in a wind, etc. It is found for targets of this type that the echo fluctuations are rapid, even in comparison with the short time interval required for the beam of a searching radar to sweep across them. Consequently, on any particular scan, although large momentary (pulse-to-pulse) echo variations occur, an average signal strength, resulting from summation of the pulses, is presented on the radar indicator at a roughly constant level. Since this average level does not vary greatly from scan to scan, but increases steadily with decreasing target range, the blip-scan ratio for targets of Class A usually increases from zero to unity over a rather short range interval. It is apparent that this corresponds approximately to a "definite range law," the definite range being that at which the average signal first becomes perceptible. The effect is illustrated in Figure 7.

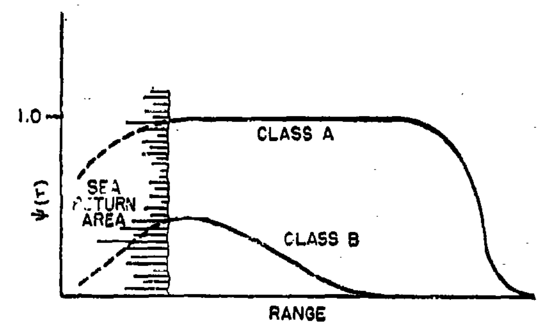


FIGURE 7. Blip-scan ratio for Class A and Class B targets.

The second type of radar target, which we shall call Class B, consists of relatively smooth, continuously curved surfaces, having few sharp corners or angles. Examples are surfaced submarines in bow or stern-aspect, periscopes, and schnorchels (cylindrical submarine "breathing" mechanisms). The energy return from such targets also varies with their momentary aspect, but the fluctuations are much slower than those observed for Class A targets. Indeed, the rate of variation in this case, being usually of the order of magnitude of the scanning rate, is sufficiently slow that the energy return per pulse does not, in general, change radically during the time the radar beam is sweeping across the target. From scan to scan, however, there is wide variation in signal intensity. As illustrated in Figure 7, the blip-scan ratio for such small, smoothly shaped targets is in no way suggestive of a definite range law.

Although most targets fall into one or the other of the classifications described above, a few, such as intermediate-aspect surfaced submarines, cannot readily be regarded as either large, complex targets or small, simple ones. It is convenient therefore to extend our definitions of these classes: Class A will include those targets for which, to the desired degree of approximation (with specified values of the long-term variables) a definite range law can be defined; Class B will include all others. Methods of dealing with the definite range situation, including averaging for distributions of parameters, may best be illustrated in connection with sonar search, which forms the subject of Chapter 6. For the remainder of the present study, therefore, we shall confine ourselves largely to the study of Class B targets.

We are now ready to formulate one of the assumptions, previously mentioned, regarding variations of target aspect. We shall assume that they are of such a nature that the blips returned from a target at a certain range are distributed *at random* among the radar scans (with a relative frequency corresponding to the blip-scan ratio at that range), in other words, that $\psi(r)$, the probability that a blip be presented on a particular scan, is independent of what may be known about the results of previous scans. This assumption, as we have shown, represents a reasonable approximation to operational facts, for most targets and for common scanning rates. It greatly simplifies the mathematical treatment of radar detection problems, to be dealt with in the following section.

The effect of changes in general target aspect, resulting from relative travel, has been found to be of secondary importance. For symmetrical targets such as schnorchel, this is obviously the case. And although, for other targets, unusually long ranges can be obtained at certain general aspects, the angular regions, in the horizontal plane, over which this is true are found in practice to be relatively small. (The situation is indicated schematically in Figure 8, for several types of targets.) We are therefore justified, in most cases, in assuming that these angular regions may be ignored and that the blip-scan ratio may be regarded as approximately independent of general horizontal aspect.

The problem of predicting blip-scan ratio from theory would be found (if it had to be dealt with) to be a formidable one, indeed. Not only would it be necessary to make computations of effective radar cross section σ—dependent on the size, shape, and materials of the target, and on the wavelength and polarization of the radar radiation—but, in addition, assumptions as to the distribution of values of σ with respect to time would be required. A theory involving such variables and assumptions would necessarily tend to become overcomplicated and artificial. (Nevertheless, efforts made to show with sufficiently simple geometrical figures, such as cylinders, that the blip-scan ratio should vary with sea state in the observed manner have met with some success; their extension to more complicated figures, however, would be prohibitively difficult.) It is fortunate, therefore, that we are not forced to rely upon theory for our knowledge of blip-scan ratio. *Test data* are available, in considerable abundance, which provide reasonably accurate values of $\psi(r)$ for specified sets

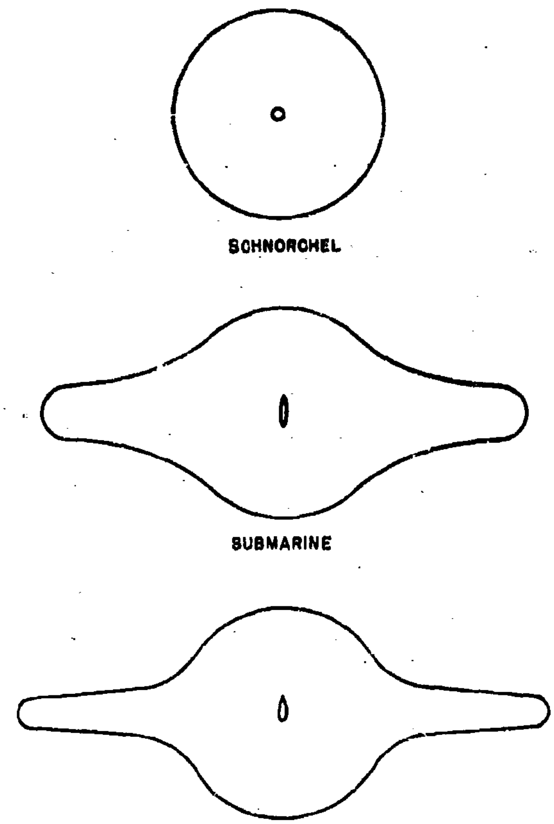


FIGURE 8. Typical polar diagrams, showing contours of constant detection probability as a function of general horizontal aspect for various kinds of targets.

of values of the long-term variables, and, as a practical matter, it is this availability that makes the concept of blip-scan ratio useful.

The effect on $\psi(r)$ of one long-term variable, the type of target, has been indicated in Figure 7. In Figure 9 the effects of two others, sea state and direction of search relative to that of the wind (upwind and downwind search) are illustrated for AN/APS-15A radar used against schnorchel target. The operational significance of variations in these and other basic parameters may perhaps be

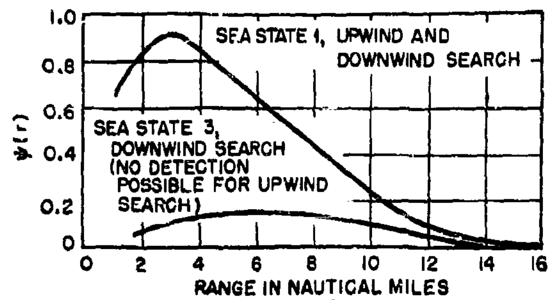


FIGURE 9. Effect of sea state and wind direction on blip-scan ratio for AN/APS-15A, altitude 500 feet, schnorchel target (experimental data from ASDevLant Project 547).

better understood in terms of their influence on radar search width, to be discussed in Sections 5.6 and 5.7.

5.6 RANGE DISTRIBUTIONS AND SEARCH WIDTH

We shall now study the application to radar detection problems of the methods of analysis outlined in general terms in Chapter 2 for the determination of sighting range distributions and search width. The coordinate system (Figure 7, Chapter 2) and much of the notation previously employed will be retained; in particular, it will be recalled that w is the relative speed of target and searching craft, T the glimpsing (radar scanning) period and $g(\sqrt{x^2 + y^2})$ the "glimpse" probability of detecting on a particular scan a previously undetected target located in the neighborhood of the point (x,y).

In Section 2.5 it was shown that the probability of first detecting in the area $dxdy$ a target that has moved along in relative space parallel to the y axis to the neighborhood of the point (x,y) is

$$e^{-F(x,y)}\, g(r) \frac{dy}{wT}; \tag{6}$$

and that the average detection rate in $dxdy$ for unit target density is

$$\rho(x,y) = e^{-F(x,y)} \frac{g(r)}{T}. \tag{7}$$

In both cases,

$$F(x,y) = -\sum_{i=1}^{\infty} \log\left[1 - g\left(\sqrt{x^2 + (y - iwT)^2}\right)\right], \tag{8}$$

the summation extending to infinity because the relative track is assumed to be of indefinitely great extent in the direction toward which the searcher is traveling. The distribution of true ranges of first detection is expressed [see equation (35), Chapter 2] by

$$\rho(r) = \int_{\pi-\theta}^{\pi+\theta} r\rho(r \sin \zeta, r \cos \zeta)\, d\zeta, \tag{9}$$

the integration extending over the angular scanning range, assumed symmetrical about the y axis (θ represents the radian measure of half this range; for all-around scanning $\theta = \pi$). The function $\rho(r)$ may be alternatively regarded as the rate of detection in a unit range interval at range r for unit target density. If, as is the case in certain operational tests, the searching craft makes its approach to the target on a radial course, the sighting range distribution is expressed by equation (7) for lateral range zero, i.e.,

$$\rho(r) = e^{-F(0,r)} \frac{g(r)}{T}. \tag{10}$$

The quantity wT that appears in equation (8) represents the distance traveled by the target along its relative track between successive radar scans. Since it is *small* in most cases (wT = 0.21 nautical miles, for a 150-knot searching aircraft, stationary target, and 12 rpm scanning rate), an important simplification of the expression for $F(x,y)$ can be made (the summation can be replaced by integration). Thus,

$$F(x,y) = -\frac{1}{wT}\int_{-\infty}^{y} \log\left[1 - g\left(\sqrt{x^2 + y^2}\right)\right]dy, \tag{11}$$

the upper limit of integration lying in the scanned range of $y = x \cot \zeta$, for ζ in the range $(\pi - \theta, \pi + \theta)$.

Writing $\Gamma(r)$ for the quantity $-\log\,[1 - g(r)]$, equation (9) becomes

$$\rho(r) = \frac{1}{T}\int_{\pi-\theta}^{\pi+\theta} rg(r) \exp\left[-\frac{1}{wT}\int_{-\infty}^{y} \Gamma(r)dy\right]d\zeta. \tag{12}$$

By analogy with the corresponding equations of Chapter 2, the distribution of lateral ranges of first sighting is given by

$$p(x) = 1 - \exp\left[-\frac{1}{wT}\int_{-\infty}^{|x| \cot (\pi-\theta)} \Gamma(r)dy\right]; \tag{13}$$

the radar search width W, corresponding to the area under the lateral range curve, is

$$W = \int_{-\infty}^{\infty}\left\{1 - \exp\left[-\frac{1}{wT}\int_{-\infty}^{|x| \cot (\pi-\theta)} \Gamma(r)dy\right]\right\}dx; \tag{14}$$

and the average detection range is

$$\bar{r} = \frac{1}{wW}\int_0^\infty r\rho(r)dr. \quad (15)$$

$\Gamma(r)$ can readily be expanded,

$$\Gamma(r) = -\log[1 - g(r)] = g(r) + \tfrac{1}{2}g^2(r) + \tfrac{1}{3}g^3(r) + \cdots. \quad (16)$$

Under the condition

$$g(r) << 1, \quad (17)$$

$\Gamma(r)$ may be replaced by $g(r)$ in equations (12) to (15), thus completing the analogy (writing γ for g, $1/w$ for $1/wT$, and π for θ) with equations (1), (3), (31), and (36) of Chapter 2, derived for the case of continuous all-around looking. Although condition (17) is sometimes not satisfied for all ranges, it is often satisfied at ranges of particular interest. Thus, in the calculation of lateral range distributions for the larger targets, it is found that $p(x)$ approaches unity while $g(r)$ is still small. The approximation

$$\Gamma(r) = g(r) \quad (18)$$

may therefore, if used judiciously, yield much useful information and shorten computation considerably.

In order to obtain an analytical expression for $g(r)$, the single-scan probability of detecting a target at range r, it is necessary to determine the relationship between this *detection* probability and the blip-scan ratio (Section 5.5), representing the single-scan probability of obtaining a recognizable *blip* on the radar indicator. We shall discuss in detail a particular set of assumptions regarding this relationship, and also consider briefly the effects of slight alterations in these assumptions.

Concerning the radar operator, we shall make the assumption that if he has seen no blip on a particular scan his probability of noticing a recognizable signal on the succeeding scan is p_0; and that if he has seen a blip he will be alert on the following scan, certain to detect any blip presented. If, however, no blip is presented on this second scan, the operator is assumed to lose interest, his chance of noticing a new signal reverting to p_0. The value of this probability is dependent on the state of training and fatigue of the operator; we shall, for the time being, regard it as a known constant and as independent from scan to scan. We shall also assume that the operator must see a certain number n of successive blips for detection of a target to occur. For rapidly scanning airborne search radars, experience indicates that n has the value 2 or 3 in most cases.

Under these assumptions, we see that a necessary and sufficient condition for the detection of a target on a particular (ith) scan, given that no previous detection has occurred, is that blips be returned on that scan and on the preceding $(n - 1)$ scans, and also that the operator see the first of these n successive blips. Expressed in symbols,

$$g_i(r) = p_0 \prod_{j=i-n+1}^{j=i} \psi\left(\sqrt{x^2 + (y - iwT)^2}\right). \quad (19)$$

If wT is small, if n is small, and if ψ is a slowly varying function of range (all of these conditions are generally satisfied in practice), then the ψ's in the above expression are all nearly equal. Therefore, dropping the subscript, we have approximately

$$g(r) = p_0\psi^n(r). \quad (20)$$

Note that the multiplication of the ψ's in equations (19) and (20) is permissible, from the probability viewpoint, only if the independent probability assumption mentioned in Section 5.5 is justified.

Condition (17) under which $\Gamma(r)$ may be replaced by $g(r)$ becomes

$$p_0\psi^n(r) << 1. \quad (21)$$

It is satisfied at ranges for which the blip-scan ratio is small, and at all ranges if the operator is very inattentive. It should be noted that this "inattentiveness" may be of an effective kind, the result not only of actual distractions and fatigue but also of difficulty in finding the target blip among others of a random nature, dependent on noise level and characteristics of presentation. The value of p_0 will therefore in operations frequently be very small.

These assumptions regarding operator efficiency and the criterion of detection have been chosen because of their reasonable nature and the simplicity of the result they yield. They are by no means the only ones that might be made, and are, indeed, immediately suggestive of several similar ones, of perhaps equal validity. For instance, it may be that the detection requirement, instead of the occurrence of n *successive* blips, is the occurrence of n blips distributed in any way among s scans (s being a small number, greater than n). In this case, treating ψ as constant during the s scans, we have

$$g(r) = p_0C_n^s\psi^n(1 - \psi)^{s-n} = p_0C_n^s\psi^n + \text{higher terms}, \quad (22)$$

where C_n^s is the binomial coefficient

$$C_n^s = \frac{s!}{n!\,(S-n)!}.$$

If ψ is sufficiently small that the higher terms of equation (22) can be neglected, we again have an expression for $g(r)$ of the same form as that of equation (20), differing from it only by a constant factor. We conclude that for small values of ψ—those, as we have pointed out, which are usually of greatest interest—this *form* of $g(r)$ is insensitive to the exact nature of the assumptions made in determining it. We shall therefore use as our radar detection law

$$g(r) = k\psi^n(r), \qquad (23)$$

in computing range distributions and sweep width from equations (12) to (15). The exact value of the constant k (as in the analogous visual sighting case) and also of n, is best determined by comparison with test (or, for some purposes, operational) results.

5.7 COMPUTATIONAL METHODS

The application of the formulas derived in the preceding section to the solution of specific radar detection problems may be illustrated by a brief discussion of computational methods. In particular, we shall be interested in the utilization of test data for the predicting of operational results and in

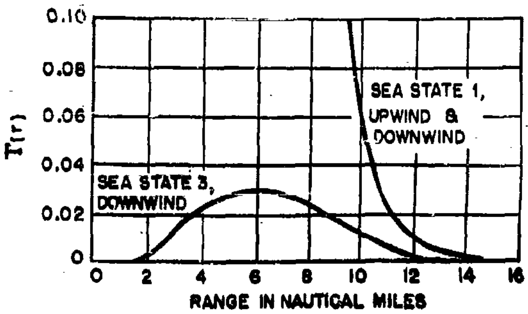


FIGURE 10. Graph of $\Gamma(r) = \log[1 - \psi^2(r)]$ as a function of range, using blip-scan data of Figure 9. Computed as an intermediate step in the determination of the true range distribution for search ahead only.

methods of correlating theoretical results with those of past operations.

The distribution of detection ranges under a particular set of conditions for search ahead only (zero lateral range), characteristic of certain convenient test procedures, may readily be computed by means of a single integration for each range value, provided blip-scan data obtained from tests made under the specified conditions are available; and provided assumptions are made regarding the values of the constant k and the number n of recognized blips necessary for detection. The integration to obtain $F(0,r)$ is performed graphically, in accordance with equation (11), first plotting $\Gamma(r)$ against r (see Figure 10). Use can be made, if desired, of the simplification em-

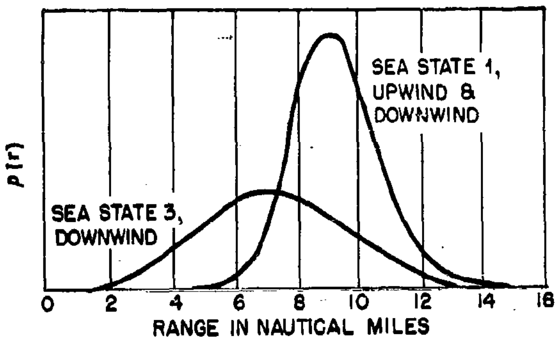


FIGURE 11. True range distribution for search ahead only. Computed using blip-scan data of Figure 9.

bodied in equation (18), for most values of r. The true range distribution is then computed directly from equation (10). Figure 11, as an example, illustrates the results obtained using the blip-scan ratios of Figure 9 and the assumptions $k = 1$ and $n = 2$ [i.e., $g(r) = \psi^2(r)$]. The particular value of such calculations is that range distributions computed under various assumptions as to the values of k and n may be compared with the range distributions actually obtained in the tests (provided these are statistically significant), as a means of determining which of these assumptions provide the best fit. Such trial-and-error calculations offer the most practical method of evaluating these parameters. It should be noted, however, that parameter estimates based on tests in general give optimistic results, as compared with those of operations; it may therefore be preferable, when such are available, to employ operational (rather than test) data, in the manner outlined in Section 4.10.

The computations involved in determining true range distributions for the case of sector or all-around search for targets uniformly distributed at all lateral ranges are similar in principle to those outlined above, but are more complicated, since additional integrations are required. The method of procedure, in brief, is to express equation (11) in polar coordinates by means of the familiar relationships $x = r \sin \zeta$, $y = r \cos \zeta$; then, for fixed values of ζ and of r, to determine by graphical integration the value of this expression and consequently of the integrand of equation (12); finally, for each value of r, to perform graphically the ζ integration required by the latter equation. If this is done for enough values of r, the desired range distribution may be determined as accurately as required. Figure 12 illustrates the type of distribution obtained. As before, AN/APS-15A blip-scan data are supplied by Figure 9, and the assumptions $k = 1$, $n = 2$ are made. As regards the general shape of the distributions, the results are seen to be not far different from those obtained for the previous case of search straight ahead. Sector scanning radars (such as AN/APS-3, which scans 150 degrees forward) also give true range distributions of a similar character.

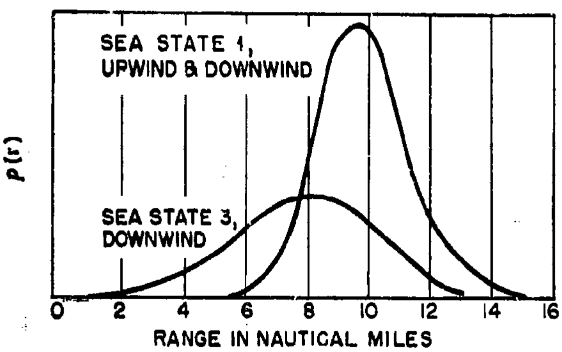


FIGURE 12. True range distribution for all-around search. Computed using blip-scan data of Figure 9.

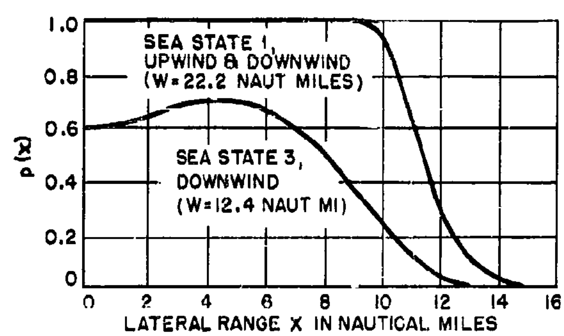


FIGURE 13. Lateral range distribution. Computed using blip-scan data of Figure 9.

The distribution of lateral detection ranges is most easily computed in terms of rectangular coordinates. Equation (11) is evaluated much as before, with the integration extending over all values of y in the scanned region, for fixed values of lateral range x. Equation (13) then gives the corresponding ordinates of the lateral range distribution. Graphical integration of the curve so obtained yields the value of the search width W [equation (14)] directly. Figure 13 shows lateral range curves obtained under the same assumptions as before; corresponding values of sweep width are indicated.

By collecting test data on $\psi(r)$ for known sets of conditions, varying only a single parameter, we can readily determine the influence of this parameter on search width. As an example, the effect of aircraft altitude on W is shown in Figure 14, for the case of

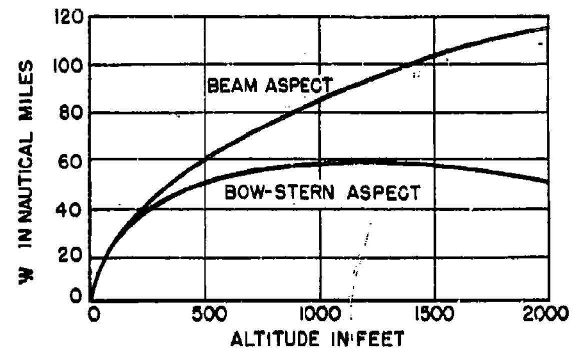


FIGURE 14. Search width W as a function of aircraft altitude. Search for surfaced submarine by AN/ASP-15A, sea states 1 and 2.

search for surfaced submarines, with the usual assumptions, $k = 1$, $n = 2$. The differences in search width for beam and bow-stern runs against this type of target are clearly shown. As previously indicated, however, true beam runs are encountered in operations with relative infrequency, so that the results for bow-stern runs give a truer average picture.

The effect on W of variations in another parameter, direction of search with respect to that of the wind, has been indicated roughly in Figure 13, for search against schnorchel in sea states 1 and 3. In sea state 1, it will be observed, wind direction is unimportant; in sea state 3, however, upwind search is already impossible, and the search width is greatly reduced even for downwind search. Results of an intermediate character are obtained for sea state 2. If area search is to be conducted in sea states greater than 1 by means of parallel sweeps in one direction (with a number of searching craft), it is clearly most advantageous to choose that direction as downwind. Usually, however, search is conducted on a roundtrip or shuttle basis; if this is the case, it might prove best to search at an angle to the wind. (Confirmation of this conjecture, however, awaits the obtaining of further test data on blip-scan ratio for crosswind search.) It should be noted that these remarks apply only to very small Class B targets, such as periscopes and schnorchels. For larger targets, including surfaced submarines, differences between upwind and downwind search are usually found to be of little significance. A useful empirical relationship regarding search width for these larger targets is that W is roughly equal to twice the range at which $\psi(r) = 0.1$, $\psi(r)$ being approximately independent of wind direction. (Since the blip-scan ratio for such targets rises rapidly in this range, W is not sensitive to the exact value $\psi = 0.1$.)

We have thus far dealt only with situations such as those to be encountered in future operations, in which conditions may be regarded as homogeneous, i.e., situations in which specific values may be assigned to each of the parameters in our equations. If we desire that our theory duplicate actual past operational results, however, the analysis does not proceed so simply. In this case it is necessary to adopt methods of *averaging* for distributions of the basic parameters. Occasionally these distributions are known; more frequently they must be estimated. The reader is referred once more to the final paragraphs of Chapter 2, in which the topic of long- and short-term parameter variations in relation to the analysis of past and future operations is summarized.

Owing to the inadequacies of available operational data and of knowledge concerning distributions, we shall not here enter into the discussion of averaging methods, but defer this until the next chapter. We may, however, mention again, for emphasis, a few of the long-term factors, the variations of which are particularly significant in the analysis of past operations. Namely, (1) set performance, which may change slowly over considerable periods of time (as in the case of seepage of moisture into wave guides or radomes, causing a gradual decline in the performance of certain types of radar) or more rapidly with changes in set tuning; (2) propagation conditions, which as we have pointed out may vary rather radically and unpredictably in certain localities; and (3) operator performance, which depends on training, experience, and alertness. *Qualitatively, the cumulative effect of these factors is always to increase considerably the dispersion of operational range distributions, and usually to reduce average detection ranges and search widths* (sometimes by a factor of 2 or 3). It is expected that analyses of operational data on Class A targets, for which the dispersion is largely due to such unassessed factors, will lead to a better understanding of these effects. Figure 15 shows an example of the type of range distribution obtained

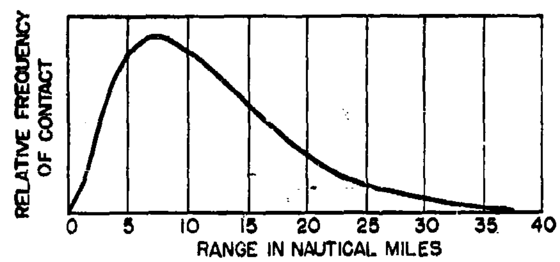


FIGURE 15. Operational true range distribution. Observed ranges of radar first contacts on surfaced submarines leading to attacks for the period July 1943 to March 1944. (Includes all types of radar, day and night service, under all weather conditions.)

in operations, during the earlier part of the past war, for surfaced submarines.

We have omitted consideration of a number of important topics relating to radar search, such as, for example, the effects of forestalling due to countermeasures (search receivers) in search for enemy units. Such forestalling, might possibly be met by operating the searching radar intermittently; but the effect of such operation on the radar search width must also be taken into consideration. We have likewise omitted discussion of visual forestalling and the calculation of combined radar and visual search widths (necessary for the determination of optimum search altitude). Such topics represent in themselves separate subjects, which must await more detailed study elsewhere.

Chapter 6

SONAR DETECTION

6.1 SONAR SEARCH—GENERAL

In subsurface warfare, reliance must be placed on sound (or supersonics) for search and detection, since sea water is virtually opaque to electromagnetic waves. Neither visual nor radar search is possible in a water medium. Consequently sonar must be used when searching for submerged submarines, torpedoes, mines, or other underwater objects. Magnetic detection is also possible, but the range of detection is normally much less than for sonar. Sonar search, therefore, is of importance in the operation of submarines and of antisubmarine forces.

Sonar detection involves either listening or echo ranging. Simple listening gear consists of a receiver and amplifier which pick up sounds generated by the target and present them to the sonar operator's ear. Echo-ranging gear has a transmitter in addition, which sends sound into the water; the sound received is then an echo reflected from the target. In general, listening gear has the advantages of simplicity and long detection range on a noisy target, but is not effective if the target runs quietly. Echo-ranging search has the advantage that it cannot be defeated by slow-speed quiet running. In addition, echo ranging provides information on *range* to the target, which listening does not. As a result, both listening and echo ranging are used for search, with preference sometimes given to the former, sometimes the latter.

In this chapter examples of both listening and echo-ranging search will be discussed, but only as examples to illustrate the type of problems involved. The aim of the operational analysis of sonar search is to determine how search gear or searching craft should be used in any particular situation to give the best result. This result may be expressed as a lateral range curve or a sweep rate in accordance with Chapter 2. There are a great many factors which determine the lateral range curve, involving characteristics of the gear, its operation, the target, its behavior, and sound transmission in the ocean. Some variables, such as speed of searching ship, can be determined by the searcher, whereas others, such as sound conditions, are beyond his control. The values of these uncontrollable variables often determine how the values of the others should be chosen. In any particular case, the various factors must be considered in detail, and the lateral range curve obtained in accordance with estimates of the physical situation. Before analyzing any typical problems, however, it is worth while giving a general outline of the factors which come into play.

In detecting the target the sonar operator must distinguish the signal from the everpresent background noise. Hence the factors of interest can be divided into the three following classes:

1. *Those which influence the strength of the signal which it is desired to detect*, the signal being either noise incidental to the operation of the target, sound transmissions by the target, or an echo reflected from it. These include characteristics of the sonar gear being used, of the target, and of the ocean.

2. *Those which determine the background level against which the signal must be heard*, including noise from own ship, noise from waves and animal life in the ocean, and, in the case of echo ranging, reverberation.

3. *Psychological factors and characteristics of the sonar data presentation* which determine the probability of detecting a given signal in the presence of a given background. Each of the subdivisions must be considered for both listening and echo ranging. They are exhibited in parallel columns, echo ranging on the left, listening on the right. Corresponding topics appear side by side, and when identical considerations apply to both cases, their treatment is written across the columns (i.e., the column division is temporarily abandoned).

ECHO RANGING

The scheme of echo-ranging detection is shown in Figure 1A:

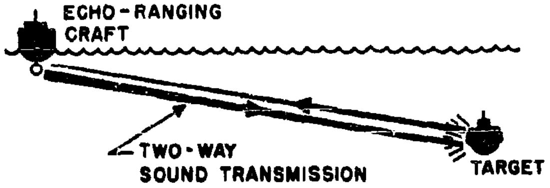


Figure 1A. Two-way sound transmission.

LISTENING

The scheme of listening detection is shown in Figure 1B:

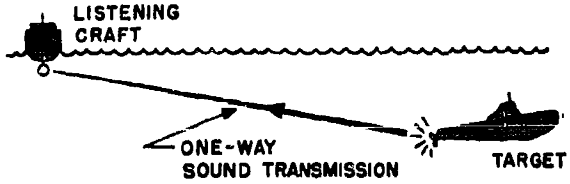


Figure 1B. One-way sound transmission.

Factors Influencing Signal Strength

1. Intensity of transmitted pulse:

The intensity of the transmitted pulse at a given range (say 1 yard) depends on the acoustic power output of the gear, and its directionality, the more directional the transmitter the greater the intensity for a given total power output. Standard echo-ranging gear now used by antisubmarine ships has an intensity of about 182 decibels above 0.0002 dynes per square centimeter at 1 yard on the axis of the projector.

1. Sound output of target:

The sound output of a ship depends primarily on the type of ship and its speed. At very low speeds machinery noise often predominates, which is almost entirely low-frequency sound. Propeller cavitation noise containing the higher sonic and low supersonic frequencies becomes important at normal speeds. A submerged submarine, however, produces this cavitation noise less readily the deeper it submerges. In addition, individual ships vary considerably from the average performance, especially in the details of their sound output. The graph in Figure 2A shows some typical sound levels in a 1-cycle band at 1 kc.

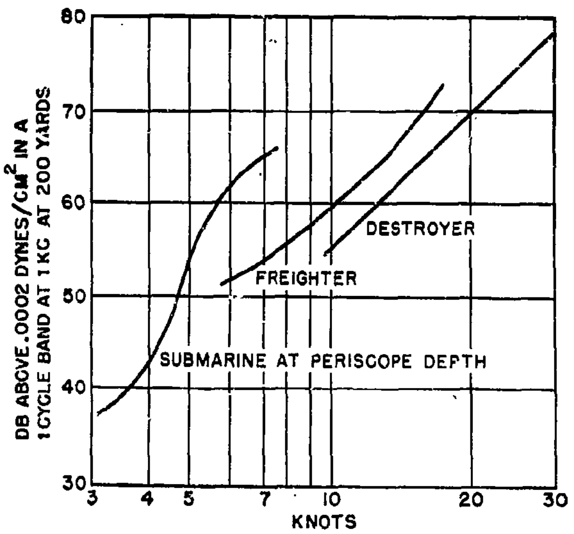


Figure 2A. Sound output of various ship targets.

ECHO RANGING *LISTENING*

2. Sound transmission:

These sounds suffer a considerable loss in intensity in transmission through the water, a two-way transmission in the case of echo ranging, one-way for listening. During a passage from ship to target or vice versa, a certain loss is suffered due to spreading, attenuation, and refraction. Calculation of the transmission loss in any particular situation is a complicated problem. In the absence of refraction or reflection the intensity in decibels, I, can be expressed as a function of range,

$$I(r) = I(1) - 20 \log_{10} r - ar. \quad (1)$$

The term $I(1)$ in equation (1) is the intensity at a distance of 1 yard from the source, $20 \log_{10} r$ is the loss of intensity due to geometrical spreading of the sound (inverse square law), and ar is the loss due to absorption of energy by the water, a is the attenuation in decibels per yard (assuming r to be expressed in yards). This equation is strictly valid only if there is no reflection or refraction of the sound beam. In many cases, however, the effect of refraction can be represented by an increase in the value of a: assigning an "effective attenuation constant," a may then be regarded as an empirical constant which depends on the frequency of the sound and temperature distribution in the ocean. The graph in Figure 2B gives values of the transmission loss, as calculated by equation (1) for some typical values of a. The effect of reflections from surface and bottom is neglected. A complete analysis of sound transmission would take into account reflection from various types of surface and bottom, and also refraction by temperature gradients whose effect cannot be represented by an effective attenuation constant. Such an analysis is outside the scope of this chapter, but can be found in Volumes 6A, 7, and 8 of Division 6.

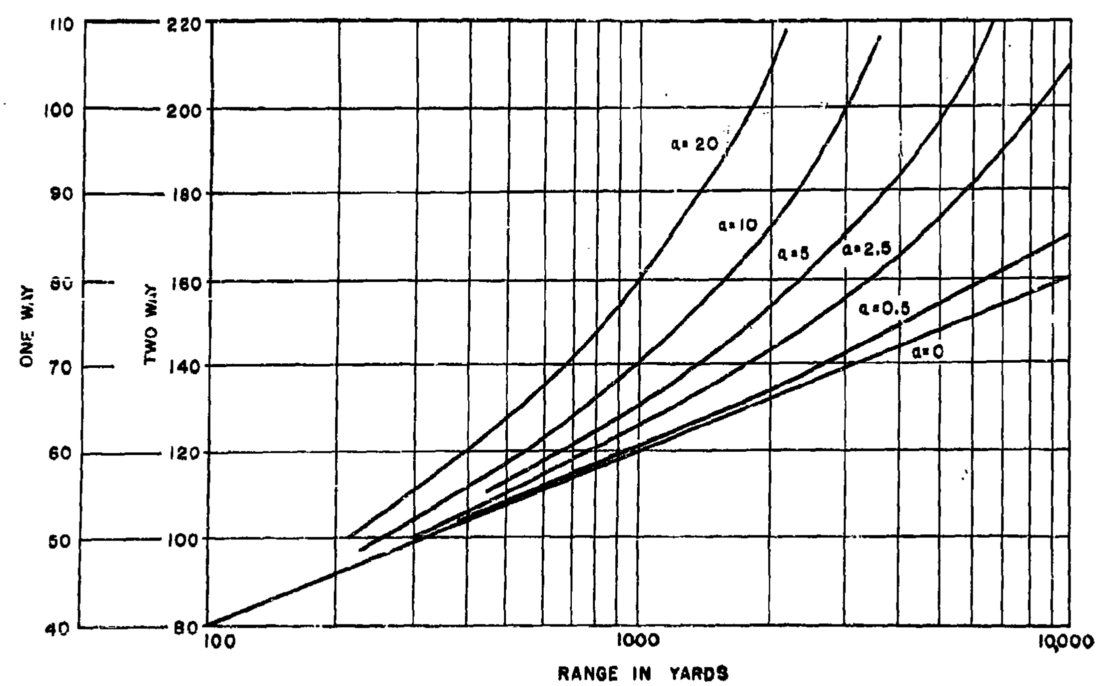


FIGURE 2B. Transmission loss from 1 yard.

ECHO RANGING *LISTENING*

3. Reflecting power of the target:

When the sound beam strikes the target, the amount reflected depends on the size and shape of the target, the nature of the target material, and also its orientation. The frequency and ping length of the sound being reflected are also of importance. These various factors determine the "target strength" which gives the intensity of the echo (reduced to 1 yard from the target) relative to the intensity of the outgoing ping when it hits the target. For a submarine, typical values would be

Bow aspect 16 db
Beam aspect 25 db
Stern aspect 8 db

4. Receiver characteristics:

The factors above determine the signal which arrives at the receiver. The characteristics of the gear, however, have a great deal to do with the nature of the signal that is presented to the sonar operator. In general the gear will have a sensitivity that depends on frequency and the direction from which the sound is approaching. At any frequency the directionality of the gear is determined by the physical properties of the receiving microphone (transducer) and its mounting. Both the transducer and selective elements in the receiver circuits come into play in determining the frequency response of the receiver.

Factors Influencing Background Level

1. Ambient noise in the ocean:

Like the air which we inhabit, the ocean is not normally completely quiet, but full of various noises, man-made and natural. Chief among the man-made noises is that due to the searching craft, which will be dis-

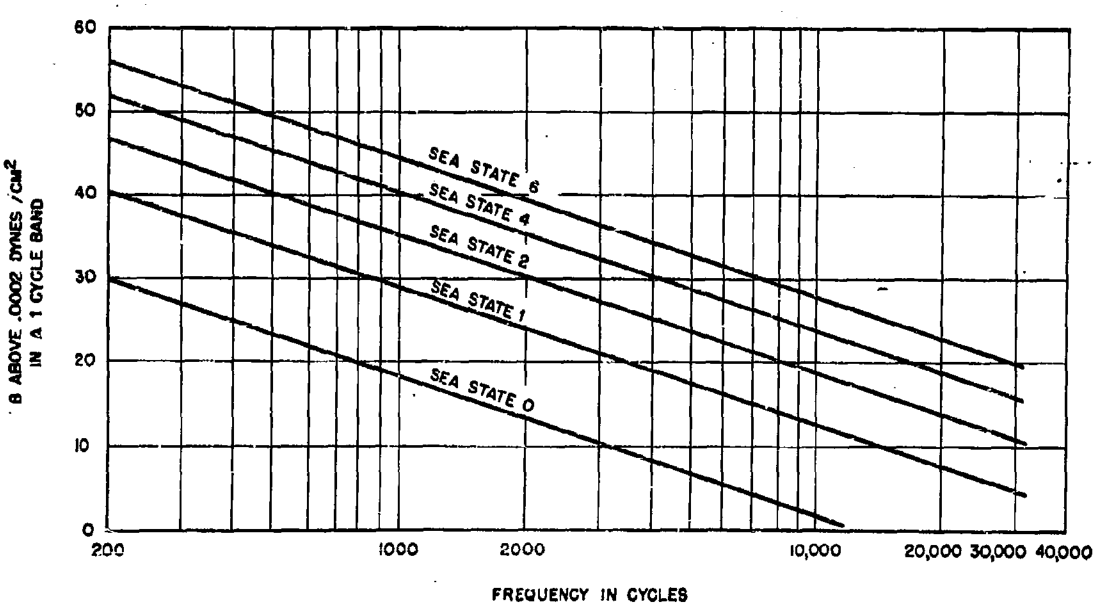


Figure 3. Ambient noise.

cussed later. The true ambient noise due to natural causes is also considerable, being generated primarily by wave action. The average frequency distribution of this noise for various sea states is shown in Figure 3.

Marine animals may also contribute to the ambient background. Famous for such activities is the snapping shrimp, a bed of which may produce a level of about 40 decibels above 0.0002 dyne per square centimeter in a 1-cycle band above 5 kc.

2. Self-noise created by searching craft:

Since the receiving transducer is necessarily in close proximity to the searching craft, any noise generated by it will be heard as a contribution to background level. This self-noise may be caused by the propellers, by moving machinery in the ship, or by the rush of water about the face of transducer or dome. In any case, the self-noise increases rapidly with increasing speed (see Figure 4). At low speeds, ambient water noise often overrides it, but at high ship speeds, self-noise is the important factor.

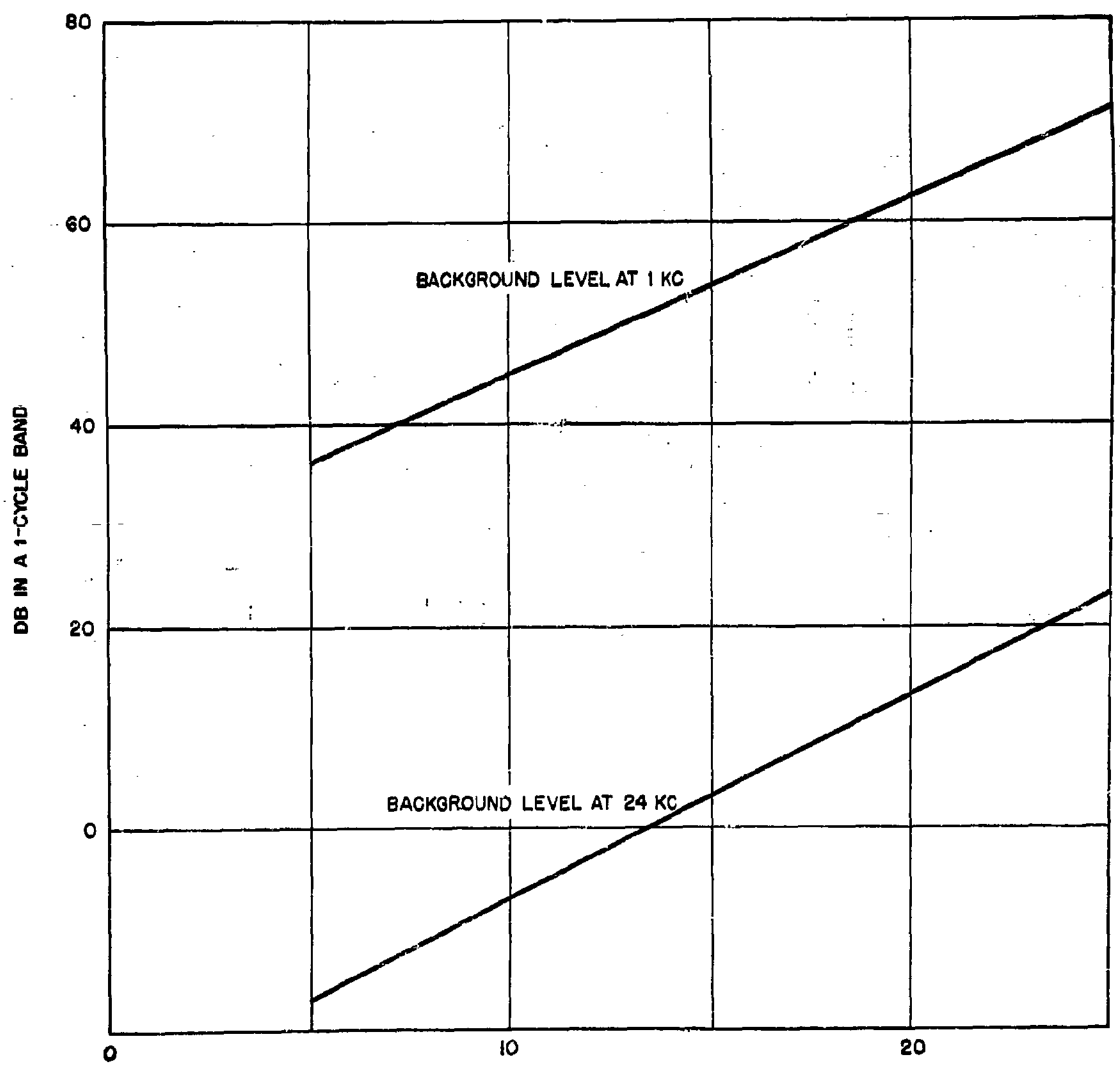


FIGURE 4. Background level for typical DD with dome.

ECHO RANGING | *LISTENING*

3. Reverberation:

The signal which is sent out by echo-ranging gear is reflected by many objects in the ocean besides the desired target. Irregularities of surface and bottom are the most important reflectors, but there are slight echoes from the body of water itself. The totality of these many false echoes is called *reverberation.* When reverberation is severe it may override other types of background noise and be the chief factor limiting the sonar range. Bottom reverberation from a rough or rocky bottom is the source of highest reverberation level. Surface reverberation may also be considerable when the surface is rough and bubbly. Other factors such as ping length, frequency, and modulation also influence the reverberation level, as does the type of sound transmission to be found in the ocean at the time.

4. Characteristics of sonar:

The background level presented to the sonar operator depends on the sensitivity of the sonar, its frequency selectivity, and directionality. Since the sources of background noise do not in general have the same frequency distribution or the same bearing as the target, the frequency response and directionality of the gear affect background noise and signal differently; this may either facilitate or hinder target recognition. By choosing the frequency response so that the signal from the target is at the frequency of maximum sensitivity and training a directional type of gear so that its direction of maximum sensitivity is oriented toward the target the range of detection can be considerably increased.

Factors Influencing Recognition of Signal

For any given signal strength and background heard by the operator, there is a certain probability that he will recognize the signal. A very loud signal will surely be detected, a weak one has only a small chance. The level of signal relative to background for which the probability is 50 per cent is called the *recognition differential.* There are various means by which the signal is presented to the operator. In aural recognition the sound is presented to his ear by phones or a loudspeaker. Sometimes visual schemes are involved in which the signal is displayed as a spot on an oscilloscope or on sensitized paper. In any case the chief factors influencing it are the following:

Echo ranging:

1. Type of signal:

Length of ping, for example, is important in aural detection. A very short ping is not recognized as readily by the ear as a longer one. For visual detection, as on an oscilloscope, this is no longer the case. Doppler frequency shift is valuable in differentiating aurally between signal and reverberation but is not useful with visual data presentation. Both the type of signal and the way it is presented to the operator are of importance.

Listening:

1. Type of signal:

The signal to be detected by listening gear varies widely in character from one target to the next. In broad-band listening, characteristic sounds such as propeller beats, gear whine, and machinery noises can be recognized even when their level is much lower than the background, and the recognition differential depends on the extent to which such noises aid in recognition. If the gear used for listening is sensitive only to a narrow band of frequency, however, the signal must be approximately equal to the background to be recognized.

ECHO RANGING

2. Type of background:

As has been indicated, the recognition differential is different according as the background is reverberation or water noise. The doppler shift aids in discriminating against reverberation, since the ear can detect the difference in pitch. Water noise, however, contains all frequencies, so that doppler is of no help. Typical requirements for recognition differential are:

a. With 0.1—second ping vs background in 1-kc band: −7 db;

b. Vs reverberation, no doppler: +3 db;

c. Vs reverberation 50 cycles doppler: −4 db.

3. Data presentation:

Gear in which the operator actually hears the signal normally have different recognition differentials from those in which a spot on an oscilloscope or on chemical paper is the means of detection.

4. Operator skill:

Aural acuity and pitch perception enter into the question of recognition differential. Training in doppler recognition is of importance. Operator attentiveness and fatigue are undoubtedly significant in routine search operations.

LISTENING

2. Type of background:

The type of background may often influence the recognition differential. For instance, shrimp crackle, which is mostly high frequency, would not be very effective at masking low-frequency machinery sounds, and listening to them would have an unusually favorable differential with a background of that type.

4. Operator skill:

In the recognition of typical sounds from the target the skill of the listener plays a large part. He must know what he is listening for, a knowledge acquired only by training. Hence the recognition differential depends a great deal on the state of training and skill of the operator. Fatigue and inattention can undoubtedly materially change the differential.

This enumeration of factors having to do with the effectiveness of sonar search gear is by no means complete. It is presented merely as an indication of the type of factor which should ideally be taken into account. If all these factors (and any others that are not mentioned) were accurately known at any instant, then it might be possible to decide precisely whether a target in a given position would be detected. As pointed out in Chapter 2, however, detection is never certain because of the human factor involved. We cannot tell whether detection will occur in a particular instance without knowing the detailed processes going on in the brain of the sound operator, and in addition we would have to be able to predict the acoustic behavior of the waves ahead of time, and to have already made extensive acoustic measurements on every target (usually an enemy craft) that might be encountered and to *know* when searching for it just which enemy craft was being sought. In any practical situation we can only estimate the probability that the target would be detected on the basis of the average values of various factors and their expected variation.

Methods for expressing the probability of detection in quantitative terms are given in Chapter 2. They involve the use of an instantaneous detection probability coefficient γ to take care of the human variable and the "short-term" fluctuations, viz., factors whose changes take place in a time that is short compared with the time taken by the full search operation. Normally there are also slowly varying factors that may, for instance, vary from day to day but are constant during the time of the operation. If these alone are of predominant importance, so that the psychological and short-term effects can be ignored, it is convenient to assume that there is, in effect, a *definite range* at any time, and it is the distribution of these individual definite ranges which gives the overall lateral range curve. This is in accordance with Section 2.9 of Chapter 2.

In order to show how some of these factors enter into the picture, two examples will now be discussed in some detail. The first of these is the *expendable radio sono buoy* [ERSB], a nondirectional listening device, the second is the standard directional echo-ranging gear used by antisubmarine ships. Similar treatments could be made and have been made for many other types of gear.

6.2 EXPENDABLE RADIO SONO BUOY

The *sonobuoy* is a nondirectional sonic listening device which is normally dropped from aircraft, floats at rest on the surface, transmits the sounds heard in the water to the plane via a radio link. Its use is to enable the aircraft to obtain sound contact with a submerged submarine. Because of its simplicity, it is suitable as a first example to illustrate the problems of sonar search, though it is not an altogether typical example of sonar search gear.

There are two possible approaches to the problem of determining the lateral range curve for this, or any, form of detection. Either the curve can be calculated on an a priori basis from estimates of the factors involved, or it can be determined from operational data. The latter is more reliable when sufficient data have been gathered, but it is not always possible to do this. Calculations of the first sort are always valuable in that they throw considerable light on the importance of the various factors involved and help in the interpretation of operational results. In the following discussion a lateral range curve will be obtained on an a priori basis, and will then be compared with available operational data.

The factors referred to in the previous section determine the chance of detection in any particular case. It is therefore necessary to assign values to them. The quantities required are:

P_1 = sound pressure of source at distance 1 yard in the sonobuoy band (0.1 to 10 kc);

μ = transmission loss in traveling from source to sonobuoy; according to equation (1) this transmission loss $[I(1) - I(r)]$ is given by $20 \log_{10} r + ar$;

B_L = background level received by sonobuoy, both water noise and any self-noise or circuit noise in the buoy itself (in same band as used for P_1);

Δ_L = recognition differential, i.e., required signal level with respect to background for detection to occur.

The target will just be detected (i.e., detected with a 50 per cent probability), if (values in decibels):

$$P_1 - \mu = B_L + \Delta_L, \tag{2}$$

or

$$\mu = P_1 - B_L - \Delta_L.$$

If all these quantities were definitely known, this equation would specify precisely a range within which the chance of detection exceeds 50 per cent, being considerably greater for most of this range. Approximately, then, it would define a definite maximum range. Each of the quantities has a considerable range of variation, however. For instance, background level is quite different for a sea of force 1 from what it is with force 5. Hence the expected sonobuoy range depends on sea state. An overall lateral range curve must give overall results, however, for those sea states met in practice. Each of the other factors will be variable also. The chief causes of variation in them are

P_1 actual speed of submarine is unknown, and sound output at a given speed varies widely from one submarine to the next;

B_L background varies with sea state;

Δ_L differential depends on the skill of the operator and the type of noise made by sub;

$\mu(r)$ depends on sound conditions.

While these factors may change widely from one case to the next, it is evident that in any particular case they are more or less fixed. There is no major cause of variation which would be expected to give large changes during the time that the submarine is near the sonobuoy. (The opposite case is true for echo ranging, when the echo level may fluctuate broadly between successive pings.) For sonobuoys, however, we here neglect the small short-term fluctuations and the effect of "human fallibility," and treat each combination of the variables as defining a definite range. On this basis we proceed to calculate the range for each combination of values for the important factors and obtain a distribution of these ranges corresponding to the assumed distribution for these values. This can be done conveniently if the values of P_1, B_L, and Δ_L are assumed to be normally distributed. From available data, it is reasonable to assume values as follows:

	Mean	*Standard deviation*
P_1 (decibels above 0.0002 dyne per centimeter at 1 yd in sonobuoy band)	128	5
B_L (decibels above 0.0002 dyne per centimeter in sonobuoy)	76	4
Δ_L decibels	−8	4

There is also considerable possible variation in $\mu(r)$. This can be represented by using a spread of attenuations, taking the attenuation equally likely to be anywhere between 0 and 3 db per kiloyard. We can combine the values of P_1, B_L, and Δ_L to give

$$P_1 - B_L - \Delta_L = 60 \pm 8 \text{ db.} \qquad (3)$$

Equation (3) can be solved graphically by a diagram of the sort shown in Figure 5.

The horizontal lines are drawn so as to divide the vertical axis into regions of 10 per cent probability, between 1 and 4 knots and individual variations of about 4 db in sound output of individual submarines. B_L corresponds to a sea state of $2\frac{1}{2}$ with two-thirds of the cases falling between state $1\frac{1}{2}$ and $3\frac{1}{2}$. As far as the overall result goes, however, it does not matter just what the source of variation is so long as the net result is a σ of 8 db. In order to show the effect of changes in sea state, for example, similar calculations

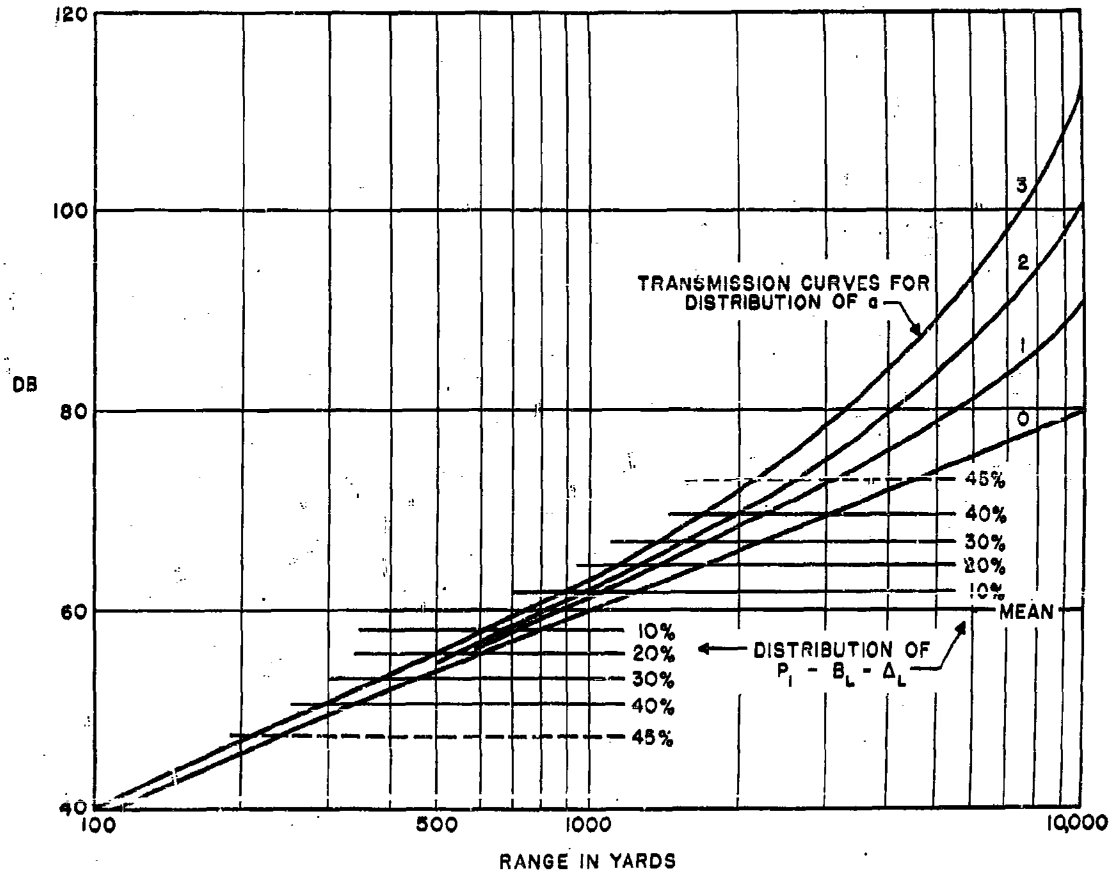


FIGURE 5. Graphical solution of equation 3.

and the transmission curves have the same property, that in any particular case there is one chance in three of the desired point lying in either of the three regions. These lines intersect forming cells, each of probability 1/30. If we assign to each cell the range of its midpoint, we have a theoretical distribution of ranges, and correspondingly a lateral range curve.

The curves in Figures 6 and 7 represent expected overall results for a considerable range of submarine speeds, sea states, and other factors. The value of P_1, for instance, corresponds to a speed of about 3 knots, with about two-thirds of the cases falling would be made for a series of fixed values of the sea state. This can be done for any factor whose influence it is desired to study closely.

We will, however, pass to the second, or operational, aspect of the problem without going into further detail. The theoretical calculations are, in actual fact, not purely theoretical, but are based largely on tests of the gear and experimental runs under controlled conditions. Such tests must, of course, be made to give an idea of the operational performance of the gear and a firm basis for theoretical predictions of its effectiveness. It is assumed that theory and test data will always be reconciled and put in agreement. The final check on the performance of the gear lies in the results of actual operations. This aspect of the problem will now be considered.

Data on operational sonobuoy ranges cannot be obtained directly. If the report of the sonobuoy contact gives the buoy pattern, the period of time that the submarine was heard on one buoy or several buoys, and evidence, such as propeller beats, of submarine's speed, then the detection range of the sonobuoy can be estimated. This has been done on the basis of data reported in connection with aircraft attacks on German U-boats, and on the basis of these estimates the following distribution of contact ranges was obtained.

Range in Yards	*Number of Cases*
0- 499	0
500- 999	4
1000-1499	4
1500-1999	9
2000-2499	9
2500-2999	2
3000-3499	5
3500-3999	3
4000 and over	3

This is the actual distribution of ranges at which contacts were made, and is consequently prejudiced in favor of the longer ranges. If detection ranges of 500 to 5,000 yards were all equally likely, many more 5,000-yard contacts would be made than 500—about ten times as many. Consequently, the distribution of potential contact ranges is obtained by dividing each of the above figures by the range. This curve of po-

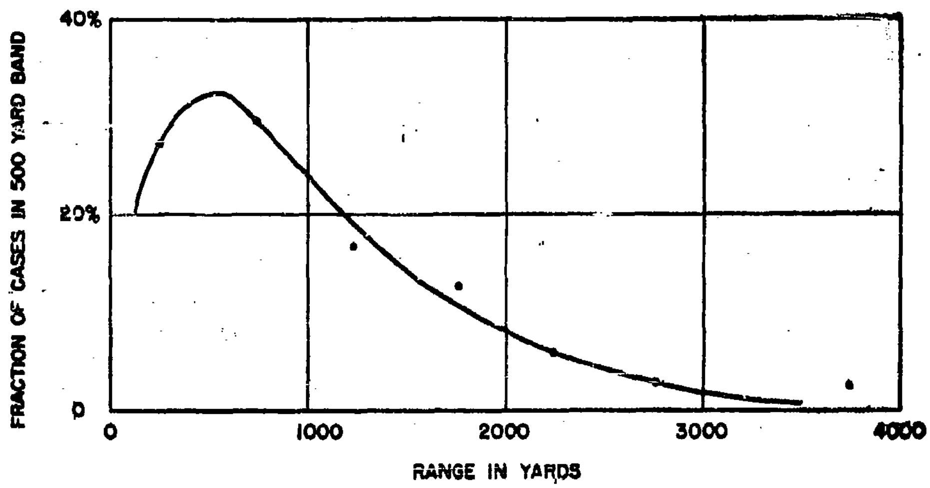


Figure 6. Distribution of ranges (a priori).

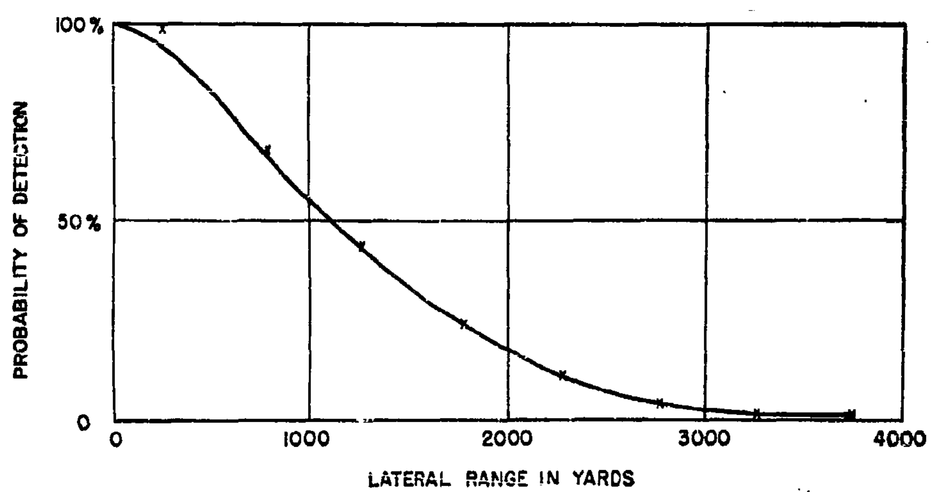


Figure 7. Lateral range curve.

tential contact ranges corresponds to the theoretical curve for distribution of ranges, and is plotted below in Figure 8. A lateral range curve is also shown in Figure 9. It is evident that the operational ranges are, on the whole, better than those predicted on a theoretical basis. There are a number of possible explanations of this discrepancy.

The most obvious explanation is to assume that some of our estimates of factors such as U-boat sound output have been too conservative. An increase of four or five decibels in assumed mean U-boat sound output would account for most of the observed difference. Similarly actual sea states may have averaged somewhat lower than was assumed, or the recognition differential may have been more favorable. It is not possible to decide just what the cause is without more detailed analysis or further information.

It may be, however, that the reason for the discrepancy lies in our interpretation of the operational data. Perhaps a larger number of potential short-range contacts were lost than was assumed in adjusting the distribution of actual contacts to obtain the "potential" range distribution. This would be the case, for instance, if contacts were lost because the submarine got past the sonobuoy while the observer was listening to a different one. Since it takes about 10 minutes, on the average, for a two-knot U/B to pass through the detection circle if the range is 500 yards, the chance of missing a U/B that does pass through the circle should be small. On the whole it seems most reasonable to conclude that operational results indicate that these theoretical values are somewhat too pessimistic, and they were, indeed, intended to be rather conservative.

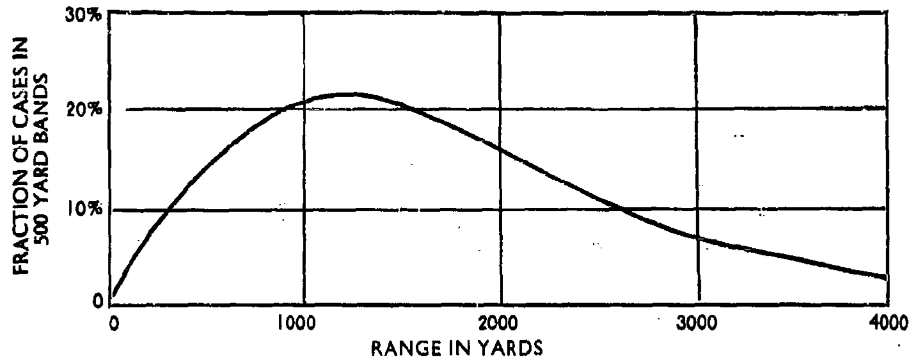


FIGURE 8. Operational range distribution (adjusted).

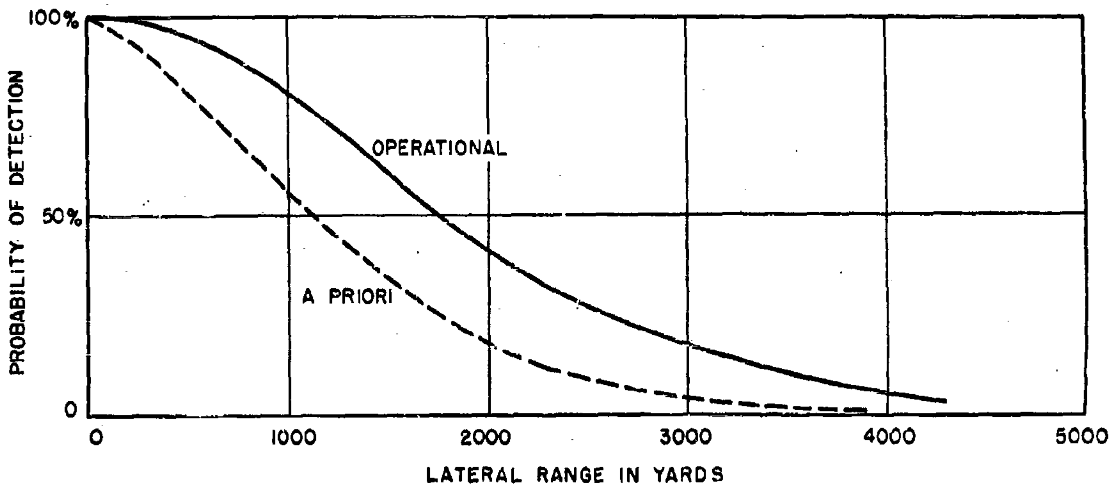


FIGURE 9. Lateral range curves.

6.3 STANDARD ECHO-RANGING GEAR

Echo-ranging gear such as is used by antisubmarine ships is characterized by a slow intrinsic search rate due to its directionality and the slow speed of sound (compared with electromagnetic waves). Only a small segment of ocean is searched by the gear in any short time interval (say one second) (see Figure 10). Consequently there are only a small number of *glimpses* at any particular target: the point of view of *glimpses*, as outlined in Chapter 2, Section 2.2 must be maintained.

As was the case for sonobuoys, both a priori calculations and operational results must be used. More factors enter into the former, however, and the process outlined in Chapter 2, Section 2.9 must be applied. First, we deal with a situation in which the variable physical factors are fixed, i.e., the sound

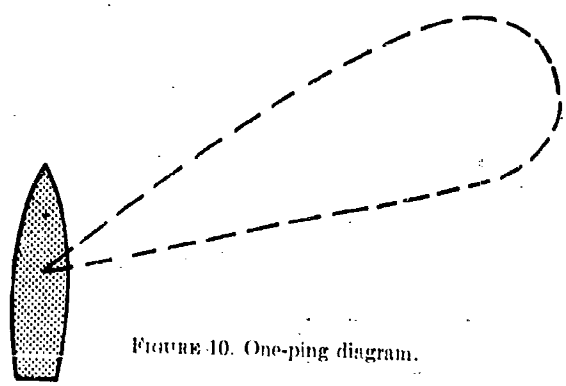


Figure 10. One-ping diagram.

conditions, type of sound gear, orientation of sound beam, etc., are unchanging. There are, however, certain factors which are bound to fluctuate, in particular the intensity of echo on a particular ping, the intensity of background, and the operator's ability to discriminate between echo and background. Even for given values of the former conditions, these latter fluctuations impart an uncertainty to the detection; and it becomes necessary to consider the probability of detection for the particular ping in question, the "one-ping probability." Second, these values for the "one-ping probability" must be combined in accordance with Section 2.4 (in the existing state of the science, the formal calculations being, of necessity, replaced by graphical methods) to give the probability of detection by the particular search in question, a "fixed conditions probability." In this way we can get, for example, a "fixed conditions lateral range curve." By doing this we assume that the *fixed* conditions are indeed constant throughout the search and the fluctuating conditions do, in fact, fluctuate from ping to ping, so that the values for each ping may be chosen at random and independently from a suitable distribution. Third, the "fixed conditions probabilities" which result must be averaged over appropriate distributions of values for the conditions to correspond to our knowledge of them. If, for example, we desire to plot a lateral range curve for a particular ship, *certain* ship speed, bathythermograph pattern, sea state, but *uncertain* U-boat depth, speed, and operator alertness, we would pick the proper values of the certain factors, average over the assumed distributions of the uncertain ones, and obtain a curve. An overall average curve for all ships and conditions would require much more extensive averaging, and surely would appear to be quite different.

We will now proceed to go through these steps for a typical example; first the one-ping probability.

In this one-ping probability all conditions are regarded as fixed except for fluctuations in echo, back-

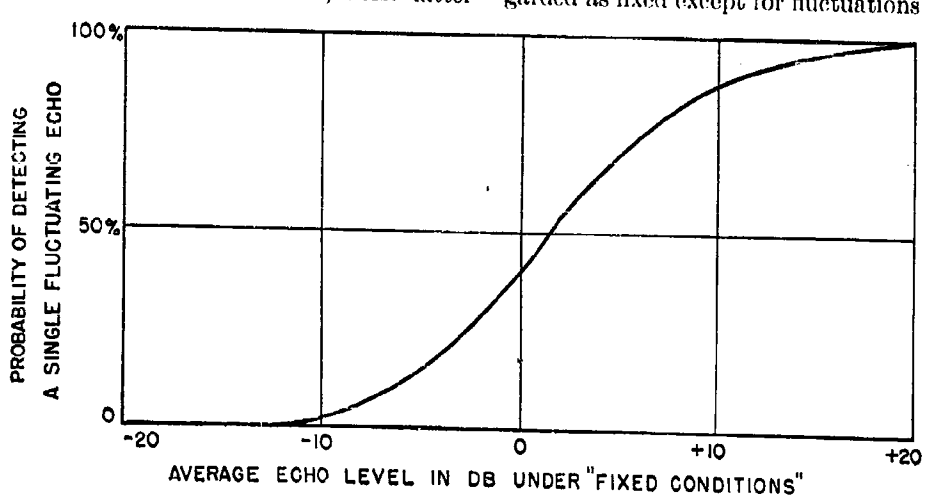


Figure 11. Dependence of probability of detection of one ping on the sound level above background.

ground, and operator alertness. The effect of these fluctuations can be represented by a curve (Figure 11) which gives the probability of recognizing the fluctuating echo as a function of its average level. This average level is given on a scale relative to the background. The zero on the scale is the echo level which would ideally be required for recognition if no fluctuations occurred. It is now necessary to calculate the echo level with respect to background as a function of target position in the sound beam.

The target's position in the sound beam may be specified by two variables, range and relative bearing. Range introduces a transmission loss as in Figure 2B, and bearing a loss relative to the axis because of directionality. As in Figure 5, various factors must be taken into account to determine the admissible loss which will just give recognition (i.e., with a 50 per cent chance), fluctuation neglected. These factors are

- P_1 transmitted signal strength at 1 yard;
- B_p background level heard in gear if reverberation is limiting; B_p is higher for short-range echoes than long, i.e., a function of range;
- T average target strength;
- Δ_p ideal recognition differential for nonfluctuating case;
- $\mu(r,\beta)$ transmission loss relative to unit range on the axis (see Section 6.2).

$$P_1 - \mu + T = B_p + \Delta_p \qquad (4)$$

$$\mu = P_1 + T - B_p - \Delta_p.$$

It is necessary to assume values for these factors in accordance with the specific case in question. For simplicity, we will assume that the transmission loss can be represented by inverse square law with 7 db per kiloyard attenuation, as might be the case for a target below the thermocline. Using this assumption and the typical directivity pattern of Figure 12, we can draw in the curves for μ as a function of r for a number of values of β.

The solid curves of Figure 13 are drawn on this basis. Values of $(P_1 + T - B_p - \Delta_p)$ must also be plotted to give a graphical solution of equation (4). Taking into account the probability of detection curve of Figure 11 we can add or subtract suitable values from $(P_1 + T - B_p - \Delta_p)$ to give the transmission loss for 0 per cent, 20 per cent, 40 per cent, · · · probability of contact. Intersections of these dotted levels with the transmission loss curves give the value of the probability of detection, i.e., the one-ping probability, for the various ranges and bearings.

On the one-ping probability curves (Figure 14), all curves have been terminated at a minimum range of 500 yards. Since the submarine is thought to be deep, the likelihood of contact at ranges shorter than

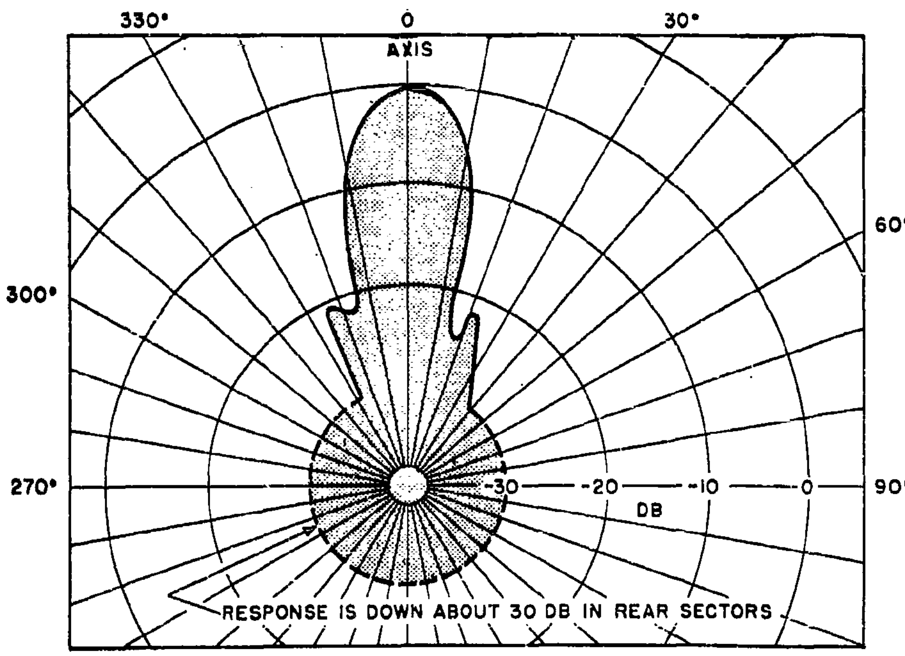


FIGURE 12. Directivity for typical projector.

500 yards is not great. There is a small area in which contact is theoretically certain, but the greater part of the beam has probabilities of 0 per cent to 80 per cent. The assumptions upon which this one-ping probability function is based are the following:

P_1 180 db, corresponding very closely to the signal output of standard gear.

B_p 40 db in a 1-kc band, heard in sound gear. This takes into account the effect of directionality in discriminating against noise background; it corresponds to typical self-noise at a speed of about 15 knots. No reverberation is considered.

T 15 db, a value which is typical of a submarine except at beam incidence.

Δ_p −7 db, the value obtained in laboratory tests for recognition differential relative to noise in a 1-kc band.

Having obtained the one-ping probability for the conditions under consideration, we must now obtain the corresponding lateral range curve. The method of Chapter 2 applies in principle, but cannot readily be carried out exactly in terms of formulas in the present case; it is necessary to resort to graphical methods of calculation. Suppose that the one-ping probability function is represented by a group of

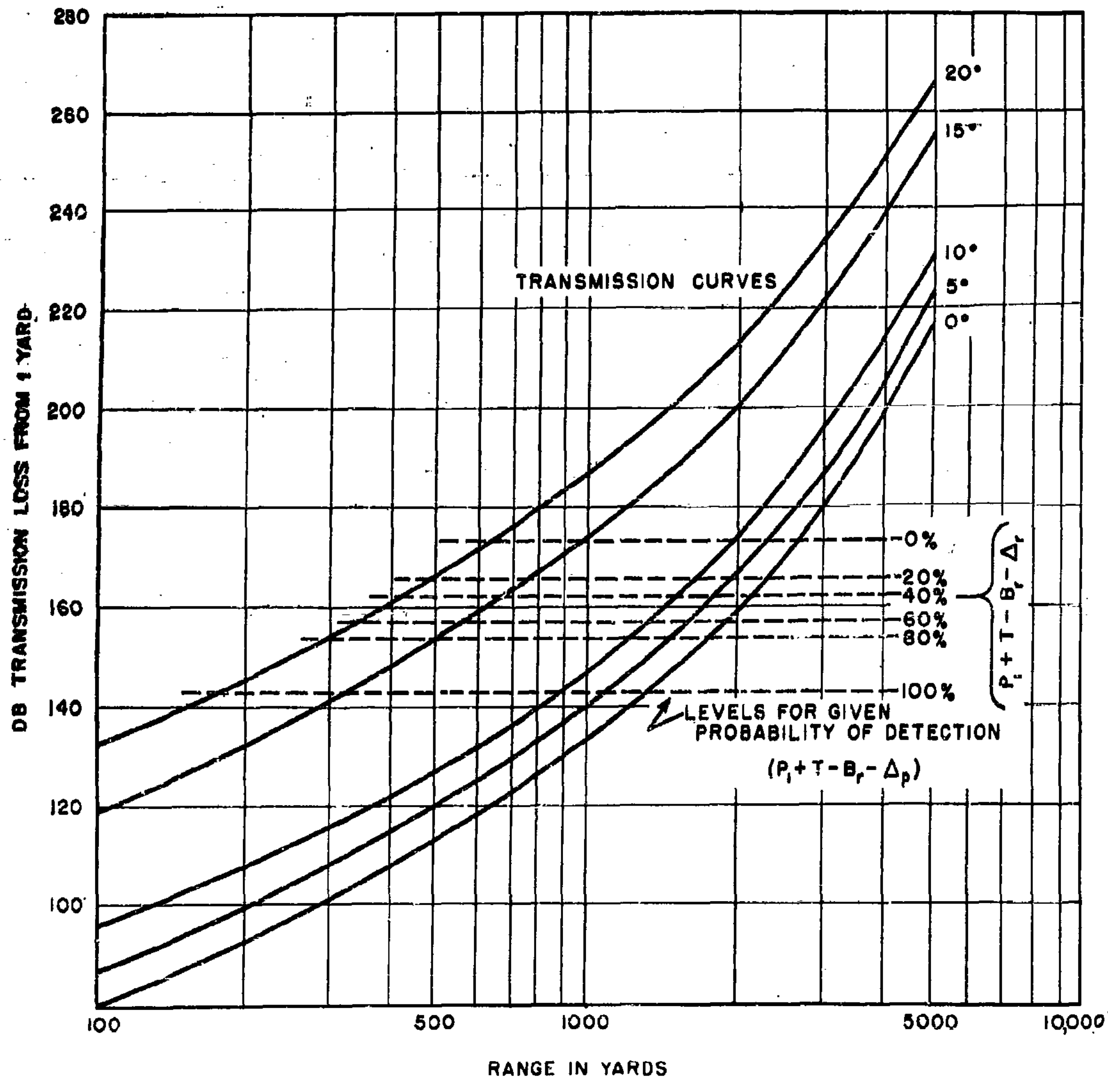


FIGURE 13. Graphical probability calculation.

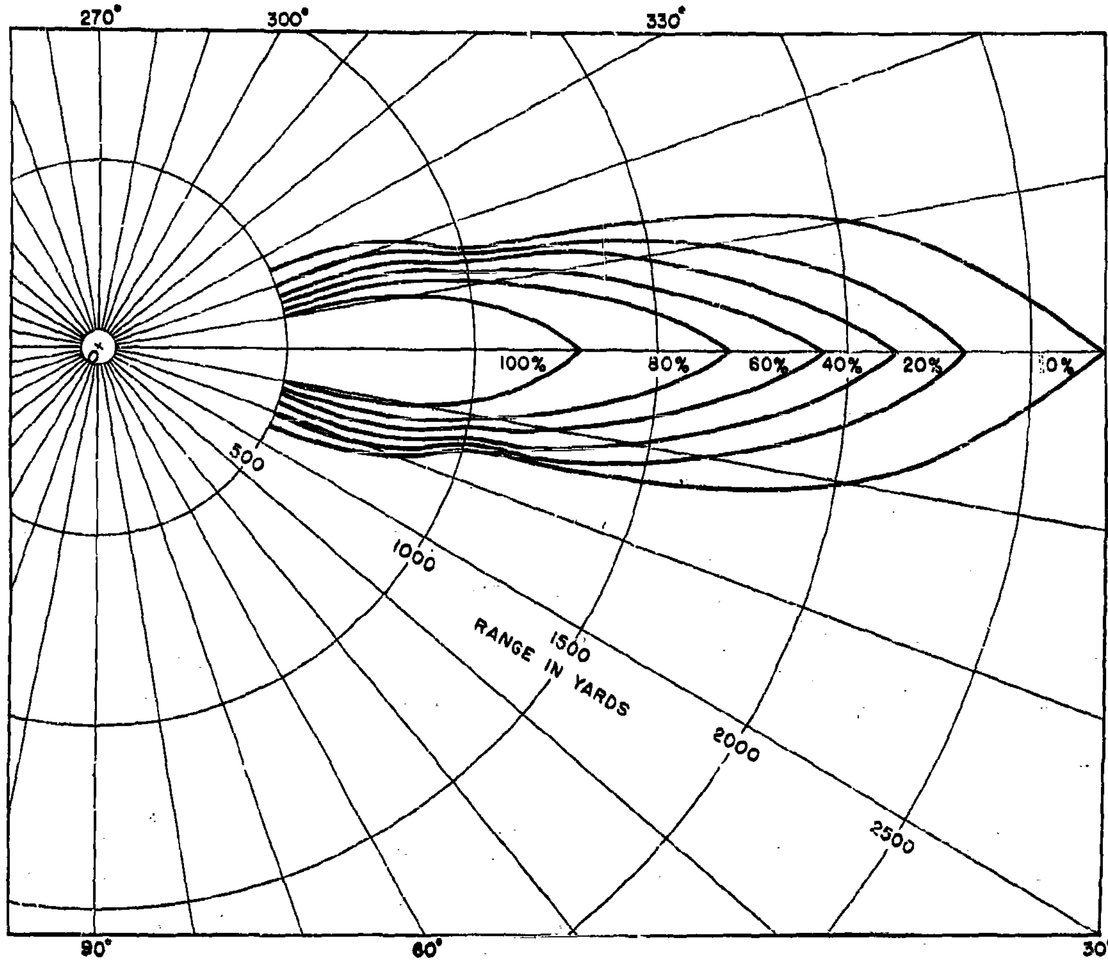


Figure 14. One-ping probability contours.

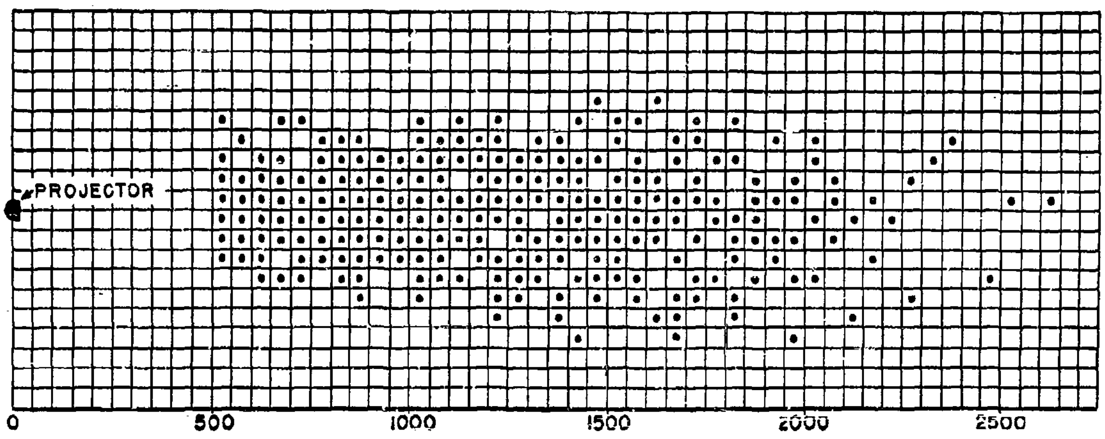


Figure 15. One-ping probability diagram.

dots, the density of dots being proportional to the probability. For example, in the diagram (Figure 15), 10 per cent of the squares between 0 per cent and 20 per cent contours have dots in them, 30 per cent of those between 20 per cent and 40 per cent, and so on.

We now decide on a planned sequence of pings, i.e., starting with projector trained abeam and training forward 5 degrees between pings until the bow is reached, and then slewing aft to the beam on the other side and pinging forward again by 5 degree steps. If we assume that the one-ping probability is not changed by training the projector from one bearing to the next, and assumed a relative velocity for searcher and target, then we can lay out successive pings in target space (i.e., space fixed relatively to the target). Target space is divided into 50 yard squares, and the event of a dot on the one-ping probability diagram filling in a square means that a submarine in that square would be detected. This process can be followed mechanically in the following fashion. Transparent paper divided into 50-yard squares is placed over the one-ping probability diagram in position for the first ping. Each square including a dot is blacked out. The diagram is moved to position for the second ping, and the process repeated for all pings. This gives an area (see Figure 16) black near the center and white at long range, the fraction of squares blacked out giving the probability of catch at any region. Such a diagram will be periodic if the pinging plan is regular, and by averaging the probability over one period for a given lateral range, we can obtain the lateral range curve. The result for this case is shown in Figure 17 at the same scale as previous lateral range curves. This accordingly completes the second step in the required calculations.

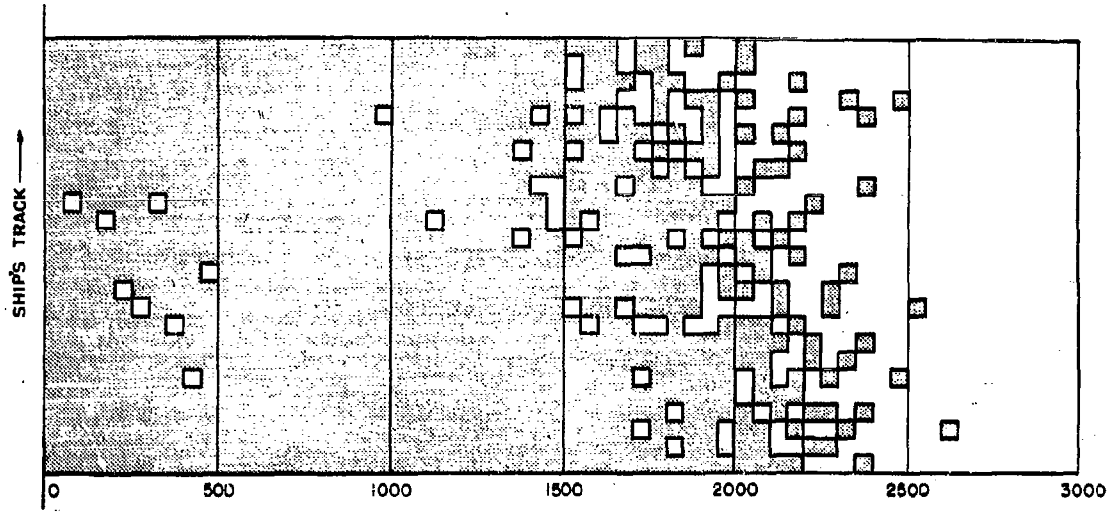


FIGURE 16. Coverage diagram (for one cycle of plan).

Before discussing the third phase of the echo-ranging gear picture, it is worth while to point out the practical applications of the "fixed conditions lateral range curve" calculations. In any tactical situation some of the conditions are reasonably well known, ship speed, sea state, bathythermograph record, and so on. There are other conditions which can be varied by the searching craft, namely the plan of search, i.e., keying interval used, limits of sweep, number of degrees to turn between successive

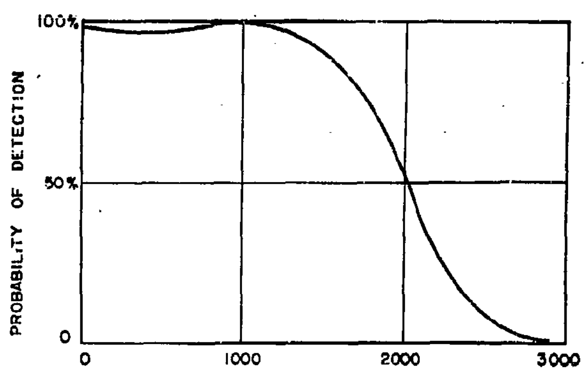


FIGURE 17. Lateral range curve (calculated from one-ping probability).

pings, spacing between ships. By calculating lateral range curves for typical values of the tactical conditions the best values of these controllable variables can be determined for the various tactical situations. In this way it can be shown, for instance, that the sweep should not normally be restricted to a small angle either side of the bow, but should be started at least as far aft as the beam, and should always tire sweep from beam to bow as a single "glimpse" with a fixed probability of detection.

When, however, operational results are to be compared with theoretical predictions, it must be kept in mind that the conditions are by no means fixed. This distinction was discussed in Section 2.9. There is a great variety in the actual conditions—the sonar involved, the sound conditions, and submarine's

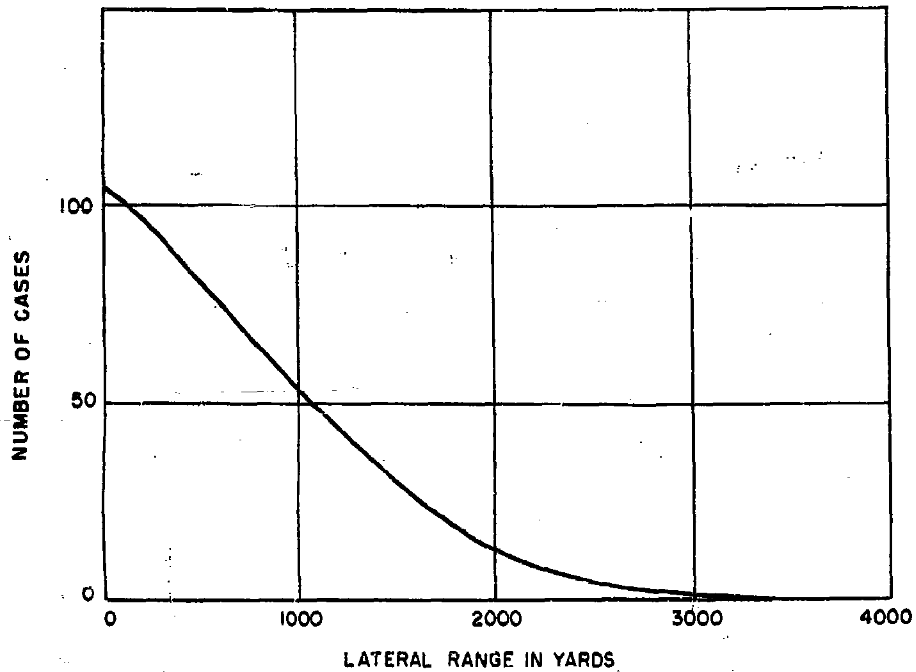


FIGURE 18. Operational lateral ranges for echo-ranging gear.

be made from aft forward rather than in the opposite direction. Many calculations of this sort have been made to determine the proper tactical use of echo-ranging gear in search. Detailed results of these calculations are not of interest at present—the general principle involved is that the gear should be used so as to give uniform coverage of the area without developing serious gaps, as must obviously be the case from Chapter 3.

In many cases calculations have been made using rather rough approximations. The one-ping probability function may be replaced by a simpler one which is zero outside a certain contour and equal to a constant (usually about 0.5) inside. This gives a more abrupt lateral range curve than the more accurate calculations, but most conclusions concerning best operation of the gear are not changed. An even simpler approximation involves considering the en-

depth, speed, and reflecting power. In principle, it is necessary to calculate lateral range curves for all values of the many variable conditions, and then average them appropriately. The labor involved in such a calculation would, however, be completely prohibitive. In addition, it would be necessary to know what distributions of all the important variables are met with in operations, and this information is not available. Consequently we will only present the operational results for comparison with Figure 17, and then determine what range of conditions might give the observed results, and whether this range does, indeed, appear reasonable.

As a fair sample of operational data, 235 first contacts by echo ranging in the period 1 January 1943 to 30 June 1944 can be quoted. From these contacts the number of contacts versus lateral range is plotted in Figure 18.

This curve indicates considerably shorter ranges than that of our previous example which led to the lateral range curve of Figure 17. In other words, the conditions which led to the one-ping probability curve of Figure 14 were more favorable than most of those met with in practice. The operational curve of lateral ranges is, however, of the general type that would result from averaging a number of lateral range curves like that obtained for fixed conditions. Figure 19 above shows a number of lateral range curves of the same general type as Figure 17 which have 50 per cent ranges of 500 yards, 1,500 yards, and 2,500 yards. A combination of these curves in the proportion 5:3:1 leads to a lateral range curve of very much the same shape as that obtained from operational results. In Figure 20 this theoretical curve is compared with an operational curve normalized to give the same probability at 1,000 yards. As can be seen, this calculated curve is in excellent agreement with the observed data. The agreement is largely fortuitous and certainly does not imply any essential correctness of the calculated curves. It does, nevertheless, indicate that the operational results are in keeping with the physical picture if, and only if, it is assumed that short and medium sonar ranges predominate in actual operations. The sweep width obtained from these curves is 1,800 yards, which is smaller than that usually obtained from purely theoretical considerations. This no doubt reflects frequent unfavorable sound conditions, imperfect maintenance of gear, reduced operator skill, and similar factors more or less inevitable under operating conditions.

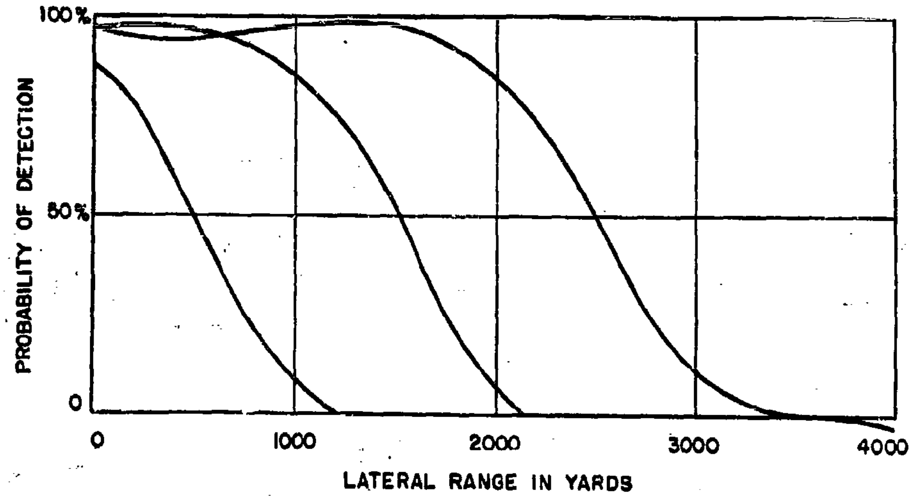


FIGURE 19. Hypothetical lateral range curves.

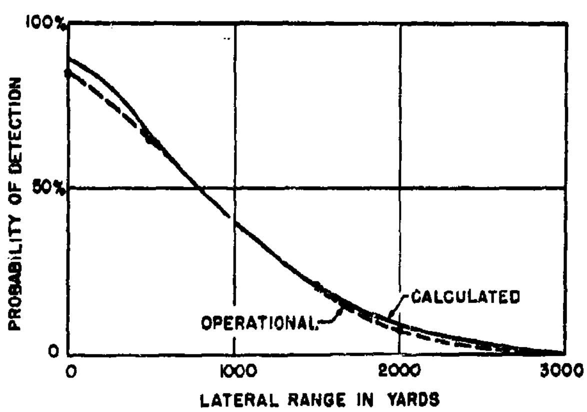


FIGURE 20. Lateral range curves, overall average.

6.4 PARALLEL SWEEPS

When a number of searching units are available, they normally operate together so that their paths in target space form parallel sweeps. For example, a line of sonobuoys which the submarine crosses effectively carry out parallel sweeps relatively to the submarine; so do a group of ships sweeping in line abreast. For such parallel sweeps, the lateral range curves of the individual units must be combined to give an overall probability of detection curve. The manner of doing this depends on the physical situation—whether the lateral range curve arises from variable or fluctuating conditions, or both. This point has been discussed in Section 2.9.

Consider two units making parallel sweeps, spaced S yards, with a submarine penetrating, somewhere between them at point x. Then the probability of detecting the submarine is given by

$$P(x) = p(x) + p(S - x) - p(x)\,p(S - x) \quad (5)$$

if the probabilities of sighting are independent. In some cases, however, the probabilities are by no means independent. In the case of two sonobuoys, for instance, the lateral range curve depends on the distribution of submarine sound outputs. A submarine that is noisy for sonobuoy at A is also noisy for a sonobuoy at B (see Figure 21). To the extent that variable conditions are in each case the same for both searchers, the two probabilities are altogether dependent. In this case the combined probability at x is simply $p(x)$ or $p(S - x)$, whichever is larger.

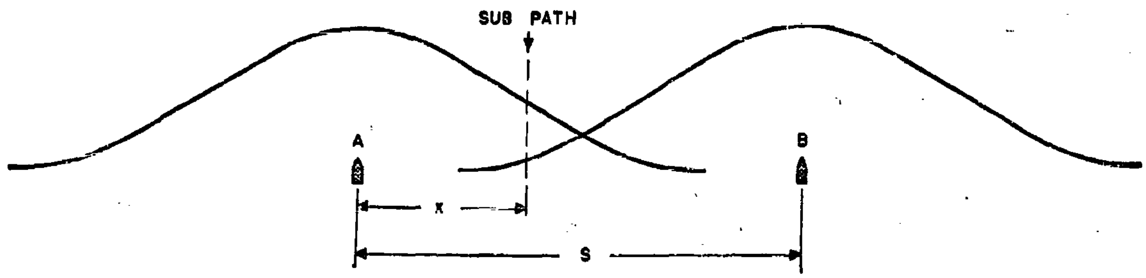


FIGURE 21. Detection probability for two parallel sweeps.

For echo-ranging search under fixed general conditions, the lateral range curve arises from rapid fluctuations in the residual uncontrollable conditions, thus resulting in independent probabilities for the searching ships; consequently the probabilities are combined in accordance with equation (5). In actual fact, conditions are rarely truly "fixed" because the depth of submarine, for instance, is not usually known. Nevertheless, calculations made for fixed conditions of typical values are useful in deciding on proper ship spacing. In doing this, the *assured range* is normally employed. This is defined as follows:

Consider the *maximum* range obtained by standard range prediction methods at a given depth; then take the *minimum* such range as the depth is varied, passing through possible depths, i.e., the maximum range at the most unfavorable depth: this is defined as the assured range. Accordingly, ship spacings based on this range will be rather tight and conservative. The lateral range curve of Figure 17 corresponds to an assured range of 2,000 yards. The derived curves in Figure 22 show the combined probability function for two ships with various ship spacings.

These curves can be interpreted in a number of ways. If the submarine can always choose the best

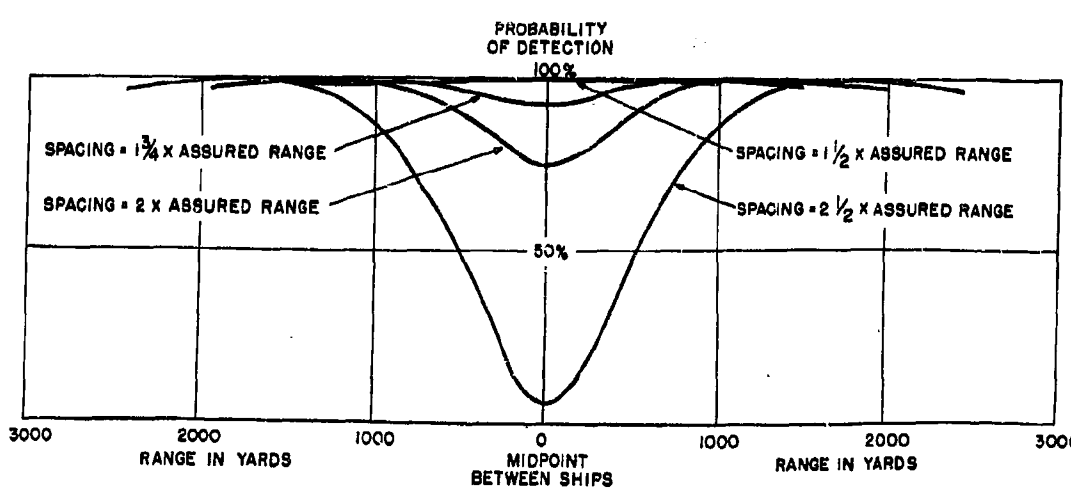


FIGURE 22. Probability versus lateral range for a pair of ships with various spacings.

point to try to sneak through, then it is the minimum value of these curves that counts, i.e., measures the tightness of the screen. If, on the other hand, he passes between more or less at random, then it is the average value which is important. If the number of ships is small, the submarine can often evade them by steering to the side at high speed and passing around the end. This type of evasion is neutralized to some extent by placing a group of ships in line abreast, the resulting front being too broad for the submarine to end-run readily. Increased ship spacing gives a broader front at the expense of holes between the ships.

From the curves in Figure 23 it is apparent that a ship spacing of 1½ times the assured range is very tight, and spacings over 2 times have rather dangerous holes. A complete analysis of the question of optimum ship spacing would involve the size of area to be searched, the submarine's ability to end-run and choose weak points in the screen, and the quantities presented in the curves. As a rough rule of thumb, a spacing of 1¾ times the assured sound range is now specified in doctrine for searches in line abreast, 1½ times (when possible) for screens.

From a practical point of view another factor should be considered. The operational data indicate that rather short ranges predominate in actual operations, and hence that the theoretically predicted "assured range" may be somewhat optimistic. On this basis the rather conservative ship spacing of 1½ to 1¾ times the assured range may be altogether justified.

Evidently the choice of ship spacing for a screen will be different from that for a search (or hunt). The former is defensive, and its primary measure of effectiveness is its ability to intercept submarines which are attempting to penetrate to the proximity of the screened units. The latter is offensive, and its measure of effectiveness is the expected number of submarine contacts which it produces. In view of the overlapping effect of close spacing, this expected number will be reduced, whereas the tightness of the screen will be increased when spacing is close. Thus screens will normally employ tighter spacing than hunts. The end-run prevention cited above is a further argument for wider spacing in the case of hunts than in that of screens. This whole question will be entered into in great detail in Chapter 8.

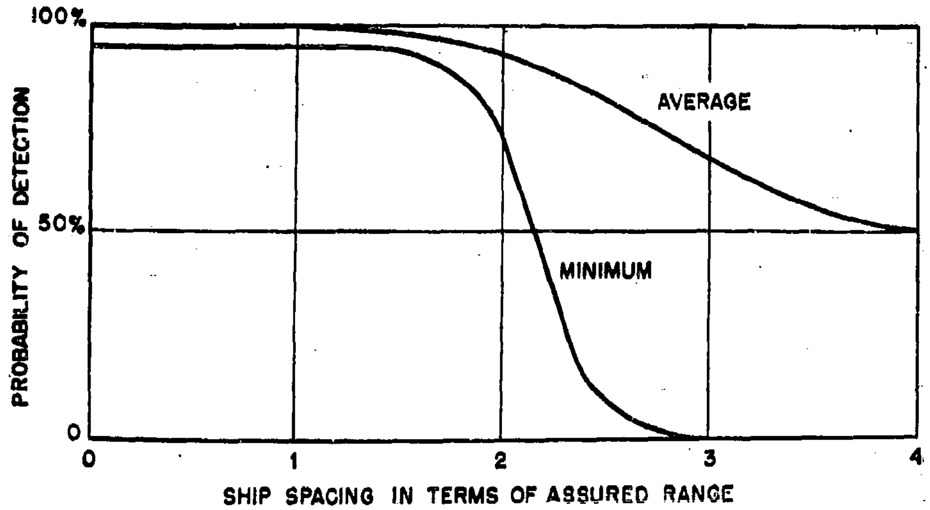


Figure 23. Probability of detection as a function of ship spacing.

Chapter 7

THE SEARCH FOR TARGETS IN TRANSIT

The General Question

In planning a search, the nature of the target is usually known, and its general position may be more or less known as a matter of probability (as in the problems of Chapter 3); but unless a fairly definite estimate of its motion can be made, the plan of search will have to be designed so as to be effective against a target having any one of many different sorts of motion, rather than being particularly effective against targets of some one special kind of motion and less so against others which are recognized as irrelevant to the tactical problem in hand. The emphasis of this chapter is on the latter situation, which arises when both the *intent* and the *capabilities* of the target are known. To know the intent of the target is to know where it is going: through what part of the ocean it passes, from what geographical locality it comes, to what place it is going, etc. And to know the target's capabilities is to have a reasonably good estimate of its speed u, as well as its endurance, etc. An essential part of such information can be put mathematically as follows: The target's vector velocity **u** is known at the different parts of the ocean where it is expedient to conduct the search. And since the main objective is to prevent the target's undetected accomplishment of its intention, success will be achieved even if the target is not detected but is forced to abandon its objective in order to avoid detection.

Attention will be confined in this chapter to the case where the detectability of the target does not change with the time, at least during long periods; thus in the case of visual or radar detection, surface craft (including surfaced submarines) are alone considered; while in the case of sonar detection, the submarine target is regarded as constantly submerged. This avoids the great complication which would occur, for example, in the case of a submarine whose tactics of submergence and emergence are not known and, since they may depend on the tactics of search, could only be evaluated by some form of "minimax" reasoning. Thus "gambits" are not considered herein.

Three cases are of great importance in naval warfare and will be studied in the three parts of this chapter. In the first, the target's intention is to traverse a fairly straight channel (which may be a wide portion of the ocean); the vector velocities at all points are parallel and equal (a "translational vector field"); the search is called a *barrier patrol.* In the second case, the target is proceeding from a known point of the ocean (e.g., a point of fix, an island, or a harbor); the vector velocities are equal in length but are all directed away from this point (a "centrifugal radial vector field"); to this class of search belongs the *trapping square* and the *retiring search* when the approximate time of departure is known, and *closed barriers*, etc., in other cases. In the third case, the target's intention is to reach a definite point (e.g., the seat of a landing operation, an island needing supplies, a harbor); the vector velocities are equal in length but directed inward toward this point (a "centripetal radial vector field"); again the method of countering this intention may be the closed barrier. There are of course various cases closely allied to the three just mentioned, such as the antisubmarine or antishipping hunts conducted by carrier aircraft as the carrier sweeps through the ocean. But when the principles of this chapter are understood, the design and evaluation of such plans offer little difficulty.

In the case where the target intends to reach a point moving on the ocean (a ship, convoy, or task force), the vector field is of an entirely different character; the form of search used is then called a *screen.* It forms the subject of Chapters 8 and 9.

Throughout the present chapter, the effect of target aspect on detection is disregarded. In the case of radar detection, for example, the possibility may exist of securing a somewhat higher chance of detection of targets of certain restricted velocity classes by using specially selected tracks; but the greatly added complication does not appear to warrant their consideration here.

7.1 BARRIER PATROLS

7.1.1 Construction of the Crossover Patrol

Under a wide variety of circumstances, the problem of detecting targets in transit through a channel by means of an observer whose speed v considerably exceeds the speed u of the target (e.g., an airborne observer and ship target) can be simplified to the following mathematical statement. Given a channel bounded by two parallel lines L miles apart (the fine vertical lines of Figure 1) and given targets moving through this channel and parallel to it at the fixed speed u (downward in Figure 1); how shall observers fly from one side of the channel to the other and back, etc., in order to be most effective in detecting the targets? It is usually necessary to attach the flights to a fixed reference point O from which they start or take their direction. Thus O may be a conveniently recognizable point at or near the narrowest part of the actual channel (which will in general correspond only approximately to the mathematically simplified parallel channel shown in Figure 1), O may be a harbor or air base, etc. It is convenient

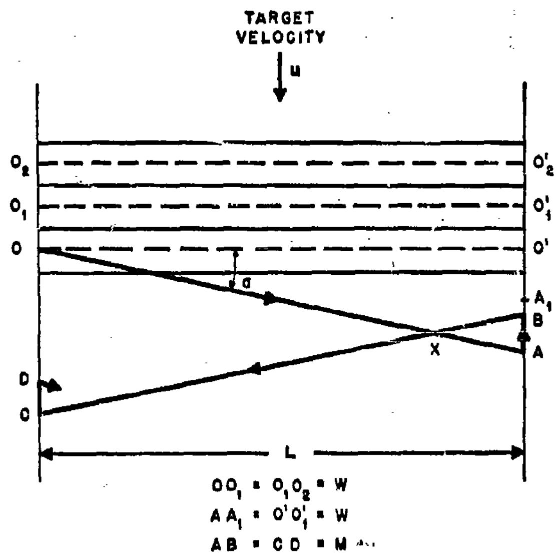


FIGURE 1. The crossover barrier patrol designed to be tight.

to draw the line OO' (dotted in Figure 1) across and perpendicular to the channel, a purely mathematical reference line called the *barrier line*.

The reasoning leading to the construction of a barrier patrol was based, historically, on the definite range law of detection (see Section 2.2; search width $W = 2R$). Since it leads in a natural manner to a form of patrol (the crossover barrier patrol) which turns out to be the optimum form from the point of view of any not-too-asymmetrical law of detection, it will be followed in detail in this section, while its more realistic evaluation will be considered in Section 7.1.2.

For convenience of wording, the target will be referred to as a ship and the observer as an aircraft. While this corresponds to the most important case, others will be considered later. The same mathematical ideas apply in all cases.

Consider those targets which at the initial epoch ($t = 0$) are on the barrier line OO'. An observer starting from O when $t = 0$ and wishing to fly over these targets will not succeed in doing so if he flies along OO', except in the excluded case of targets at rest ($u = 0$). He will have to fly along OA, where the angle $O'OA = a$, called the *lead angle*, is determined by the requirement that the observer reach each point on AB at the same time that the target which was initially on the point of OO' directly above it reaches that same point, i.e., $\sin a = u/v$,

$$a = \sin^{-1}\frac{u}{v}. \tag{1}$$

When the observer flies along OA, he detects not only the targets initially on OO', but, in virtue of the definite range assumption, all those within a distance of $W/2$ miles on either side of OO': the band of width W centered on OO' (Figure 1) is swept clean; i.e., all the targets which may have been in this band when $t = 0$ are detected.

The observer now wishes to detect, on his return flight, all those targets which were, when $t = 0$, in a second band of width W adjoining the first one and directly above it. This band is centered on the line $O_1O'_1$ of Figure 1, where $OO_1 = O'O'_1 = W$. The observer reaches A when

$$t = \frac{OA}{v} = \frac{L}{v\cos a} = \frac{L}{\sqrt{v^2 - u^2}}.$$

At this epoch, a target initially at O'_1, W miles above O', will have moved down to a point A_1, but continue to be W miles above the point A which the target initially at O' will have reached by this time; $AA_1 = W$. This is because if one target is W miles behind another, and if they both have the same speed and course, it will always be W miles behind. Obviously if the observer were to fly directly back to the left bank of the channel, he would not fly over the O'_1 target (now at A_1), and hence not accomplish his purpose of sweeping the $O_1O'_1$ band of targets. To do this, he must fly up the right bank, until he meets this O'_1 target at a point denoted by B and determined by the condition that the time taken for the observer to fly from A to B equals the time taken for the target to go from A_1 to B, i.e., $AB/v = A_1B/u$. This equation together with the fact that $AB + A_1B = W$ determines the *length of upsweep* $M = AB$; solving these equations we find

$$M = \frac{v}{v + u} W. \tag{2}$$

This flight takes $M/v = W/(v + u)$ hours, so that the observer is at B when

$$t = \frac{L}{\sqrt{v^2 - u^2}} + \frac{W}{v + u}.$$

At this epoch, the targets, which when $t = 0$ occupied the $O_1O'_1$ band, will be in a band of width W centered on the line (not shown in Figure 1) through B perpendicular to the channel. From then on the situation is precisely similar to what it was initially: the observer will sweep this band clean if he flies back to the point C of the left bank, where the lead angle of BC has the same value a as before. And having arrived at C, he must make the upsweep $CD = M$ if he is to detect on his third crossing the targets which when $t = 0$ were in a third W width band adjacent to and above the $O_1O'_1$, i.e., the band centered on $O_2O'_2$, Figure 1.

The time taken by the target to fly the *basic element OABCD* is denoted by T_0; the time computations of the preceding paragraph show that

$$T_0 = \frac{2L}{\sqrt{v^2 - u^2}} + \frac{2W}{v + u}. \tag{3}$$

There is another interval of time which it is useful to consider: the time T_t which a target takes in moving from O_2 to O. Since $OO_2 = 2W$ we have

$$T_t = \frac{2W}{u}. \tag{4}$$

Figure 2 illustrates the three possible cases. They are as follows:

Figure 2A. D is *below* O; then T_t is *less* than T_0 (since the O_2 target and the observer reach D simultaneously), and the first crossover point X is to the *right* of the center of the channel (the second, to the *left*, etc.); the flights *if continued by the same aircraft* would take place farther and farther down the channel and thus lead to a *retreating element barrier*. This is the case for which Figure 1 has been drawn.

Figure 2B. D coincides with O; then $T_t = T_0$, and the point X is at the center of the channel and bisects OA and BC; the flights *if continued by the same aircraft* would repeat themselves exactly; the path $OABCD = OABCO$ would be flown over and over again, and thus the barrier would remain stationary; this is called the *symmetric* crossover barrier patrol, or, if one will, the *stationary element barrier*.

Figure 2C. D is *above* O; then T_t is *greater* than T_0, and the first crossover point X is to the *left* of the center of the channel or, in extreme cases, may not occur at all (the second, to the *right*, etc.); the flights *if continued by the same aircraft* would take place farther and farther up the channel and thus lead to an *advancing element barrier*.

It is to be emphasized that all three barriers may be flown as stationary barriers by the device of repeating the elements by having successive aircraft

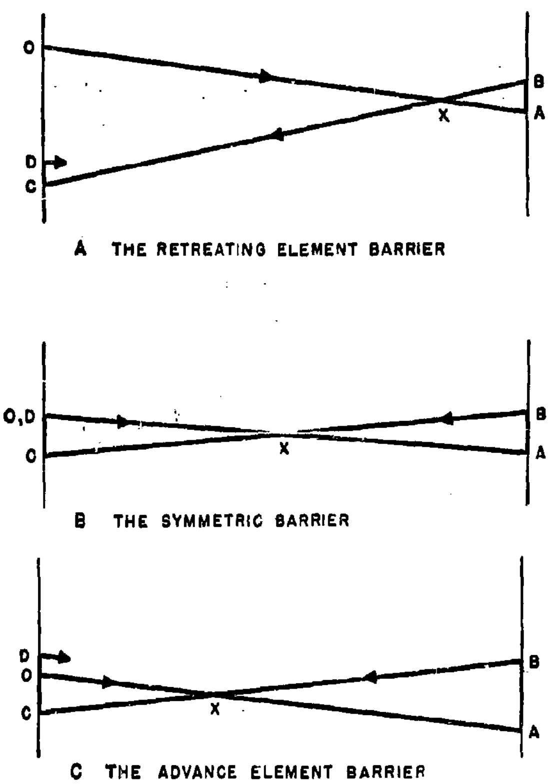


FIGURE 2. The three cases of crossover barrier patrol.

start from O at epochs of T_t after one another: While the elements themselves may retreat or advance, the geographical position of the flights, and hence of the barrier as a whole, remains stationary. This will be illustrated by later examples.

The advancing barrier represents a situation in which more than enough flying is available (assuming the single aircraft's endurance sufficient for the repeated flights) to produce the required coverage. Advantage can be taken of this circumstance to fly only during the favorable hours of the twenty-four, in daylight, if the greatest chance of detection or certainty of recognition is the main desideratum, at night with radar, if the element of surprise is more important; etc. For after advancing the barrier sufficiently during the favorable period, flights can be discontinued and the unswept waters (target positions) can be allowed to come down until their lower boundary reaches OO' (Figure 1), whereupon the flights are recommenced. To find the time of no patrolling, one can reason as follows: If one basic element $OABCD$ is flown, but not a second, the central axis of the first unswept strip (the $O_2O'_2$ strip of $t = 0$) will require the further time $T_t - T_0$ to reach OO'; if N basic elements are flown, the further time $N(T_t - T_0)$ to reach O. However, if the last upsweep is not flown, it will require $M/v = W/(v + u)$ more time, since the aircraft completes its patrol flights that much earlier. Hence the interval of time (after the aircraft reaches the left bank for the last time) during which no patrol flights (as distinguished from the return flight to base) need be flown is $N(T_t - T_0) + W/(v + u)$. With the aid of (3) and (4) we obtain the expression.

No patrolling period =

$$N\left[\frac{2vW}{u(v+u)} - \frac{2L}{\sqrt{v^2 - u^2}}\right] + \frac{W}{v+u}. \qquad (5)$$

The symmetric barrier represents a situation in which the continued flying of the single aircraft is exactly enough to ensure the required coverage. It has a great advantage of simplicity over the asymmetrical cases; and the measures described later for making it applicable are frequently taken.

The retreating barrier represents a situation in which a single aircraft, even if its endurance would permit it to fly a large number of basic elements, is insufficient to maintain the required coverage, since the unswept area invades positions farther and farther down the channel. Under these conditions its flight has to be supplemented by that of other aircraft. One obvious way of doing this is to have a second aircraft leave O at the time when the target initially at O_2 reaches O, i.e., when $t = T_t - 2W/u$. This will be $T_0 - T_t$ hours before the first aircraft reaches D, and still longer before it returns to the base O. The second aircraft flies $OABC$, etc. But a better plan is to have n aircraft fly simultaneously abreast in a line perpendicular to their course and at the distance W apart. This has the effect of increasing the search width to the value $W' = nW$, and thus, if n is sufficiently great, leads from a retreating element barrier to a symmetric or an advancing one. Let us assume that with one aircraft $T_t < T_0$, i.e., the barrier element is retreating. What is the least number n of aircraft flying as described which will give rise to a nonretreating one? The answer is found by imposing the condition $T_t \geqq T_0$ and replacing T_0 and T_t by their expressions in (3) and (4) *in which W has been replaced by $W' = nW$*. We have

$$\frac{2nW}{u} \geqq \frac{2L}{\sqrt{v^2 - u^2}} + \frac{2nW}{v+u},$$

which is transformed algebraically so as to give the condition

$$n \geqq \frac{L}{W}\frac{u}{v}\sqrt{\frac{v+u}{v-u}}. \qquad (6)$$

Since n is an integer, it is taken as the lowest integer greater than or equal to the expression on the right, which is in general not an integer. Thus the number of aircraft is proportional to the width of channel and inversely proportional to the search width; and when v is so much greater than u that the radical can be regarded as unity, it is proportional to the target's speed and inversely proportional to the observer's speed.

It is remarked that when the width of channel L is not overwhelmingly greater than the search width W, an attempt is sometimes made to base the cross-over patrol not on the boundary lines of the channel as in Figure 1 but on two lines parallel to them and at a distance $W/2$ to the right of the left-hand boundary and $W/2$ to the left of the right-hand boundary, respectively. But all the discussion and formulas previously given apply to this case, provided L is replaced throughout by $L' = L - W$. Thus (6) becomes

$$n \geqq \left(\frac{L}{W} - 1\right)\frac{u}{v}\sqrt{\frac{v+u}{v-u}}.$$

And we derive all the formulas needed to consider the altered channel case together with the case of n aircraft abreast, simply by replacing W by nW and L by $L - W$ in formulas (2) and (5). Thus, the length of upsweep formula (2) becomes

$$M = \frac{v}{v+u} nW,$$

and the conditions for cases of Figures 2A, 2B, or 2C are that $T_t < T_0$, $T_t = T_0$, or $T_t > T_0$ take on the forms

A. $$\frac{nW}{u} < \frac{L - W}{\sqrt{v^2 - u^2}} + \frac{nW}{v+u},$$

B. $$\frac{nW}{u} = \frac{L - W}{\sqrt{v^2 - u^2}} + \frac{nW}{v+u},$$

C. $$\frac{nW}{u} > \frac{L - W}{\sqrt{v^2 - u^2}} + \frac{nW}{v+u}.$$

We have seen in Figure 1 how the bands swept clean according to the definite range law exactly cover the channel without overlapping or holes. That a plan of parallel flights, which, being intuitively simpler, might be tried instead of crossover flights, produces overlapping and holes and thus a loss both of efficiency and effectiveness, is illustrated in Figures 3A, 3B, 3C, 3D, and 3E. The geometric object lesson contained in these figures was of historical decisiveness in an important operation of World War II. Here the problem was to set up a barrier effective against a 24-hour run at about 12 knots of enemy blockade runners in a known direction.

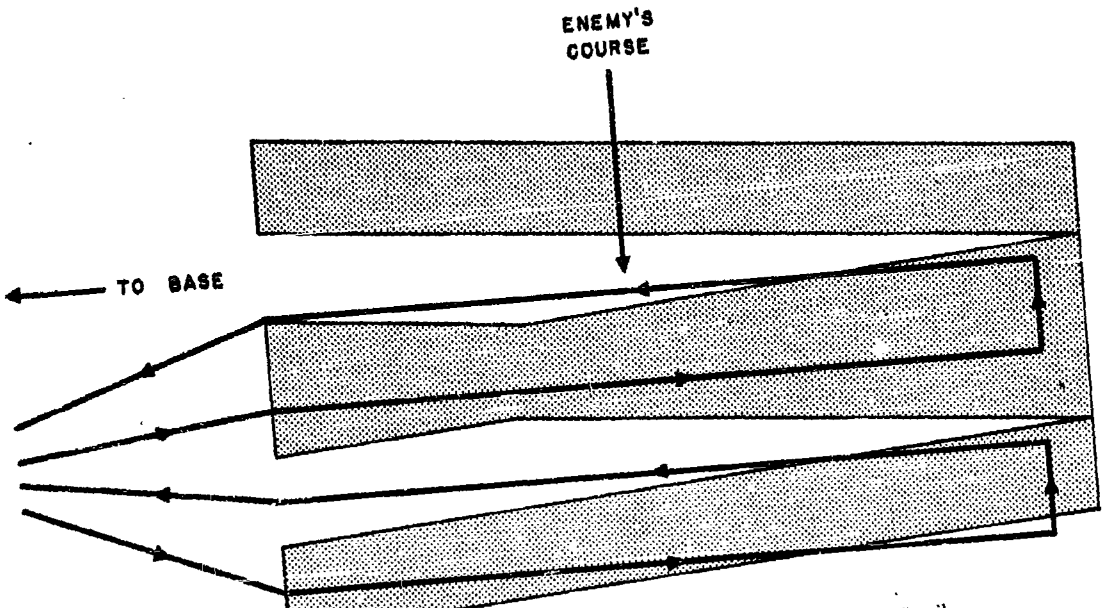
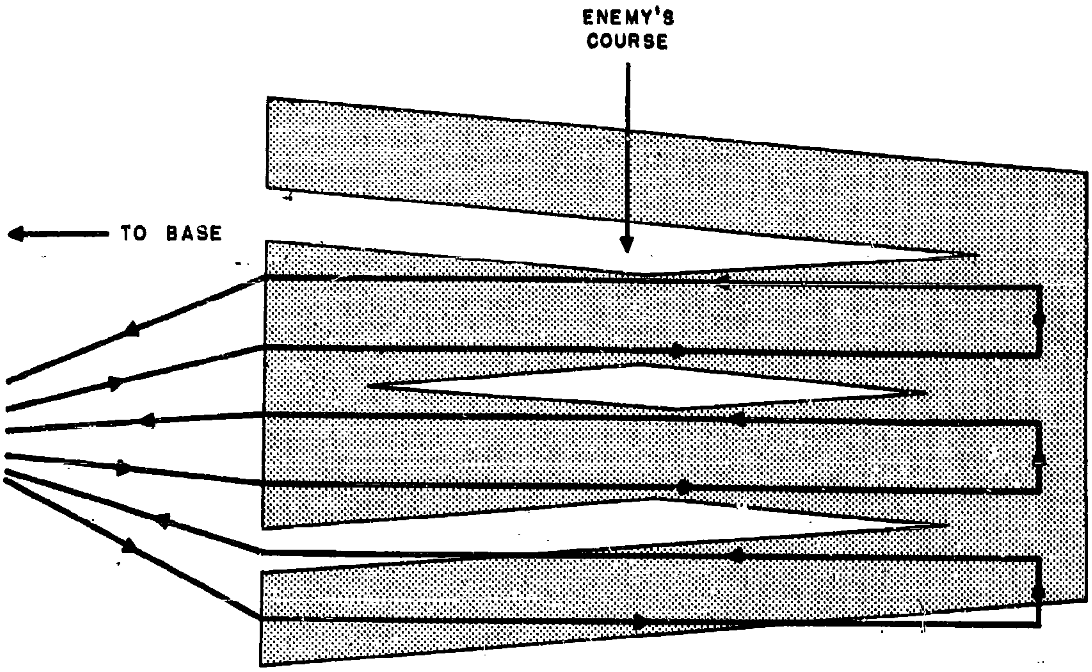
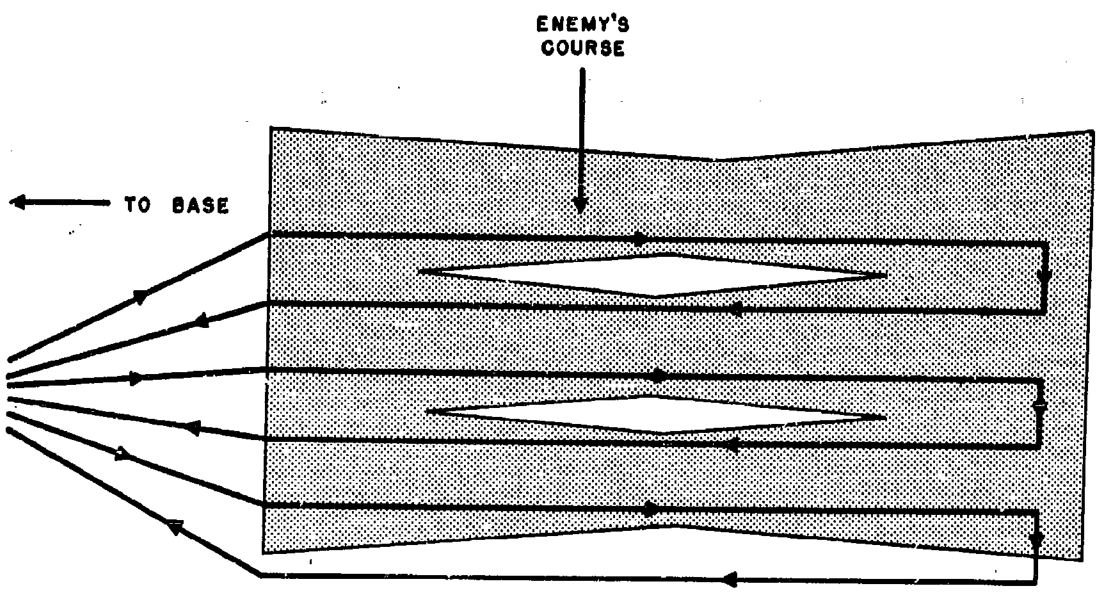


FIGURE 3A. Ineffectiveness of parallel search courses as barrier (legs separated by twice detection range).



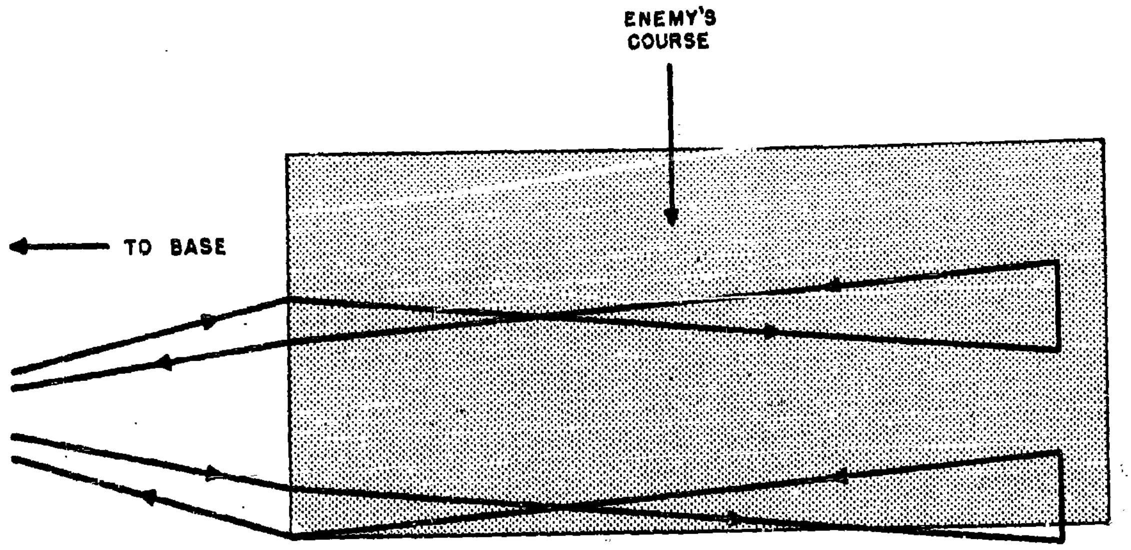
speed of plane 150 knots
speed of ship 11.7-12 knots
reliable detection range 35 miles
line patrolled 600 miles
plane separation 52½ miles (3 planes)

FIGURE 3B. Ineffectiveness of parallel source courses as barrier (legs separated 1.5 times detection range).



speed of plane 150 knots
speed of ship 11.7-12 knots
reliable detection range 35 miles
line patrolled 600 miles
plane separation 52½ miles (3 planes)

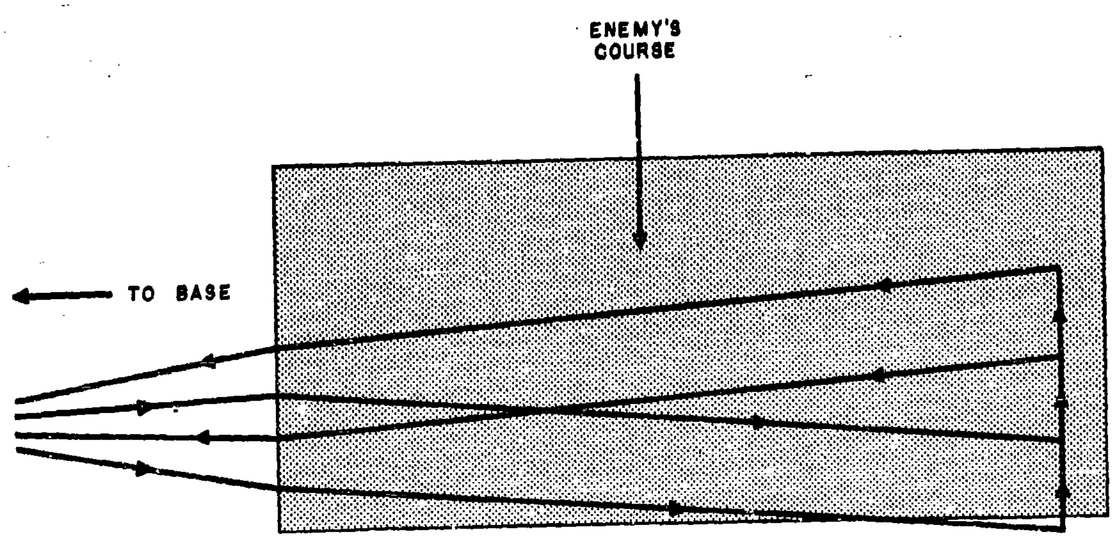
FIGURE 3C. Ineffectiveness of parallel search courses as barrier (direction of flight reversed).



speed of plane 150 knots
speed of ship 11.7-12 knots

reliable detection range 35 miles
line patrolled once daily 600 miles

FIGURE 3D. Barrier patrol without holes.



speed of plane 150 knots
speed of ship 11.7-12 knots

reliable detection range 35 miles
line patrolled once daily 600 miles

FIGURE 3E. Preferred type of barrier patrol.

7.1.2 Crossover Barriers with Any Law of Detection

The assumption of a definite range law of detection, while affording a convenient basis for the construction of barrier patrols, leads to one fallacious conclusion, namely, that if designed as in Section 7.1.1 it is absolutely tight, providing a 100 per cent chance of detecting the target; and, on the other hand, if it is designed with a slight overestimation of the search width W, it has holes. This in turn may lead to the practically disadvantageous procedure of exerting great effort to basing the barrier on a preconceived value of W, believing that nothing short of this tightness is adequate, while at the same time having a false sense of security when this ideal is achieved. The matter at issue is, in other words, just what it was in Chapter 2, when the importance of considering various more realistic laws of detection was stressed. The fact is that no barrier is 100 per cent tight, whereas one falling far short of the ideal of the cleanly swept adjacent strips of Section 7.1.1 may have very real value; it may provide a very useful probability of detection. Thus if a barrier detects on the average even one-quarter of all targets, it will make it a very dangerous and costly procedure for the enemy to send his shipping through the channel.

In order to apply the machinery of Chapter 2 to these barrier patrols, we shall consider how the flights of Section 7.1.1 appear in space relative to the targets, i.e., in a horizontal plane moving down the channel with the speed u, "u-moving space," as we shall say for brevity. In such a space, all the targets are fixed. Thus the lines of targets OO', $O_1O'_1$, $O_2O'_2$, etc., are stationary horizontal lines, maintaining for all values of t their positions as shown in Figure 1 for $t = 0$. And since, as we have seen, an observer flying the basic element $OABCD$ of the plan passes directly over the OO' and O'_1O_1 targets in the order of these letters, as well as up segments of the boundaries of the channel in passing from the O' target to the O'_1 target, and from the O_1 target to the O_2 target, the basic element flight will be along the path $OO'O'_1O_1O_2$ of u-moving space, as shown in Figure 4. And as long as the crossover patrol is flown, more and more of the horizontal lines and their connecting segments of Figure 4 will be traversed.

But this is simply the case of detection of a randomly placed stationary target by means of parallel sweeps, the problem considered in Section 2.7. The fact that the distance between parallel tracks comes out as W in Figure 4 is merely a consequence of the assumption of the definite range law made in Section 7.1.1 above, along with the desire of making the barrier 100 per cent tight. The point to be emphasized is that *the crossover barrier patrol gives the best distribution of flights—uniformly spaced parallel sweeps*; this is why, in spite of rejecting one part of the assumptions upon which its design was based (definite range law), the form of patrol is still regarded as optimum. But in the interest of flexibility, to provide coverage by parallel sweeps when a sweep spacing of $S = W$ cannot be used, one must get away from this particular value of S.

Let the sweep spacing S be an arbitrarily chosen constant, one chosen without necessary relation to W or any other parameter of visibility. Is it possible to

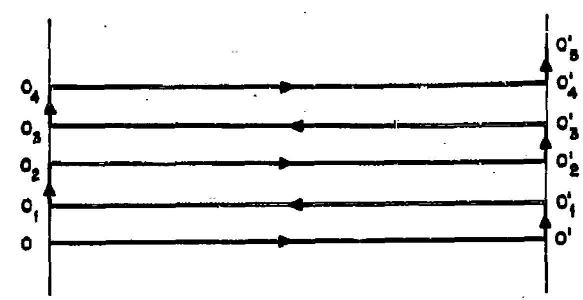


FIGURE 4. The barrier flights relative to the targets (u-moving space).

fly in such a way that the observer's path in u-moving space be of the form shown in Figure 4, but with

$$OO_1 = O_1O_2 = O_2O_3 = \cdots = OO'_1 = O'_1O'_2 = O'_2O'_3 = \cdots = S, \tag{7}$$

and how are the flights to be described in the geographical space of Figure 1?

The answer is simple: Fly the same crossover paths as in Section 7.1.1, only with W replaced throughout by S. For the relation between the geographical and the relative paths depends in no wise on the fact that S in Section 7.1.1 had the value W. The horizontal bands centered on OO', $O_1O'_1$, etc., are now of width S, and merely lose their meaning of "regions swept clean." Moreover, all the formulas of Section 7.1.1 apply to the present case, provided W is replaced throughout by S. For convenience we shall write them down in the new (general) form here.

Length of upsweep for one and for n observers abreast spaced S miles apart,

$$M = \frac{v}{v+u} S \quad \text{(one observer)}$$
$$M = \frac{v}{v+u} nS \quad (n \text{ observers abreast}). \tag{8}$$

Time of flight of one basic element,

$$T_0 = \frac{2L}{\sqrt{v^2 - u^2}} + \frac{2S}{v+u} \quad \text{(one observer)}$$
$$T_0 = \frac{2L}{\sqrt{v^2 - u^2}} + \frac{2nS}{v+u} \quad (n \text{ observers abreast}). \tag{9}$$

Time taken for first target (on left bank) not flown over in the flight of the first basic element but which will be the lowest one flown over in the second basic element to reach O:

$$T_t = \frac{2S}{u} \quad \text{(one observer)}$$
$$T_t = \frac{2nS}{u} \quad (n \text{ observers abreast}). \tag{10}$$

Time of no patrolling after N basic elements (the last one without upsweep) are flown by n observers abreast, from the end of the flights until the patrol flights must be resumed in order to maintain the uniform coverage of sweep spacing S in u-moving space:

No patrolling period

$$= N\left[\frac{2vnS}{u(v+u)} - \frac{2L}{\sqrt{v^2 - u^2}}\right] + \frac{nS}{v+u}. \tag{11}$$

This assumes an advancing barrier, and no account is taken of the time for the aircraft to return to base.

As was seen in Section 7.1.1 the barrier element will be of the (A) retreating, (B) stationary (symmetrical), or (C) advancing types, according as $T_t < T_0$, $T_t = T_0$, or $T_t > T_0$ respectively. Using equations (9) and (10) and transforming the results algebraically, the following criteria are derived in the case of n aircraft flying in line abreast at spacing S.

Writing $k = \sqrt{\frac{v+u}{v-u}}$

the condition for

A. Retreating barrier is $n < k \frac{L}{S}\frac{u}{v}$.

B. Stationary barrier is $n = k \frac{L}{S}\frac{u}{v}$. (12)

C. Advancing barrier is $n > k \frac{L}{S}\frac{u}{v}$.

The method of evaluating the probability $P(S)$ of detecting a particular target attempting to cross the barrier is given in Section 2.7, equation (43) in particular. Figure 12 of Chapter 2 shows the relation between this probability and S (although the abscissa is the sweep density "n" $= 1/S$) in typical cases. If in particular an inverse cube law of sighting is assumed, $P(S)$ is given by formula (47) of Chapter 2

$$P(S) = \text{erf}\left(0.954 \frac{E}{S}\right), \tag{13}$$

E being the effective visibility. In general, $P(S)$ has to be derived by approximate formulas or graphical methods, in connection with the material set forth in Chapters 4, 5, and 6.

It may be remarked that if the target's speed u has been overestimated, the barrier flights will appear relatively to the target not as the horizontal lines of Figure 3 but as two sets of parallel lines, one set, corresponding to flights from the left bank to the right, being tipped slightly down to the right, the other set, corresponding to flights in the reverse direction, tipped down to the left. Also, all upsweep legs will be a trifle shortened. And thus the flights, while not giving as regular a picture relative to the target, will give one of more crowded paths, and hence the chance of detection will be *increased*. It will, however, not be as great as it would have been had the searcher planned his flights for the correct value of u.

If the target is moving obliquely down the channel, instead of exactly parallel to the banks as we have been assuming, the effective value of its speed as far as the tightness of the barrier is concerned is its downward component. This again tends to reduce the effective speed, with the result noted above.

There are, generally speaking, two methods of applying the results derived here. *First*, the probability $P(S)$, i.e., the tightness of the barrier, can be given; required the number of aircraft needed to maintain a patrol of a given sort (e.g., a symmetric one), the corresponding length of upsweep M (which determines the basic element), etc. *Second*, given the number of aircraft and type of patrol, find the probability of detection. In the first case, S is derived from $P(S)$ by an equation like (13), n is determined by (12), e.g., (B) or (C), then M is calculated from (8). In the second case, S is determined from (12), and then M and $P(S)$ from (8) and an equation like (13). But there are various mixed cases; the situation is illustrated in Section 7.1.3.

We are now in a position to solve the problem, foreshadowed in Section 7.1.1, of the use of an advancing element barrier to avoid 24-hour flights. Suppose that for A hours out of the 24 it is expedient to fly the patrol, whereas during B hours it is inexpedient; B might be the hours of darkness in cases where much dependence is placed on visual detection or recognition; $A + B = 24$. Let us try to determine the number n of aircraft patrolling abreast and the number N of circuits they must patrol together so that coverage at the sweep spacing S in u-moving space be maintained constantly. For any given N and n, the time of no patrolling is given by (11), while the total patrolling time is $NT_0 - M/v$ (the $-M/v$ term, because the last upsweep is omitted). The sum of these two periods of time must equal 24 hours; using (8), (9), and (11), one finds that

$$nN = \frac{12u}{S}. \tag{14}$$

Now since the permissible no-patrolling time must be at least as great as B, (11) leads to the inequality

$$N\left[\frac{2vnS}{u\,(v+u)} - \frac{2L}{\sqrt{v^2 - u^2}}\right] + \frac{nS}{v+u} \geqq B. \tag{15}$$

By multiplying this through by $N(v + u)$, introducing the quantity k of (12), and replacing nN by its value given in (14), one derives the quadratic inequality

$$2LkN^2 - (Av - Bu)N - 12u \leqq 0,$$

which (on completing the square, etc.) is readily shown to be equivalent to the following:

$$N \leqq \frac{1}{4Lk}\left[Av - Bu + \sqrt{(Av - Bu)^2 + 96Lku}\right]. \tag{16}$$

It is noticed that this inequality does not involve S. And now we have the conditions (14) and (16) in all respects equivalent to (14) and (15) (one set is a necessary and sufficient condition for the other): Thus (14) and (16) are the necessary and sufficient conditions which N and n must satisfy to be solutions of the problem. They do not, however, fully determine N and n. The method for doing this is as follows: *First*, choose for N the largest integer satisfying (16). *Second*, choose a value for S which on the one hand makes $12u/NS$ an integer, and on the other hand gives an acceptably high value to $P(S)$ [by the use of (13)]. This last involves some trial and error; moreover, a value of S which has these properties may turn out to impose too great force requirements, and so one may have to be satisfied with a somewhat lower probability $P(S)$, i.e., use a larger S. When the value of S has finally been chosen, (14) gives n as a positive integer, and the problem is solved. It will be illustrated in Section 7.1.3.

7.1.3 Practical Applications

The following examples are illustrative:

1. It is desired to close a 600-mile channel by a barrier giving a 90 per cent chance of detection. The speeds are $v = 130$ knots and $u = 12$ knots. The effective visibility is 20 miles, and equation (13) is assumed.[a] How many aircraft are needed in order to have a symmetric element barrier, and how should the flights be specified?

Using (13), $\text{erf}(0.954 \times 20/S) = 0.9$, and it is found from a table of error functions that $\text{erf}(1.163) = 0.9$, hence $0.954 \times 20/S = 1.163$, so that $S = 16.4$. Since the inequality (A) of (12) must not hold, the number n of aircraft must be the smallest integer, not less than

$$\frac{L}{S}\frac{u}{v}\sqrt{\frac{v+u}{v-u}} = \frac{600 \times 12}{16.4 \times 130}\sqrt{\frac{142}{118}} = 3.7;$$

in other words, four aircraft are necessary. But with four aircraft, case (C) of (12) is in effect, not case (B): The barrier advances up the channel. To have a stationary one, the four aircraft may fly closer together than S, by an amount determined by solving the equation

$$4 = \frac{600 \times 12}{130S}\sqrt{\frac{142}{118}},$$

i.e., equation (12), case (B). We obtain $S = 15.15$ miles. With this reduced spacing, the barrier gains in tightness; in fact the probability found from (13) now becomes a 92.5 per cent chance—all to the good. The length of upsweep given by (8) is $M = 55.5$ miles. Finally, the lead angle given by (1) is $a = 5° 18'$. These quantities determine the fundamental element, or rather elements, as four congruent symmetrical crossover paths flown shown in Figure 5.

But in determining force requirements it is not

[a] The value of 20 miles is too low for ships but is about correct for surfaced submarines.

sufficient to have only the fundamental flights given; we must find how long their flying takes and consider questions of aircraft endurance. It is found from (9) (with $n = 4$, $S = 15.15$, etc.) that $T_0 =$ 10.14 hours. Now only a long-range patrol aircraft such as a PBM or PB4Y could have an endurance

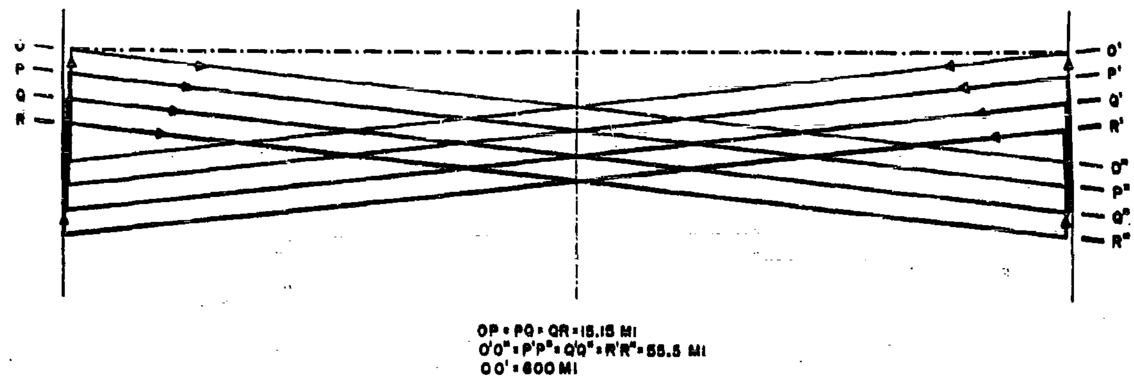


FIGURE 5. Symmetrical barrier flown by four aircraft abreast.

sufficient for this one circuit of the fundamental element; it will have to be capable of well over 10 hours, since time must be allowed for investigation of contacts, the trip to and from base (which may not be at O, P, Q, R) etc. And it cannot be expected to make more than one circuit. Thus a new flight of four fresh aircraft must be readied and waiting at O, P, Q, R in order to take up the flights as soon as the first set return to these points. The operation will therefore require eight aircraft, assuming maintenance to be quick and perfect. Actually, a few more should be on hand, as well as enough pilots and lookouts to ensure their being well rested at the outset of every new flight—an essential condition for their efficient operation, without which the effective visibility will fall far short of the assumed 20 mile figure.

2. Under the assumptions of the last example, let it be required to fly a barrier of the advancing type with n aircraft abreast during the $A = 12$ hours of light, to be discontinued during the $B = 12$ hours of darkness. Applying the method at the end of Section 7.1.2, we obtain from (16) that $N \leqq 1.16$, and hence we take $N = 1$. Next we must take an S which gives $12u/NS = 144/S$ an integral value and provides an acceptable probability $P(S)$, while at the same time not using an undue number of aircraft. We have seen that the value 16.4 gives a 90 per cent chance of detection, and this suggests taking $S = 16$, for which $144/S = 9$ and $P(S)$ given by (13) is a probability of nearly 91 per cent. Then $n = 9$: nine aircraft must be flown spaced at 16 miles abreast, one flight only being made per day. The time of patrolling is $T_0 - M/v$ (the last upsweep being omitted) which has the value [given by (8) and (9)] of 10¼ hours, the remaining 13¾ hours (including the 12 of darkness) not needing any patrol. The fact that the two periods are not each equal to 12 hours is of course due to the circumstance that since N and n are integers, (14) and (15) cannot be solved as equations but rather (14) as an equation and (15) as an inequality. The operation thus requires about the same forces as in the previous example. Which of the two methods is to be used depends on considerations of equipment (how good the radar is for search at dark, etc.) and tactics.

3. A channel 300 miles wide is to be barred by a symmetrical barrier flown by one aircraft of 6-hour endurance. How tight is the barrier, and how frequently does the aircraft have to be relieved? Assume again $u = 12$, $v = 130$, $E = 20$, and equation (13).

From equation (12), case (B), with $n = 1$, we obtain $S = kLu/v = 30.4$. From equation (13), $P = 0.644$: a 64.4 per cent chance of detection. Eliminating S from equations (9) and (12), case (B), we have

$$T_0 = \frac{2Lk}{v}, \tag{17}$$

which, in the present case, gives $T_0 = 5.06$ hours. Evidently with an endurance of 6 hours, just one circuit can be flown, so we shall require about five flights a day, and between two and five aircraft available at the very minimum.

7.1.4 Barrier When Target Speed Is Close to Observer's Speed

So far it has been assumed that v considerably exceeds u; indeed, when $v \leqq u$, the crossover type of barrier is kinematically impossible. This is no obstacle when the observer is airborne and the target is a ship, but when both observer and target are units of the same type (both ships or both aircraft), the situation excluded heretofore becomes important. Although many plans of barring a channel can be devised for this case, attention will be confined here to the very simple case in which the observer moves back and forth across the channel on a straight path perpendicular to its (parallel) banks: such a patrol is always possible and its design evidently does not involve the speed ratio u/v.

This back-and-forth barrier will be compared with the symmetrical crossover (when $u < v$), and since only a rough comparison is sought here, the definite range law will be assumed (range $= R$ in each case). A more accurate detection law is not likely to alter the comparison appreciably.

The two diagrams on Figure 6 show the geographic as well as the relative tracks for the two types of patrol.

In each relative track a half cycle has been selected and the area swept shaded. The probability of detection for each case has been taken as the ratio of the shaded area to the total area in the channel between the two dashed lines marking off the half cycle. It is convenient to introduce two new variables to describe the probability of detection, $r = v/u$ and $\lambda = L/W$. For the case of the crossover patrol, the probability $P_\times$ of detection is given by

$$P_\times = \min\left[1, \left(1 + \frac{r\sqrt{r^2 - 1}}{r + 1}\right)\frac{1}{\lambda + 1}\right].$$

For the back-and-forth patrol the probability P of detection is given by

$$P_{\leftrightarrow} = \begin{cases} 1 - \left(\lambda - \dfrac{\sqrt{r^2+1} - 1}{2}\right)^2 / \lambda(\lambda+1) & \text{for } r \leqq 2\sqrt{\lambda(\lambda+1)} \\ 1 \quad \text{for } r > 2\sqrt{\lambda(\lambda+1)}. \end{cases}$$

In Figure 7 the values of P for the two cases are plotted as functions of r with λ kept fixed for a given curve. In comparing crossover patrols with back-and-forth patrols, curves bearing the same value of λ

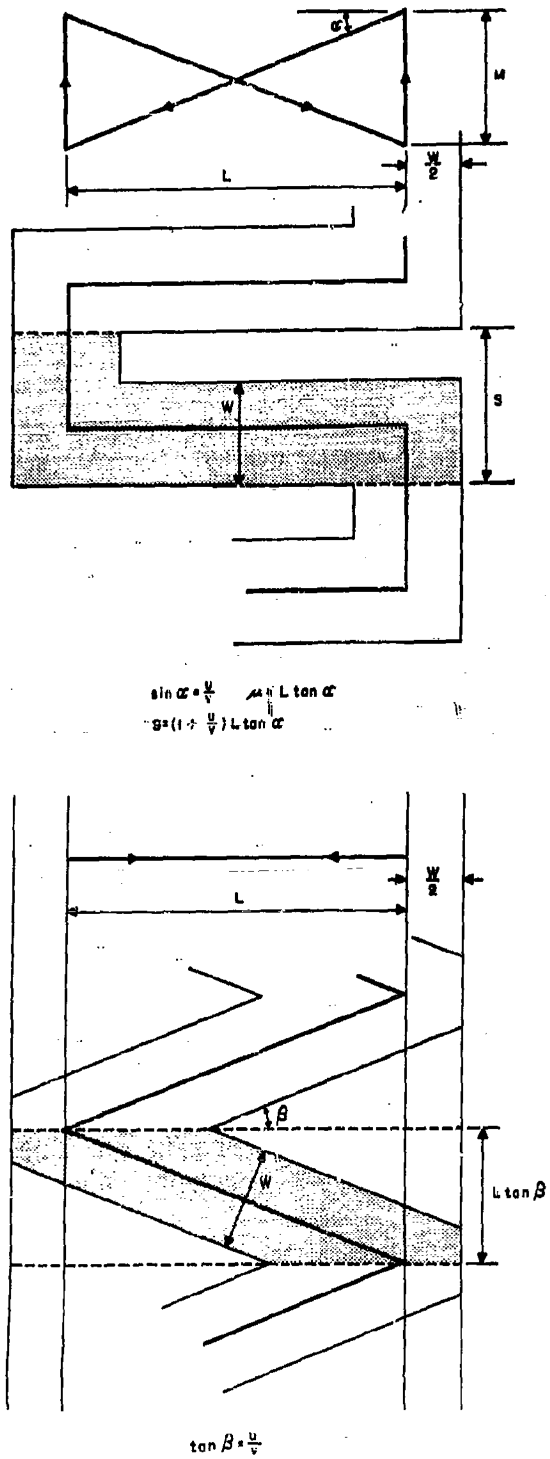


FIGURE 6. A comparison of barriers.

should be compared. The solid curve passes through the points of intersection of the curves being compared and marks the boundary between the regions

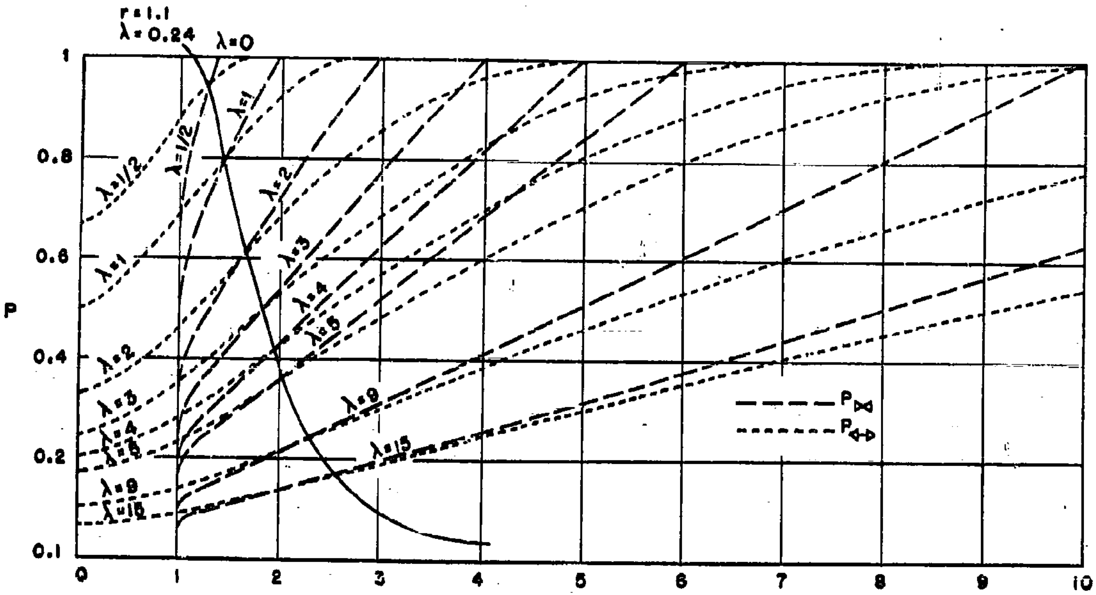


FIGURE 7. The comparative effectiveness of back-and-forth and crossover plans.

where back-and-forth is preferable and where crossover is preferable.

In order to facilitate the selection of the preferable type of patrol, Figure 8 is included. This curve shows the relation between λ and r for the points of intersection of curves in Figure 7.

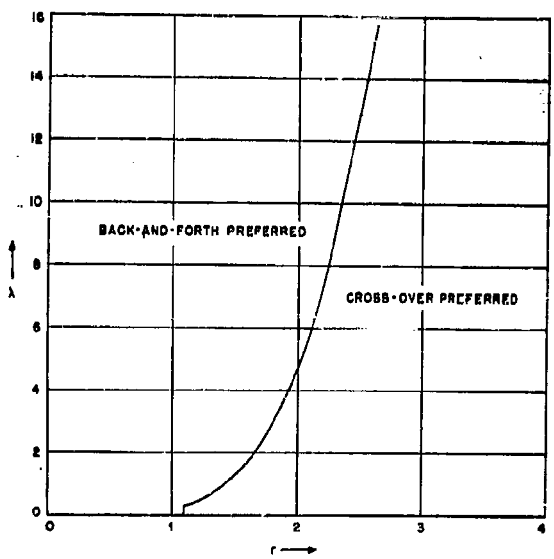


FIGURE 8. Regions of effectiveness of back-and-forth and crossover plans.

An example will illustrate the use of the curves. Suppose a ship making 12 knots is trying to prevent undetected penetration of a barrier by a 6-knot submerged submarine. Assume further that the channel being guarded is 8 miles wide and that the sonar search width W is 2 miles. Then $L = 7 - 2 = 5$, $\lambda = 5/2 = 2.5$, $r = 12/6 = 2$. Entering Figure 8 with these values for r and λ one discovers that a crossover patrol is preferred.

7.2 CIRCULAR BARRIERS

7.2.1 Constant Radial Flux of Hostile Craft

In the case where enemy surface craft or submarines are attempting to leave a point of the ocean, such as an island or exposed harbor, and in the case in which they are attempting to approach such points or to close positions at which our forces are conducting landing operations, the vector velocity pattern is a radial one: it is "centrifugal" (directed away from the central point) in the former case and "centripetal" (directed toward the center) in the latter. But in each case it can be regarded as *constant in time*: over long periods, the density of outgoing or incoming craft is not expected to vary. This sets the present situation in strong contrast with that considered in Section 7.3, in which the unit to be detected is, to be sure, proceeding radially away from a point of fix, but in which its likelihood of being at various distances from this point depends strongly upon the time elapsed since the fix. Corresponding to this latter circumstance, the layout of the search provides for a progressive variation of the searched positions with lapse of time; the theory of such plans is far more complicated than that of the ones considered at present.

7.2.2 Targets Moving Toward a Central Objective

The type of operation considered here may be contrasted with that of protecting a convoy against submarine attack considered in Chapters 8 and 9. The defended position is in the present case at rest; and thus the submarines have no tracking problem and, moreover, can approach at any speed at which they find it convenient to operate, and from any relative bearing. The defended position is consequently exposed to attack in any direction from which it is not effectively shielded by land masses or shoal water. As in the case of convoy escort, the protection given will be of two kinds. Aircraft will engage in barrier patrols outside distances within which wholly submerged approach is possible (more than 60 miles out with the submarines of World War II). An intermediate aircraft screen may be provided to pick up submarines which may be surfaced at distances under 60 miles, but this screen will be less important. An inner screen of surface ships will patrol a barrier for submarines (submerged or surfaced) which may elude the aircraft screen. Details of each screen follow.

It is advantageous to abandon the design described in Section 7.1 in favor of a simpler plan. The lead-angle a which was a key element of the basic design given in Section 7.1 is required only because of the necessity to search parallel strips of space moving with enemy velocity u, patrolling with own velocity v, when the direction of search was *necessarily opposed* in alternating members of each pair of strips. If search of adjacent strips is always carried out in the same direction there is no need to use a lead-angle at all. The resulting searched strips in u-moving space will still be parallel but they will be inclined to the target track at an angle $(90° - \tan^{-1} u/v)$ instead of at 90 degrees.

When the barrier is set up *around* an objective the closed barrier path may be made to contain the objective, and the lead-angle a can become zero. The resulting path is a simple circle, or more practically, a square. Such a path is much easier to navigate, or to evaluate, than a set of barriers set up on barrier lines which form the sides of a polygon.

The size of the square will usually be determined by tactical considerations, i.e., the distance at which interception should be achieved. For example, against submarines this may signify the distance within which approach while wholly submerged becomes possible. The distance the target can travel while the searcher makes one circuit, or between sweeps by equally spaced searchers on the same square track, becomes the track spacing S in u-moving space. For the simple case of a circular track this is

$$S = \frac{2\pi r}{n} \cdot \frac{u}{v} \tag{18}$$

where r is the radius, and n the number of search craft employed. Since with a given effective visibility there is a contact probability corresponding to any value of S, formula (18) determines the number of searchers required to give any desired tightness.

For the more practical case of a square track of side b

$$S = \frac{4b}{n} \cdot \frac{u}{v}. \tag{19}$$

This value of S corresponds to the *minimum* contact probability since with a square the target track will not always cross the search track at 90 degrees.

There is one important difference between this type of barrier and certain cases of the crossover type. S for a given size of square is determined wholly by relative target movement and the number of equally spaced searchers (S is proportional to u/nv), and no explicit choice as to its value determines the search path. In this respect it is like the continuous symmetrical crossover patrol. This raises the question as to how two searchers patrolling abreast should be spaced. If the definite range law is applied, search abreast at a distance apart equal to twice the range would be equal in effectiveness to search by each singly, equally spaced on one and the same square path. With any other kind of contact probability law the corresponding track spacing in search abreast is the value of S which is found by using formula (19) and the appropriate value of n. It will be noted that this is not equivalent to placing two or more searchers closely abreast on exactly the same track. They must be either equally spaced on an identical track, or, if abreast, spaced at right angles to that track by exactly the right amount. The effects of irregularities in spacing are, however, rather small. It should be noted that if the value of S/n is at all large, search abreast on square or circular paths cannot be carried out effectively because of the difference in the length of track for each observer.

It is apparent that the maximum size of square is fixed by (a) the endurance of the searcher (endurance $> 4b$) and by (b) the number of searchers continuously available. When surface craft are employed against underwater targets it is very desirable that two or more patrol abreast. Since a large number will be required to cover a square of any size, it will be more practical to employ small groups (three to six) in crossover patrols on barrier lines which form a polygon with a convenient number of sides, rather than have them all patrol a single large square.

7.2.3 Air Patrols Against Incoming Submarines

Detailed application of the foregoing considerations to a particular situation depends mainly on the range of detection to be expected from the radar gear installed on the aircraft. This will determine the interval at which aircraft patrol on the basic square track. The longer the radar range the fewer aircraft need be employed, and the easier it will be to navigate along the track. With S-band or X-band radar, navigation will be made very easy by the constant presence on the screen of check points on the island.

The distance from the position being protected from submarines at which the patrol track will be placed will be the estimated submerged run of submarines (not over 70 miles) plus the effective visibility for the radar used, as given in Chapter 5. With ASG radar used on patrol planes in World War II this will result in flying a square with legs of about 180 miles to each side. With PBM aircraft the complete circuit of the square will then require 6 hours. Two complete circuits can be made by each sortie in 12 hours. Reference to formula (19) and standard tables of effective visibilities shows that only two 120-knot aircraft need patrol at one time equally spaced in order to give a 50 per cent chance of contact with a 15-knot surfaced submarine. Four aircraft will give a probability of contact of 77 per cent and six aircraft will give a probability of contact of about 90 per cent. With the older and less effective ASB radar ($E = 8$ miles), however, six 150 knot aircraft patrolling a square with 150 mile sides will only give a probability of contact of 70 per cent, and each sortie will be limited to a single circuit (four hours). Thus 18 TBF sorties give less protection than four PBM sorties.

Consideration of these figures leads to the conclusion that PBM aircraft should be used where available. A single squadron can probably offer six aircraft a day, at least for a limited period. These can be divided between four by night and two by day, to give a nearly uniform probability of contact of about 77 per cent, if due account is taken of the added possibility of sighting submarines visually in the daytime and the need to maximize the defensive value of the patrol. The probabilities given are applicable to certain types of Japanese submarines of World War II.

The force requirements can be reduced by patrolling smaller squares, but it is questionable whether this expedient should be employed. There is no gain in security if we give up the outer screen in order to patrol waters through which submarines are likely to transit largely or entirely while submerged. Screening operations within fifty miles of the beach are largely the responsibility of surface craft, except at night. Where additional forces are available, such as short range carrier-based aircraft, they may well be employed in patrolling an additional smaller square. This is particularly important at night when the surface screen may become less tight, owing to the possibility that the submarine may surface, and thus make higher speeds. Patrol of a square only 50 miles from the beach by only three TBF aircraft, even at night, gives an 80 per cent probability of contacting any surfaced submarines which may slip by an outer screen of four PBM's. This raises the overall probability of contact at night to over 95 per cent.

When the land mass involved is unsymmetrical and only one dimension is larger than 20 miles (a long, thin island), the square patrol may still be the simplest and most efficient method of setting up the antisubmarine barrier, when the landing operation being defended is made at or near one end of the island, even though a small part of the flying is over land. When all dimensions are large a barrier patrol or a combination of barrier patrols of the familiar crossover type will be set up at the appropriate distance from our forces, along the coast.

7.2.4 Surface Patrol Against Incoming Submarines

Surface craft on inner screens are assigned the responsibility of catching submerged submarines which may have eluded the outer aircraft screen. This screen will be placed outside torpedo range, but not so far outside as to require excessive patrol craft or as to thin out the tightness of the screen. Unless the patrol line has other than antisubmarine functions, it is therefore unlikely that it will be set up to surround an entire island, and the barrier patrol will be composed of straight line segments each of which follow an appropriate design based on an obvious application of the principles made familiar in earlier parts of this book.

In all cases, even when the size of the island permits patrolling all the way around it, it is very important to distribute the forces available so as to sweep as wide a swath as possible, rather than to have single ships or pairs of ships patrolling in column, or in short segments of the patrol line. The submarine can easily evade one or two ships; in addition, when ships patrol a very short segment there is the serious difficulty of continual interference by own wakes. At least three ships, and preferably four to six ships, should always patrol abreast. Under normal sound conditions, and using plans of this type, a group of four ships patrolling abreast can hold a very tight line almost 40 miles long. With the forces normally assigned to such operations, a line 200 miles long is easily maintained. The required patrol line is usually less than 100 miles long, and ten to twelve ships distributed between two or three groups should be more than sufficient. Additional forces usually available may be used to strengthen this line still further or to set up additional screens at other distances.

7.2.5 Patrol Against Centrifugal Targets

The general considerations concerning the circular (or square) barriers, as well as the closer barriers under certain circumstances, set forth in Sections 7.2.2, 7.2.3, and 7.2.4, above apply in an obvious manner to the case of targets seeking to leave the central point. No further discussion is required here.

7.3 SEARCH ABOUT A POINT OF FIX

7.3.1 The General Question

When an object of search on the ocean, such as a surface craft or downed airplane life raft, has had its approximate position disclosed to a searcher at a certain time, the searcher has the problem of disposing its subsequent searching effort (which is always limited) in such a manner as to maximize its chance of detecting the object, subject, of course, to the practical restrictions of navigation. The information regarding the object's approximate position may be derived from a DF fix, the report of a chance observation, by indirect inference, or, in the case of the life raft, from a radio communication from the aircraft about to crash. The point at which this information locates the target is called the *point of fix* and the time for which the information is given, the *time of fix*. It is assumed that the searcher is airborne, and thus has a considerable speed advantage over the target.

If the fix were a perfectly accurate one and the target were at rest or moving in a known manner, the searcher's task would be simple. He would proceed to the point of fix in the former case, and would search the locus of points to which the target, initially at the point of fix, could be assumed to have moved during the intervening time in the latter case. But such accuracy of fix is seldom if ever obtained: Only a probability distribution of target positions at the time of fix is actually given. This distribution will have its greatest density at the point of fix and fall continuously to zero at a distance. In an important group of cases, this distribution can be regarded as symmetrical about the point of fix and can indeed be taken with satisfactory accuracy as a circular normal one,

$$f(x,y) = f(r) = \frac{1}{2\pi\sigma^2} e^{-r^2/2\sigma^2}, \qquad (20)$$

where $f(x,y)dxdy$ is the probability that the target at the time of fix be in the small region $dxdy$ at the point (x,y) at the distance r from the origin O (which is at the point of fix) and where σ is the standard deviation.

It should be remarked that DF fixes usually do not give rise to circular distributions, but under certain conditions the distributions to which they lead are of this character.

Two cases are considered in this chapter. In the first, the target's motion is negligible, so that it can be regarded as at rest; equation (20) gives its distribution at all subsequent times. In the second, the target's speed is known but its direction is not but is assumed to be uniformly distributed in angle throughout the full circle; the distribution after the lapse of time t after the fix has already been derived in Chapter 1, equation (10). The solution in the second case will be derived more or less directly from the first. It is to the second case that the mathematical schema of the centrifugal vector field mentioned at the beginning of this chapter applies exactly, but in contrast to the cases of Section 7.2, the density of targets is not constant, but after being humped up about the center spreads itself out into a thick ring cut normally by the vectors, with the lapse of time.

7.3.2 Square Search for a Stationary Target

In this case, as we have seen, equation (20) gives the probability density of the distribution of the target for all later time. If the total searcher's track length during the search is L miles and his search-width W, then the quantity of searching effort as it has been defined in Chapter 3 is $\Phi = WL$. The problem of so disposing a continuous spread of searching effort of total amount Φ that the probability of detection is maximum has been solved in Section 3.4. But here we are confronted by the practical problem of designing actual navigable flights which will maximize the probability of detection.

The type of flight which it is expedient to use consists of a set of "expanding square" flights of the sort shown in Figure 9. After passing over the point of fix O to a point S miles beyond O, the aircraft turns through a right angle (e.g., to the right), and after S miles it turns again through a right angle in the same direction, continuing $2S$ more miles before turning a third time; after flying $2S$ miles it turns again, and continues in this manner, always keeping adjacent parallel tracks S miles apart. After a certain number of legs have been flown, there results an approximately square figure covered by equally spaced lines, the space between them being the sweep spacing S. Such a figure will be called a "square of uniform coverage." The underlying scheme of the search is to fly a succession of superimposed squares of uniform coverage, each centered at O, and of successively large dimensions. This will furnish a practicable means of approximating to the theoretically optimum continuous distribution of searching effort derived in Chapter 3.

The first problem is to determine a desirable value of the sweep spacing S, up to now left arbitrary. The point of view adopted here is that S should be so chosen that on the initial square the probability of detection per unit time shall be a maximum (during the important part of the search, i.e., the beginning). Clearly such an S will, for a given law of detection, be a function of the parameters of detection and of the standard deviation of the distribution σ. In the

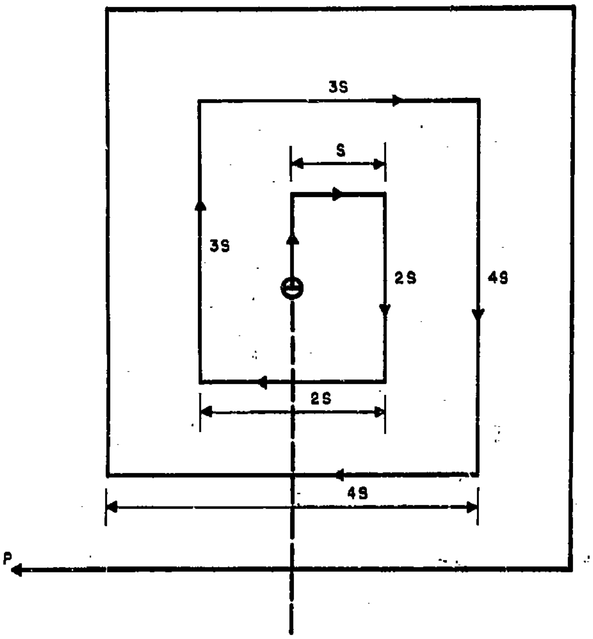


FIGURE 9. Square of uniform coverage.

case of the inverse cube law of detection, S will be a function of E and σ. Since it can be shown that S will not be sensitive to the law of detection, it is permissible to assume a convenient one. We shall assume the inverse cube law, in which, as was seen in Chapter 2 [equations (26) and (46)],

$$p(x) = 1 - \exp\left[-0.092\left(\frac{E}{x}\right)^2\right],$$

where $p(x)$ is the probability of detection of a target of lateral range x from a straight aircraft track.

Consider the distribution (20) of targets before any flights are made. The probability of the target lying on the strip parallel to the y axis between x and $x + dx$ is found (by summing probabilities) to be

$$dx\int_{-\infty}^{+\infty} f(x,y)dy = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}\,dx. \qquad (21)$$

The graph of the differential coefficient against x is the familiar normal law curve, reaching its maximum at the origin. Now suppose that an indefinite straight flight is made along the y axis and has failed to detect the target; in the light of this additional knowledge, the distribution of targets is altered, and the differential coefficient in (21) no longer represents the lateral density of targets [i.e., the probability in the $(x, x + dx)$ strip]. To find the new density, the use of Bayes' theorem is called for (see Section 1.5, footnote c). The "a priori probability" of the target's being in the $(x, x + dx)$ strip is given by (21); the "productive probability" of the event (viz., of not detecting the target on the sweep through O) is

$$1 - p(x) = \exp\left[-0.092\left(\frac{E}{x}\right)^2\right],$$

and hence the "a posteriori probability" density of probability of the target's being in the $(x, x + dx)$ strip is *proportional* to the product

$$\exp\left(-\frac{x^2}{2\sigma^2} - \frac{0.092E^2}{x^2}\right), \qquad (22)$$

which has its maxima at $x = \pm 0.65\sqrt{E\sigma}$; it is no longer humped at the origin but presents double humping with an intervening depression at the origin.

Where must a second indefinite sweep parallel to the y axis be made if it is to achieve the greatest probability of detecting the target? Since the distribution obtained above is skewed, the distance D between the first and the second track should be slightly greater than the distance $0.65\sqrt{E\sigma}$ out to the maximum of the new distribution. To find it precisely, we compute the probability of detection by multiplying the expression (22) with $p(x - D)$ and integrating over all positive x. It appears at once that in order to maximize this probability, D must make the function

$$\int_0^\infty \exp\left[-\frac{x^2}{2\sigma^2} - \frac{0.092E^2}{x^2} - \frac{0.092E^2}{(x-D)^2}\right]dx$$

a minimum. By trial, it is found, in using numerical integration, that the approximate value of D is $0.75\sqrt{E\sigma}$.

Clearly, if the searching were done by means of infinite parallel equispaced sweeps, the sweep-spacing S would be taken equal to D if the chance of early contact is to be maximized. The square search of Figure 9, while not exactly of this type, is near enough so that an obvious choice of S is to give it the same value D; accordingly, we shall use henceforth the sweep spacing

$$S = 0.75\sqrt{E\sigma}. \qquad (23)$$

Returning to the square of uniform coverage, it is seen that the three-circuit flight path of Figure 9 can be inscribed in a square of side $2 \times 3S$. More generally, an N-circuit flight of this sort can be inscribed in a square of side $2NS$. The total length of tract from O to P is found (as an arithmetic progression) to be $L = 2S \times 2N(2N + 1)/2$. Hence the average density of searching effort (the ϕ of Section 3.3) is WL/area, or

$$WS\,\frac{2N(2N+1)}{(2NS)^2} = \left(1 + \frac{1}{2N}\right)\frac{W}{S},$$

which is approximately W/S, the value to be adopted here.

Consider now a sequence of n squares of uniform coverage centered at 0 and of half-side s_k $(k = 1, 2, \cdots, n)$, where

$$0 < s_1 < s_2 < s_3 < \cdots < s_n.$$

If (x,y) is a point in some of these squares, let us say in those of side $s_{m+1}, s_{m+2}, \cdots, s_n$, then the mean density of searching effort performed by these squares is $(n - m)\,W/S$. Thus the flights give rise to a searching effort function $z = \phi^*(x,y)$, the graph of which (in xyz space) is of the form of a tapered heap of square slabs of thickness W/S, piled upon one another and centered on the z axis. The total volume of this pile, $\iint \phi^*(x,y)dxdy$, must equal the total searching effort $\Phi = WL$; thus we must have

$$4\sum_{k=1}^{n} s_k^2 = SL. \qquad (24)$$

It is by means of this function $\phi^*(x,y)$ that we must approximate the solution $\phi(x,y)$ obtained in Section 3.4 of Chapter 3. There it was shown that, outside the circle of radius a given by

$$a^4 = 4\sigma^2\frac{\Phi}{\pi}, \qquad (25)$$

$\phi(x,y)$ is zero, while within this circle, it is given by

$$\phi(x,y) = \phi(r) = \frac{a^2 - r^2}{2\sigma^2}. \qquad (26)$$

The graph of $z = \phi(x,y)$ is thus a paraboloid having the z axis as axis of revolution, cutting the xy plane in the above circle, beyond which the paraboloid is replaced by the xy plane.

Thus, graphically put, our problem is to determine the quantities $n, s_1, \cdots, s_n$, subject to (24), so that the piled slab solid shall approximate the paraboloidal one. The heights of the two solids being nW/S and $a^2/2\sigma^2$ respectively, n is determined by equating them,

$$n = \frac{a^2 S}{2\sigma^2 W}. \qquad (27)$$

Since n must be an integer, (27) means that it must be taken as the nearest integer to the right-hand member. To find s_k, consider the space between the two horizontal planes $z = a^2/2\sigma^2 - (k-1)W/S$ and $z = a^2/2\sigma^2 - kW/S$; they contain the kth slab of the piled slab solid, and hence cut from it the volume $4s_k^2 W/S$; and they cut from the paraboloid a volume of revolution readily found by integration to be $\pi(2k-1)\,\sigma^2 W^2/S^2$. On equating the two volumes, we obtain

$$s_k^2 = \frac{\pi}{4} \cdot \frac{W\sigma^2}{S}(2k-1), \quad k = 1, 2, \cdots, n. \qquad (28)$$

Thus the volumes of the piled slab solid and the paraboloidal solid are equal [and hence (24) is automatically satisfied, since it expresses the required equality of this volume with $\Phi = WL$, a requirement met by the paraboloidal volume, according to Section 3.4.], and the two solids agree in position about as closely as possible.

The number and dimensions of the squares of uniform coverage are determined by equations (27) and (28); but except for the fact that they are all centered at O, their positions (relative orientation) have been left arbitrary. We now lay down the following rule: *The second square should be tipped so that its side makes 45 degrees with the side of the first, the third should similarly be at 45 degrees with the second (and thus be parallel to the first) and, in general, the $(k+1)$ should be at 45 degrees with the kth [and parallel to the $(k-1)$].*

The justification of this rule is twofold:

First, a greater randomization of flights is achieved; i.e., there is less danger of passing over the same path twice in succession, with resulting loss of efficiency. The situation in this regard is illustrated by the following considerations. Navigational errors will not be apt to permit the second search to be flown along the optimum tract which is approximately midway between and parallel to the legs of the first search. If $p(S)$ is the probability of detection with sweep spacing S, then the probability of detection with two searches when the optimum track on the second search is attained is $p(S/2)$. However, if the second search duplicates the first search, the probability of detection with two searches is $p(S/\sqrt{2})$, (assuming the inverse cube law). If the legs of the second search are inclined to those of the first search, so that the probabilities of detection on the two searches may be considered as independent, the probability of detection with two searches is $1 - [1 - p(S)]^2$. In general this latter probability will be slightly less than $p(S/2)$ but considerably greater than $p(S/\sqrt{2})$. For example, using the inverse cube law with $E/S = 0.1$, we have

$$p(S/2) = 0.212,$$
$$p\left(S/\sqrt{2}\right) = 0.151,$$
$$1 - [1 - p(S)]^2 = 0.203.$$

Thus, there seems to be more to gain than to lose by inclining the legs of the second search to those of the first search.

Second and more important, the $(k+1)$ square flown according to this rule will sweep a maximum of important unswept water: consider the situation after one square of uniform coverage has been flown without resulting in detection; the area which it is most important to sweep is the part of the circle circumscribing the late square but outside the latter. With the next square tilted at 45 degrees, a maximum of this area is covered. Moreover, with the tilted square the region of overlapping of two successive squares is least.

It remains to evaluate the probability P^* of detection by the square search described herein, and to compare it with the probability P for the idealized search of Section 3.4. It is shown that a lower value of P^* is given by

$$P^{*n} = (1 - e^{-W/S})\, e^{-Wn/S} \sum_{k=1}^{n} e^{Wk/S} \operatorname{erf}^2 \sqrt{\frac{\pi W}{8S}(2k-1)}, \qquad (29)$$

while P is given by

$$P = 1 - \left(1 + \frac{nW}{S}\right) e^{-nW/S}. \qquad (30)$$

These two functions are plotted in Figure 10 for various values of n; the P^{**} curves are in solid line and the P curves are dotted. The P^* curves lie between these curves.

Evidently, P^* will be less than the probability P for the idealized search, since P is the maximum probability which can be obtained with the given amount of searching (provided the assumptions of Chapter 3 are made).

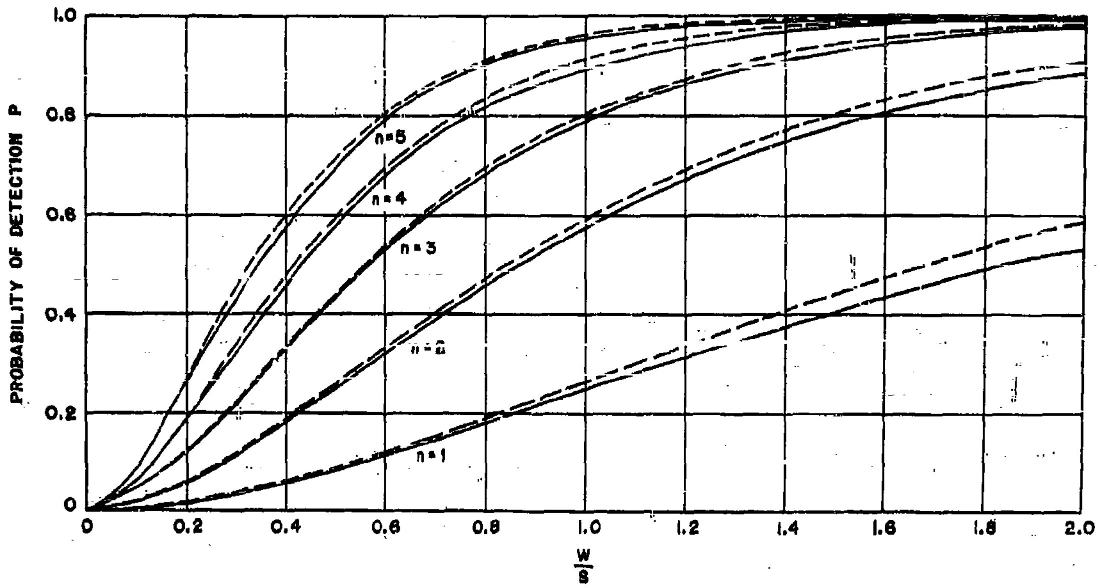


FIGURE 10. Probability of detection.

Using the target distribution as given in formula (20) and the density of searching effort as given in formula (26), the probability P of detection for the idealized case is obtained from formula (15) of Chapter 3 as follows:

$$P = \iint_A (1 - e^{-\phi(x,y)}) f(x,y)dxdy,$$

$$= \frac{1}{2\pi\sigma^2}\int_0^a 2\pi r\, e^{-r^2/2\sigma^2} (1 - e^{-(a^2 - r^2)/2\sigma^2})\, dr,$$

$$= 1 - \left(1 + \frac{a^2}{2\sigma^2}\right) e^{-a^2/2\sigma^2}.$$

Using formula (27), P may be written in the form of equation (30).

The direct computation of P^* is difficult because each square is tilted 45 degrees to the preceding square and each square, except the first, is not included completely in the following square. To obtain a lower limit to P^*, consider the square distribution before the squares are tilted, i.e., when they are all parallel, and let P^{**} be the probability of detection with this distribution of effort under the assumptions of Chapter 3. Then P^* will be greater than P^{**}, for reasons already given above. By formula (8) of Chapter 3, the probability that the target will be in the square bounded by $x = \pm s_1$, $y = \pm s_1$, and will be detected by the n coverages of the square, is

$$P_1 = \frac{4}{2\pi\sigma^2}\int_0^{s_1}\int_0^{s_1} (1 - e^{-nW/S})\, e^{-(x^2 + y^2)/2\sigma^2} dxdy$$

$$= (1 - e^{-nW/S}) \operatorname{erf}^2 \frac{s_1}{\sigma\sqrt{\pi}}.$$

Similarly, the probability that the target will be inside the square bounded by $x = \pm s_k$, $y = \pm s_k$ but outside the square bounded by $x = \pm s_{k-1}$, $y = \pm s_{k-1}$ and will be detected by the $n - k + 1$ coverages of this area is

$$P_k = (1 - e^{-(n-k+1)W/S})\left(\operatorname{erf}^2 \frac{s_k}{\sigma\sqrt{2}} - \operatorname{erf}^2 \frac{s_{k-1}}{\sigma\sqrt{2}}\right).$$

Thus, P^{**} is given by

$$P^{**} = \sum_{k=1}^{n} P_k = (1 - e^{-W/S})e^{-nW/S}\sum_{k=1}^{n} e^{kW/S} \operatorname{erf}^2 \frac{s_k}{\sigma\sqrt{2}}.$$

Using formula (28), P^{**} may be written in the form of equation (29).

7.3.3 Retiring Square Search for a Moving Target

The case here considered is that in which the initial distribution at the time of fix is given by equation (20), but with the target moving with an estimated constant speed u in a random direction. This is the second case mentioned in Section 7.2 and corresponds with what can be expected to occur when the fix has been made on a moving target (direction unknown) by a method of observation which has not imparted information to the target and thus influenced its motion. The problem of finding the distribution at the later time was solved in Section 1.6, where it was found that the probability density $f(r,t)$ at a point r miles from the point of fix and t hours after the time of fix is given by the equation

$$f(r,t) = \frac{1}{2\pi\sigma^2} e^{-(r^2 + u^2t^2)/2\sigma^2} I_0 \frac{rut}{2}. \qquad (31)$$

The function $f(r,t)$ has been plotted for various values of t as a function of r in Figure 17 of Chapter 1. (These curves are cross sections of the distribution surfaces in a vertical plane through the point of fix.) It will be seen that when ut is at least as great as 3σ, the distribution has its maximum at approximately $r = ut$. Moreover, the shape of the curve for the large values of t is approximately the same as the shape of the curve of the initial distribution. These statements can be verified by using the asymptotic approximation

$$I_0(x) \sim \frac{e^x}{\sqrt{2\pi x}},$$

for large values of x. The asymptotic approximation for $f(r,t)$ then becomes

$$f(r,t) \sim \frac{e^{-(r-ut)^2/2\sigma^2}}{2\pi\sigma\sqrt{2\pi utr}},$$

for large values of utr/σ^2. When r is close to ut, the approximation becomes

$$f(r,t) \sim \frac{e^{-(r-ut)^2/2\sigma^2}}{2\pi\sigma ut\sqrt{2\pi}}, \quad \frac{utr}{\sigma^2} \text{ large, } r - ut \text{ close to zero.}$$

The statements concerning the curves become apparent from this latter approximation.

From the statements made in the last paragraph and the results of the preceding section we can construct a search for large values of the time T which has elapsed from the time of fix to the initiation of the search. In order to obtain the maximum probability of detection per unit time, the first circuit should be flown on the peak of the distribution (the peak is *circular*—it is the top of a ring); the second circuit a distance S as given in equation (23) inside or outside—say outside—the peak; the third, a distance S inside the peak; the fourth, a distance $2S$ outside the peak; etc. Since the peak of the distribution moves out approximately at the speed of the target, the ideal track on each circuit is an equiangular or logarithmic spiral.

We shall approximate each circuit by four legs with 90-degree turns, as shown in Figure 11. Let

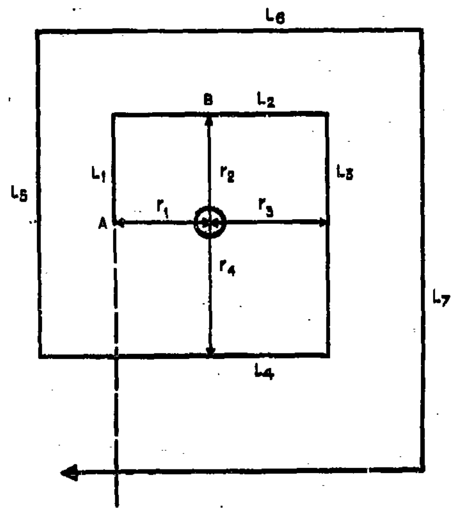


FIGURE 11. Retiring search.

the lengths of the legs be L_1, L_2, L_3, etc., and let the corresponding distances of the legs from the point of fix be r_1, r_2, r_3, etc. The time required for the aircraft to fly from A to B is $(r_2 + r_1)/v$, whereas the time required for the target to move from distance r_1 to distance r_2 from the point of fix is $(r_2 - r_1)/u$. The aircraft will just keep up with the target if these times are equal. In this way we find that

$$r_2 = mr_1,\ r_3 = mr_2,\ r_4 = mr_3,$$

where

$$m = \frac{v + u}{v - u}.$$

If we were to make $r_5 = mr_4$, the fifth leg would duplicate the first leg in space relative to the target. To make the fifth leg lie S miles outside the first leg in relative space, we must determine r_5 so that

$$\frac{r_5 + r_4}{v} = \frac{r_5 - r_4 - S}{u},$$

from which we find

$$r_5 = mr_4 + a,$$

where $$a = \frac{vS}{v - u}. \qquad (32)$$

Continuing in this way, we obtain

$$\begin{aligned} r_2 &= mr_1 & r_8 &= mr_7 \\ r_3 &= mr_2 & r_9 &= mr_8 - 2a \\ r_4 &= mr_3 & r_{10} &= mr_9 \\ r_5 &= mr_4 + a & r_{11} &= mr_{10} \\ r_6 &= mr_5 & r_{12} &= mr_{11} \\ r_7 &= mr_6 & r_{13} &= mr_{12} + 3a, \text{ etc.} \end{aligned} \qquad (33)$$

A first approximation to r_1 is uT. However, the approximation to the spiral by straight legs requires that r_1 be slightly less than this value. Equating the average distance of the aircraft to the average distance of the peak of the distribution from the point of fix, the average being with respect to time spent by the aircraft on a leg of the plan, we obtain

$$r_1 = 0.9uT.$$

Since any changes of course of the target will reduce the outward component of its velocity and since the second circuit is to be flown outside the first circuit, we shall reduce r_1 still further and take

$$r_1 = 0.8uT. \qquad (34)$$

Using equation (33) and the obvious relations between the lengths of the legs and their distances from O, we have

$$\begin{aligned} L_1 &= mr_1 & L_7 &= mL_6 \\ L_2 &= mL_1 + r_1 & L_8 &= mL_7 - 2a \\ L_3 &= mL_2 & L_9 &= mL_8 \\ L_4 &= mL_3 + a & L_{10} &= mL_9 - 2a \\ L_5 &= mL_4 & L_{11} &= mL_{10} \\ L_6 &= mL_5 + a & L_{12} &= mL_{11} + 3a, \text{ etc.} \end{aligned}$$

where r_1 and a are given by equations (32) and (34).

The above search plan has been devised for large values of T. For small values of T two questions arise:

1. What is the lower limit T_1 of T for which the search plan for large values of T may be used without any essential decrease in the probability of detection for a given amount of searching?

2. For T less than T_1, what modifications of the above plan for large values of T will give an essential increase in the probability?

The curves of Figure 17 of Chapter 1 do not answer these questions. It is seen that when t is less than σ/u the distribution does not differ much from the initial distribution. However, there is a very rapid change in the distribution as t increases from σ/u to $2\sigma/u$. For t greater than $2\sigma/u$ the distribution has its maximum at approximately $r = ut$ and moves outward at the speed of the target. This transition period between an essentially stationary distribution and a distribution moving at the speed of the target makes the problem difficult.

The following derivation of equation (31) suggests a method of handling the problem. In terms of

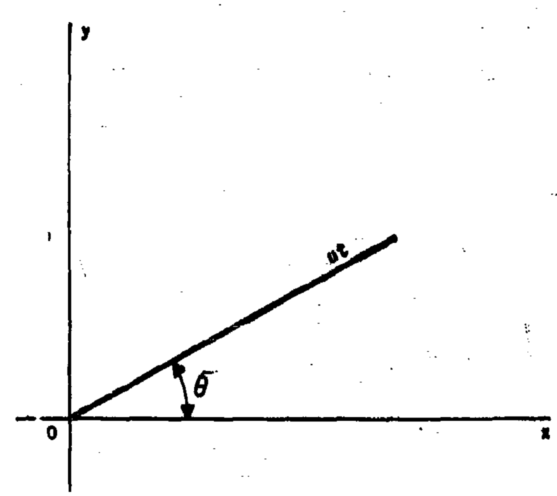


Figure 12. The coordinate system for retiring search.

the rectangular coordinates (x,y) shown in Figure 12 the initial distribution is

$$f(x,y) = \frac{1}{2\pi\sigma^2} e^{-(x^2 + y^2)/2\sigma^2}.$$

If the target were known to travel at speed u in the direction θ, the distribution at time t would be

$$F(x,y;t,\theta) = \frac{1}{2\pi\sigma^2} e^{-1/2\sigma^2[(x - ut\cos\theta)^2 + (y - ut\sin\theta)^2]};$$

in other words, distribution (20) with a simple change of origin (translation). If, on the other hand, the target motion is random in direction, the distribution at time t is

$$f(x,y;t) = \frac{1}{2\pi}\int_0^{2\pi} F(x,y;t,\theta)d\theta$$

$$= \frac{1}{2\pi\sigma^2} e^{-(r^2+u^2t^2)/2\sigma^2} \cdot \frac{1}{2\pi}\int_0^{2\pi} e^{utr/\sigma^2 \cos(\theta-\alpha)}\, d\theta,$$

where $r^2 = x^2 + y^2$ and $\tan\alpha = y/x$. Using $\theta - \alpha$ as a new variable of integration and noting that the integral of a function of $\cos\theta$ from $-\alpha$ to $2\pi - \alpha$ is equal to the integral of the same function from 0 to 2π, we obtain the distribution given in equation (31).

For a given search plan and an assumed law of detection, we can apply the above method to find the searched distribution at any time t. We first find the effect of the searching upon the distribution function $F(x,y;t,\theta)$ and then average with respect to θ. To approximate this process, choose a small number of equally spaced directions θ from 0 to 2π, the number depending upon the degree of approximation desired. For each θ write on coordinate paper numbers proportional to the initial distribution and think of this distribution as moving in the direction θ. Then lay out the search plan relative to this distribution up to the time at which the distribution is to be examined. From the assumed law of detection, multiply the distribution numbers by the probability that the target will not be detected, taking into account the number of legs on which detection may occur and the distances from these legs. The result will represent the searched distribution for a given θ. The average distribution then can be obtained by averaging the numbers representing the searched distributions for each position relative to the point of fix and a given direction of reference.

In applying the above method it was assumed that four directions would give sufficient accuracy. A check was run by computing a distribution with four directions and with eight directions. It was found that the two distributions did not differ very much. Using a number of values for t and S, search plans were laid out and tested as follows: The first circuit was decided upon from the curves of Figure 17, Chapter 1. Thereafter, each circuit was determined by examining the distribution at the end of the previous circuit. The following results were obtained.

1. If $T > \sigma/u$, the original plan for large values of T is nearly optimum.

2. If $T < \sigma/u$, the plan for large values of T is fairly good and may be used if the additional complication of another plan is not acceptable. However, an appreciable improvement can be obtained by slight modifications. Lay out the plan as in Figure 11 with

$$\begin{aligned} L_1 &= mr_1 & L_7 &= mL_6 \\ L_2 &= mL_1 + r_1 & L_8 &= mL_7 + a_2 \\ L_3 &= mL_2 & L_9 &= mL_8 \\ L_4 &= mL_3 + a_1 & L_{10} &= mL_9 + a_2 \\ L_5 &= mL_4 & L_{11} &= mL_{10} \\ L_6 &= mL_5 + a_1 & L_{12} &= mL_{11} + a_3, \text{ etc.} \end{aligned}$$

and determine r_1, a_1, a_2, etc., as follows:

$$r_1 = \frac{S}{m^2+1}, \tag{35}$$

$a_1 = a_2 = \cdots = a_k = vS/v - u, a_{k+1} = -(k+2)a_1$, $a_{k+2} = +(k+3)a_1$, $a_{k+3} = -(k+4)a_1$, etc., where k is the positive integer nearest $2ut/S$. Here, r_1 has been chosen so that L_1 and L_3 are separated by a distance S. This scheme seemed to be the best possible for the first circuit. The succeeding circuits then are flown outside this circuit until the point is reached at which it is more profitable to search on the inside of the distribution. In case T is so small that $2uT < S$, a slight improvement can be obtained by decreasing a_1 by 10 or 20 per cent.

In the case of this section it is assumed that the target speed u is known exactly. The question naturally arises as to how much the probability of detection is affected by an error Δu in estimating u. From the asymptotic approximation to $f(r,t)$ for large values of t it is seen that the distribution function will be multiplied approximately by the factor $\exp(-t^2\Delta u^2/2\sigma^2)$. This is a very rough estimate of the factor by which the probability of detection will be multiplied. Thus, if $t\Delta u$ is small compared with σ, the probability of detection is not affected very much.

The effect of a speed distribution upon the target distribution can be obtained in the usual way. Let $g(u)$ be the probability density in speed. Then

$$\int_0^\infty g(u)du = 1$$

and the target distribution at time t is

$$\int_0^\infty g(u) \cdot f(r,t)du.$$

It is evident from the way in which u is involved in $f(r,t)$ that this integral will be difficult to compute for

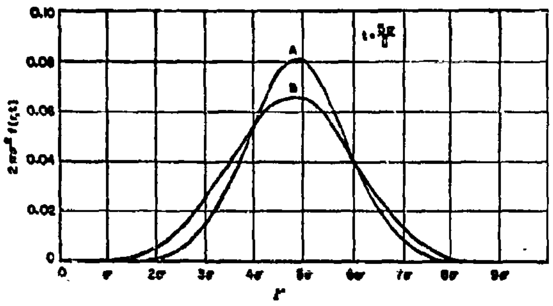


FIGURE 13. Distribution curves for two assumed distributions of target speed. A: assumed speed $\bar{u}$ only. B: speed of $\bar{u}$, $0.8\bar{u}$, $1.2\bar{u}$ with speed $\bar{u}$ times as likely as $0.8\bar{u}$ or $1.2u$.

any continuous function $g(u)$ which represents a reasonable speed distribution. Consequently, the distribution curve was plotted from the curves of Figure 17, Chapter 1 for particular values of t using the speed distribution function

$$\begin{aligned} g(u) &= 0.25 \text{ for } u = 0.8\bar{u} \\ &= 0.50 \text{ for } u = \bar{u} \\ &= 0.25 \text{ for } u = 1.2\bar{u} \end{aligned}$$

where $\bar{u}$ is the assumed speed. This curve for $t = 5\sigma/\bar{u}$ and the corresponding curve of Figure 17, Chapter 1, are shown in Figure 13. The new distribution curve has a smaller maximum value than the original distribution and is spread out more. However, the difference between the two distributions is not sufficient to justify any changes in the search plans.

Chapter 8

SONAR SCREENS

8.1 INTRODUCTION

When a formation of ships, such as a merchant convoy or task force in transit, is passing through waters in which hostile submarines may be present, the danger can be greatly diminished by providing for the detection and attack of submerged submarines which are closing to its vicinity. With present equipment, submerged submarines can only be detected at sufficient range by hydrophone or echo ranging and can be located with sufficient accuracy only by echo ranging. This detection must occur as far as possible from the convoy, both for its safety against torpedoes (whose ranges may exceed those of sonar detection) and to facilitate attacks on detected submarines. Consequently, it is necessary that the sonar equipment be carried on highly maneuverable and armed naval units of a relatively expendable nature, such as destroyers or destroyer escorts, which are to be appropriately stationed at a suitable distance from the convoy. Such a disposition of units is called a *sonar screen*. It has the dual function of detection and subsequent attack.

8.2 TORPEDO DANGER ZONES AND HITTING PROBABILITIES

In designing antisubmarine screens, it is essential to determine what the areas are from which the submarine has a good chance of scoring a torpedo hit upon one of the units which it is proposed to screen. Speaking loosely, the *torpedo danger zone* about an individual ship or group of ships is the region (thought of as moving along with the ships, i.e., it is fixed relative to them) within which a torpedo must be fired if it is to have any chance of scoring a hit. The shape and size of the zone will of course depend on the speed and type of the torpedo, as well as the speed and disposition of the ships. Speaking more precisely, there is a danger zone *for each given probability P*. It is the area from which the torpedo must be fired in order to have a chance not less than P of scoring a hit. It is bounded by a closed curve containing the ships, which is the locus of points from which the torpedo must be fired in order to have the given probability P of hitting. Such curves, one for each value of P, are the level lines of the *probability function* $p(r,\theta)$ and have as equation

$$p(r,\theta) = P.$$

The probability function $p(r,\theta)$ is the value of the probability of scoring a hit by a torpedo fired from the point of polar coordinates (r,θ) with respect to the reference point in the formation of screened ships (and in space moving along with them).

It becomes, therefore, our first object to evaluate this function $p(r,\theta)$. The present section is devoted to the description of various methods for doing this.

The primary factor involved in the determination of the enemy's chance of success $p(r,\theta)$ is the type of weapon which he uses. If the lethal coverage of the weapon is high, the overall chance of success is correspondingly great. Consider, for example, three types of torpedo. The first runs at a 50-foot depth and explodes after a given length of run; the second at 5 feet and explodes on contact; the third at 5 feet, exploding on contact, but provided with a homing device such that whenever it passes within 500 yards of a ship it will home onto it and score a hit. It is obvious that the first will have a rather small lethal coverage, since only a ship in a particular point will be affected by the explosion. The coverage of the second is greater because any ship along the entire run of the torpedo may be hit. The homing feature of the third will give it by far the greatest coverage because the ships need only be within 500 yards of the torpedo track for a hit to be scored. These cases can be demonstrated qualitatively by a diagram like that of Figure 1. The areas are shown for a torpedo proceeding at right angles to the ships. The areas shown have the property that any ship whose center lies in the area will be hit by the torpedo. For case (1) the area covered is slightly larger than the plan view of the ship, since a torpedo exploding up to about ten yards away might sink the ship.

Lethal coverage alone does not determine the enemy's chance of a hit. His firing errors must be taken into account, and if his errors are so great that there is only a small probability that the target lie in the lethal area, his chance is correspondingly small. For any given accuracy, however, the weapon with the larger lethal coverage will be the more effective.

The foregoing discussion applies only to a single target. When fired at a convoy or group of ships, the torpedo will be successful if it hits any member of the group, i.e., if any of them lies in its lethal area. In general, the torpedo will actually be fired at a particular ship, but it may miss that ship and hit another one purely by chance. For long-range torpedoes fired at large convoys, the chance of such an event may be quite considerable. In such a case a torpedo may well be fired as a "browning shot" from a fair distance, aimed at the convoy as a whole on the chance of a random hit. Because of the importance of browning shots with large, closely spaced convoys, the probability of securing a hit on a merchant convoy is quite different from that on a single ship. The following discussion will deal with the case of the large convoy. The probability of hitting a single ship will be discussed in detail (in connection with the case of the task force) later in this chapter.

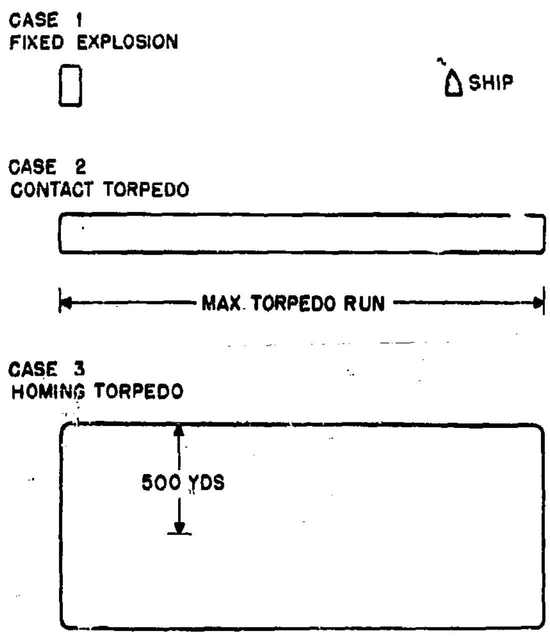


FIGURE 1. Lethal coverage for various torpedoes.

8.3 THE PROBABILITY OF SCORING A HIT ON A CONVOY WITH A SINGLE TORPEDO

While it has been the general practice to fire torpedoes in salvo, the probability that a salvo of a given number of torpedoes will hit, or sink or damage to a given extent, a ship in the convoy, depends on *the probability of scoring a hit with a single torpedo.* It depends also, of course, on the assumed law of damage produced by a given number of hits (the damage being different, e.g., for merchantmen and for battleships). The one-torpedo probability becomes, therefore, the fundamental quantity to be found as a preliminary to any study of the probability of damage to convoys.

In this section, several methods of estimating this probability of a single torpedo hit will be described. The data from which such computations start are, firstly, the structure of the convoy; secondly, the position of the firing point; thirdly, the velocities of the units involved. And since the computation must be repeated for a large number of firing positions, an essential requirement is simplicity and speed of computation.

The first method is purely graphical and involves no computational difficulties beyond mere counting. We draw the convoy to scale with each ship of proper size and position. Radiating from the firing point, we draw a set of straight torpedo paths in convoy space, so chosen as to be spaced in angle with an angular density corresponding to the dispersion or error which it is intended to assume. If, for instance, twenty torpedo paths are to be drawn and the angular distribution is considered normal (Gaussian), they would be drawn at angles which divide the normal curve into equal areas, as shown in Figure 2.

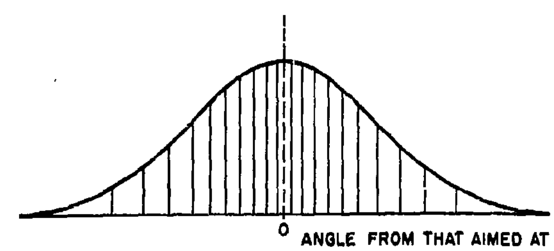


FIGURE 2. Torpedo firing errors.

Thus we consider each track drawn as having a probability of 1/20 that a torpedo aimed at angle = 0 will actually run along it. By superimposing the torpedo diagram on the convoy and counting the number of torpedo paths that hit ships, we can estimate the chance of a hit. In the example shown in Figure 3, six out of 20 score hits so that the probability of a hit from that firing point is 30 per cent.

In deciding on the probability of a hit from any particular point, however, it is necessary to pick out the submarine's best shot, whether he should aim at the ship as shown in Figure 3 or at the closer one at the head of the starboard column which presents a less favorable target aspect. It is usually possible to

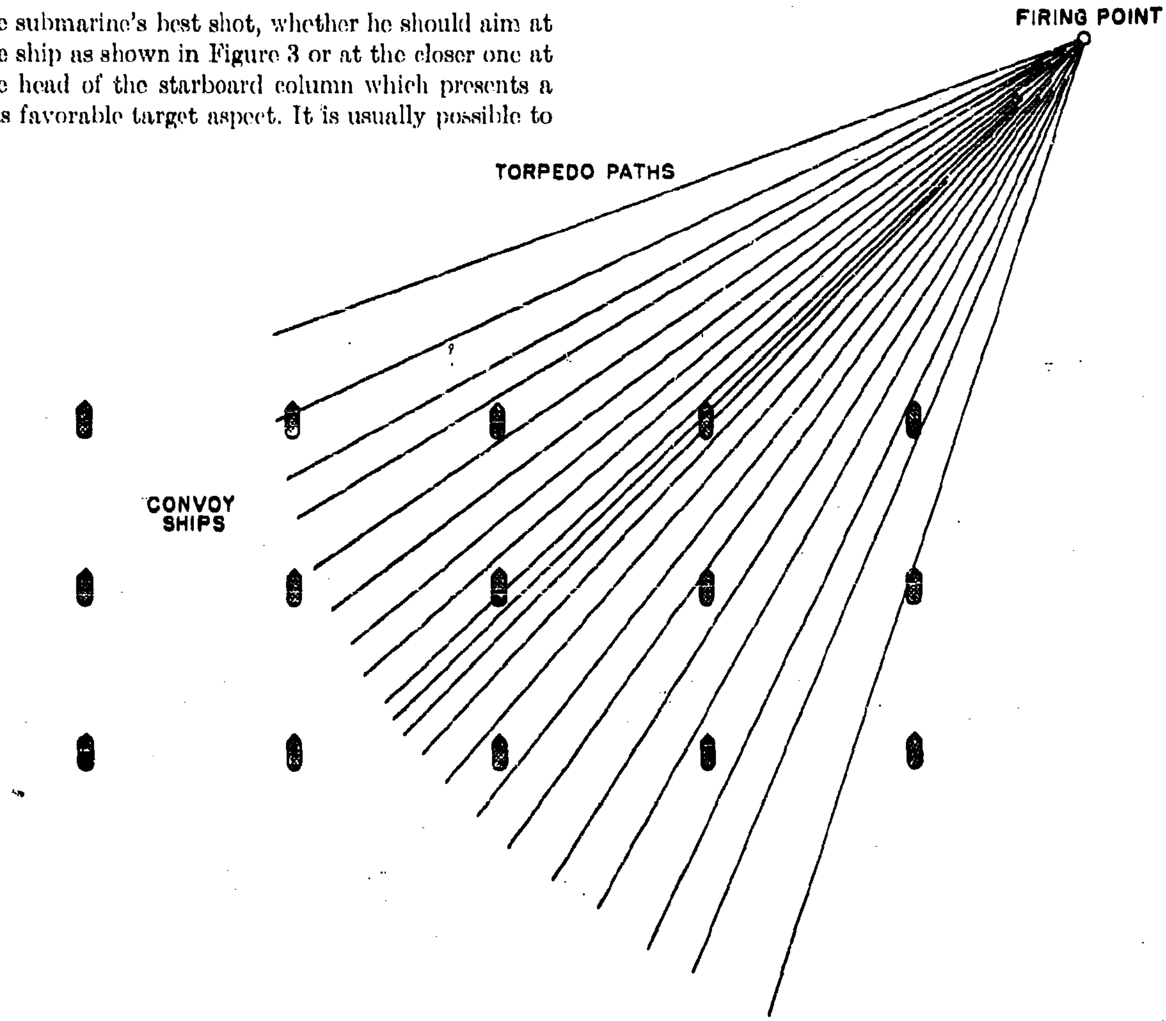


FIGURE 3. Convoy and possible torpedo paths.

pick out the best shot by eye after a little experience, but sometimes both possibilities must be counted.

This method is extremely simple and direct, but it has a number of disadvantages. In the first place, it involves careful drawing and positioning of the diagrams. A great deal of inspection of diagrams is required. Since only a rather small number of torpedo paths can be drawn conveniently, the values are only accurate to the nearest 5 per cent, and may show considerable fluctuations for very small changes in firing position or aiming angle. In addition the estimated probability shows humps and valleys due to the screening effect of a regular arrangement of ships. From certain angles one ship hides a number behind it, while from slightly different ones all are presented as targets. The arrangement of ships in any actual convoy is not orderly enough to show such an effect, so that such variations must be smoothed out more or less arbitrarily in arriving at a final result such as that shown in Figure 4.

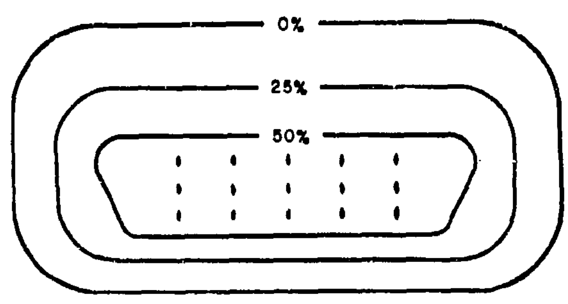


FIGURE 4. Typical probability of hit contours.

The choice of torpedo firing error function depends on torpedo and predictor characteristics and firing doctrines; it will not be discussed here. It should be pointed out, however, that this error depends in general on the track angle of the torpedo, the angular error being small from ahead, large from astern. The length of torpedo path in convoy space is also variable—long from ahead, short from astern. When the convoy fired at is large, the variation in firing error is not very important and can be neglected quite satisfactorily, but the variation in track length must be taken into account, as shown in Figure 5. (Figure

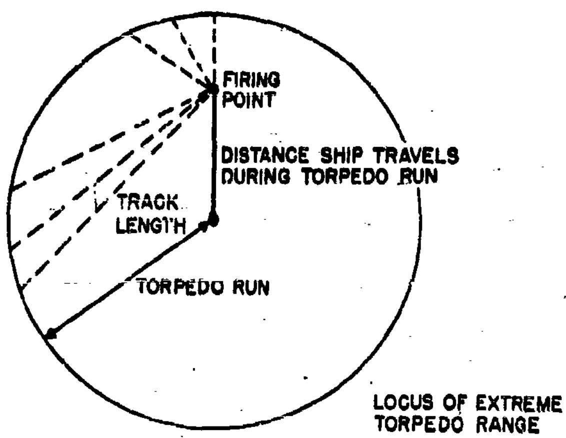


FIGURE 5. Torpedo track length L in convoy space.

3 conforms to this requirement.) Since the chance of a random hit on a ship not aimed at contributes a considerable part to the total hitting probability, the variation of the error in aiming is relatively unimportant. This contribution depends on track length but not on aiming error.

An obvious extension of this simple counting method can be used to reduce the fluctuations in calculated probability of hitting which are caused by the screening effect of a regular lattice of ships. We need only consider the possibility of irregularity in the convoy formation, that is, of ships being somewhat out of their proper convoy station. This can be accounted for by making each ship a "diffuse target" of length equal to the ship's length, plus the amount the ship varies in position (thought of, for simplicity, as a fixed amount) relative to the ship's position as aimed at. Then the probability that a torpedo passing through the diffuse target will actually hit the ship is taken as

$$p = \frac{L}{l}.$$

If we now consider the ith torpedo path ($i = 1, 2, \cdots, m$; m being 20 in the previous example) and the jth diffuse target ($j = 1, 2, \cdots, n$; n being the number of ships in convoy), we can define a number h_{ij} as the possibility that a torpedo traveling along the ith path will pass through the jth diffuse target. From a diagram analogous to Figure 3, we evaluate

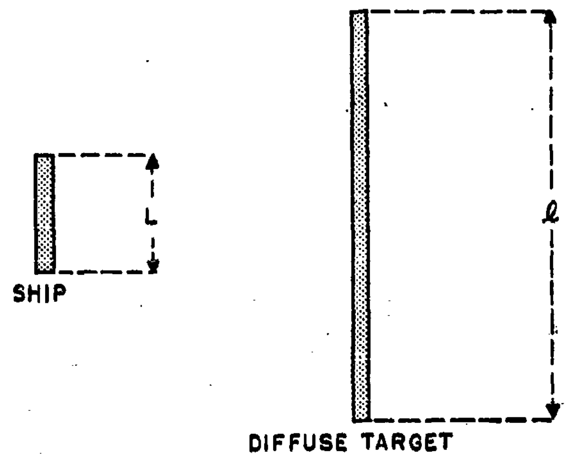


FIGURE 6. "Diffuse target."

the h_{ij} as being equal to either 1 or 0 by inspection. Then the probability that a torpedo following the ith path will hit the jth ship is

$$p_{ij} = h_{ij}\frac{L_j}{l_j}.$$

Hence the probability of *one* hit on the ith path being identical with that of *at least one* hit on the ith path is

$$p_i = 1 - \prod_{j=1}^{n}(1 - p_{ij}).$$

And the overall probability of a hit from the particular firing point is

$$p = \frac{1}{m}\sum_{i=1}^{m}\left[1 - \prod_{j=1}^{n}(1 - p_{ij})\right].$$

This probability function is considerably smoother and more realistic than that obtained by the simple counting method. The diffuse target method is considerably more laborious, however, and involves a good deal of arithmetical calculation.

It should be pointed out that the choice of the lengths of the diffuse targets is rather arbitrary. A plausible assumption is that each ship may be out of station with respect to its neighbors by a given amount, which we may call x. Then the length l for the nearest neighbors to the ship fired at will be $L + 2x$. If the station-keeping errors are assumed to be random, then the l for the next nearest neighbors will be $L + 2\sqrt{2}\,x$, for next $l = L + 2\sqrt{3}\,x$, etc., until the diffuse targets overlap and the whole column becomes essentially one large diffuse target.

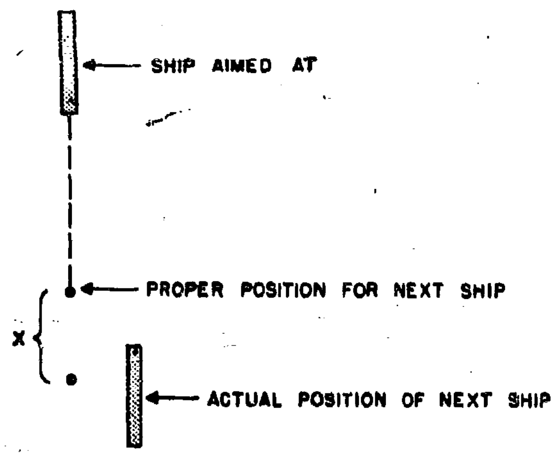


FIGURE 7. Station-keeping error.

In this picture of the convoy we think of the position of the ship aimed at as being fixed and the positions of other ships becoming increasingly uncertain the farther removed they are from the ship aimed at. It is evident that the actual state of affairs from the submarine's point of view is something of this sort, though the specific way in which the uncertainty increases may vary from that assumed above. If, however, the calculations required are prohibitively lengthy, a simplification is possible by doing away with this variable diffuse picture and treating each column of the convoy as a diffuse target in which the ships are distributed more or less at random, that is, with probability that a torpedo passing through it will hit a ship given by

$$p = \frac{L}{a},$$

where a = ship spacing in column.

This "random column" method is very simple and convenient but is not accurate at close ranges. In these cases the actual probability of hitting the ship aimed at is considerable, and the random column discounts this chance. For browning shots fired at a considerable distance from the convoy, however, this method is very satisfactory.

The simple counting, diffuse target, and random column methods are applicable to straight run torpedoes without much difficulty. If, however, we have a torpedo which enters a convoy and then zigzags or circles to increase its chance of a random hit, the resulting diagrams become almost hopelessly confusing. Such we call "curly" torpedoes. Some calculations have been carried out for curly torpedoes on a single target by direct methods, but the combination of a curly torpedo (or salvo thereof) and a multiplicity of targets is very hard to evaluate in that way. It is possible, however, to write explicit expressions for the probability of a hit by a torpedo of quite arbitrary type. While these formal expressions cannot readily be evaluated, there are certain simplifications possible which lead to useful results and permit us to estimate the effectiveness of some types of curly torpedoes when used against groups of ships.

Definitions:

1. Let $p_s(x,y)$ be the probability density of ships at (x,y) (coordinates are in space relative to the convoy throughout) in the sense that $p_s(x,y)dxdy$ is the probability that a ship have its center in the infinitesimal area $dxdy$.
2. Let $(X_0 Y_0)$ be the point from which the submarine fires its torpedo. It is assumed that only one torpedo is fired (see above).
3. Let $f(x,y,X_0,Y_0,a) = 0$ be the equation of a possible torpedo path starting at (X_0,Y_0). Variation of the parameters a gives the family of all possible torpedo paths starting from the firing point.
4. Let $\phi(a,a_0)$ be the probability density that a torpedo aimed along path a_0 actually travels along path a.

The meanings of these symbols are illustrated in Figure 8. In such a case the parameter a might be taken as the angle between the torpedo path and convoy course. The function $\phi(a,a_0)$ then describes the angular errors which the submarine may be expected to make in firing its torpedo.

In estimating the chance of a hit we will first assume that the torpedo is known to be traveling along a certain path, determined by the parameter a, and determine *its* chance of hitting. This chance is then averaged over all values of a, thus allowing for errors in firing the torpedo. This process is precisely analogous to that carried out for the diffuse target method with a straight run torpedo.

Consider first a short section of the torpedo path. Any ship which happens to be close enough to the path of the torpedo will be hit, but since we know only the probability that there will be a ship at any given point, we must think of the torpedo as "sweeping out" that probability (or density) function with a certain sweep rate. This sweep rate is determined by the type of torpedo, size of ship, and other things. For a regular contact-fuzed torpedo the area swept out (which has the property that any ship whose center lies in the area gets hit) is given by drawing a

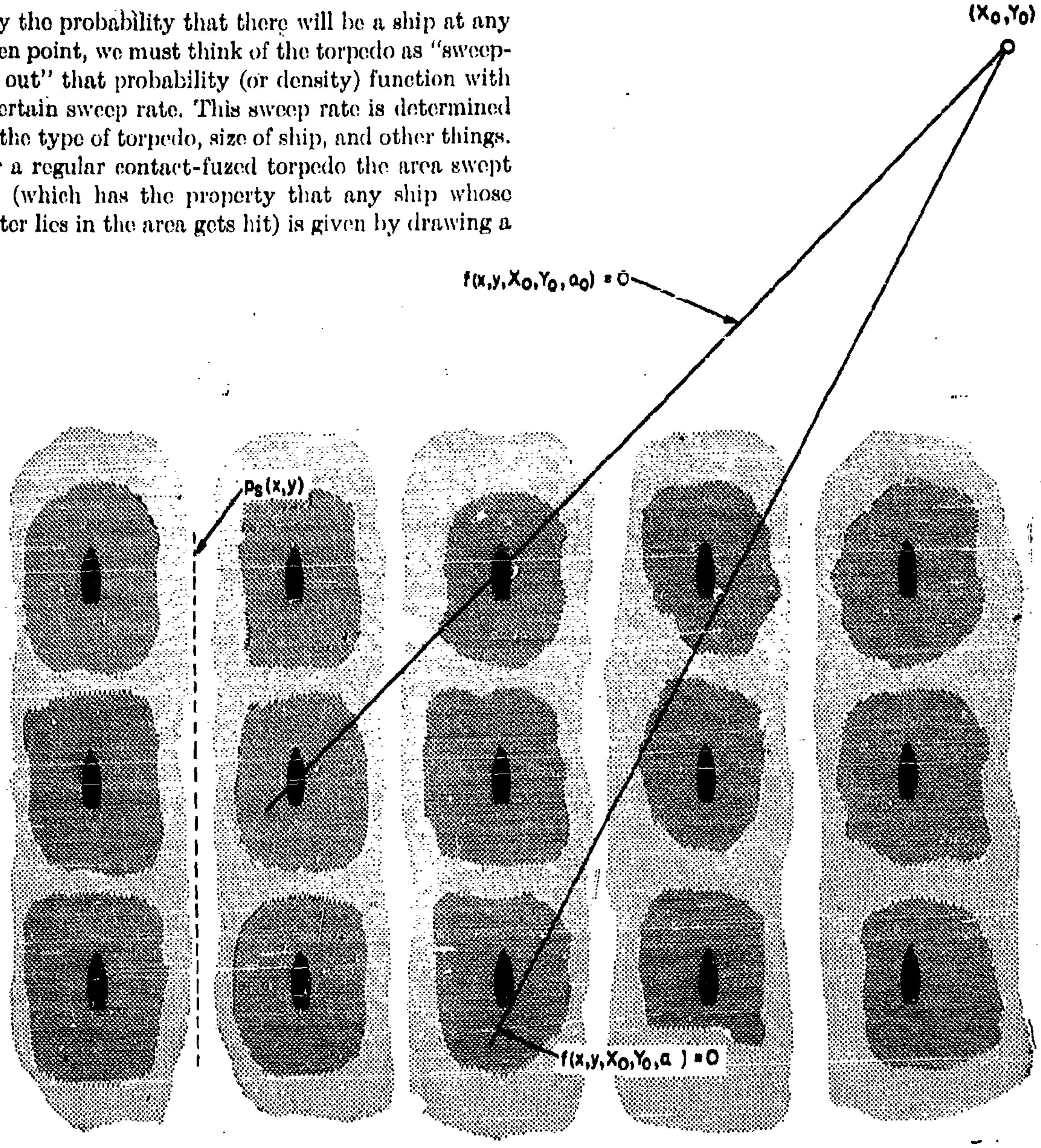


Figure 8. Convoy shown as density function.

ship in to scale, heading on the reciprocal of convoy course (corresponding with the fact that the diagram is relative to the convoy), with center on the torpedo path, and sliding its center along the relative path. This is illustrated in Figure 9. This sweep width depends on the angle between torpedo path and ship's course and is denoted by $S(\theta)$.

For a homing torpedo the effective target area is larger than the ship itself and may have a variety of forms, depending on the detailed type of torpedo and ship involved. To obtain the sweep rate for such a torpedo, we follow the same procedure, sliding a figure showing the target area rather than just the area of the ship.

As the torpedo proceeds along its path it sweeps out the density of ships p_s with sweep width $S(\theta)$. In Figure 10, l gives the distance which the torpedo has traveled, and we let $p(l)$ be the probability that

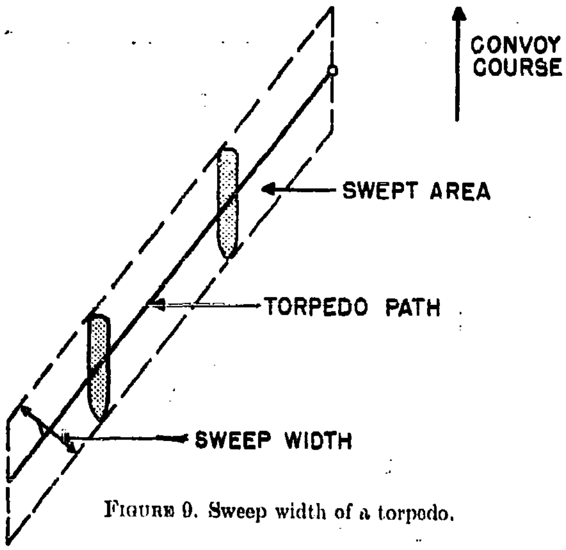


FIGURE 9. Sweep width of a torpedo.

the torpedo will have hit a ship by the time it has gone a distance l.

In going an additional distance dl it would sweep out $p_s S dl$ ships (p_s and S may be functions of l), but the chance that the torpedo actually reaches the

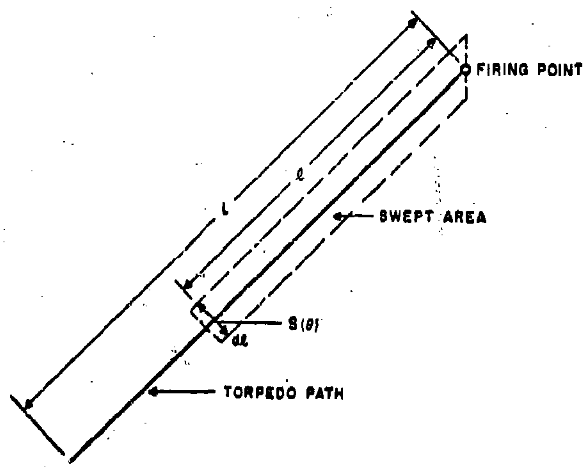


FIGURE 10. Typical torpedo path, drawn straight for simplicity.

distance "l" is $1 - p(l)$. Hence the increase in probability of hitting is

$$dp(l) = p_s S\,[1 - p(l)]\,dl.$$

Solution of this equation gives us

$$-\log\,[1 - p(l)] = \int_0^l p_s S dl,$$

or

$$p(l) = 1 - e^{-\int_0^l p_s S dl}.$$

In its complete run, then, the torpedo's chance of hitting is

$$p(L) = 1 - e^{-\int_0^L p_s S dl}.$$

We must remember, however, that this result applies only for the particular torpedo path selected and that the integration indicated is performed along that path with appropriate values for p_s and S. This may be indicated by introducing the parameter a into the notation as follows,

$$p[a] = 1 - e^{-\int_{\text{path } a} p_s S dl}.$$

Here $p[a]$ is the probability the torpedo will score a hit if it goes along the path a. It is then necessary to obtain the average value of $p[a]$ to obtain the probability that a torpedo fired from (X_0, Y_0) aimed to go on path a_0 will hit a ship, which we denote by $P(X_0, Y_0, a_0)$.

$$P(X_0, Y_0, a_0) = \int_{\text{all } a} \phi(a, a_0)\, p[a] da$$

$$= \int_{\text{all } a} \phi(a, a_0) \left\{1 - e^{-\int_{\text{path } a} p_s S dl}\right\} da.$$

It is evident that our previous methods of calculation were simplifications of this formula in which the quantity in braces was assigned one of a set of values by inspection of a diagram, and the function ϕ was approximated by a population of cases (in the examples discussed, 20 cases). It is also possible to approximate the expression by neglecting the error function ϕ if we assume that the position of the ship aimed at is not definitely known. In this picture the torpedo track is taken as perfectly definite, but

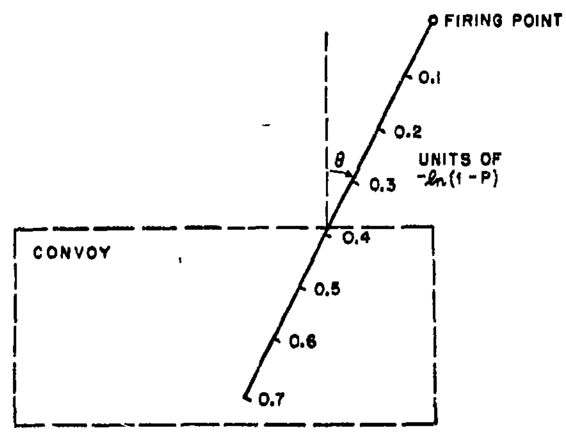


FIGURE 11. Graphical calculation of probability of hitting.

target positions relative to it are indefinite and the function p_s is made to absorb the torpedo firing errors. Then

$$P(X_0,Y_0,a_0) = 1 - e^{-\int_{\text{path } a_0} p_s S\,dl}.$$

If we make the further simplifying assumption that p_s is uniform throughout the convoy and zero outside [i.e., p_s = (no. of ships)/(convoy area)], this becomes

$$P(X_0,Y_0,a_0) = 1 - e^{-p_s A_0},$$

or

$$-\log(1 - P) = p_s A_0,$$

where A_0 is the total area swept out in the convoy by the torpedo.

The use of this formula can be illustrated by the simple case of a straight run torpedo. We draw the torpedo path at a given angle, as in Figure 11, and scale off along it distances proportional to $1/p_s S$. Mark resulting points of division by the corresponding multiple or submultiple of $1/p_s S$. Then for any position of the convoy shown dotted in Fig-

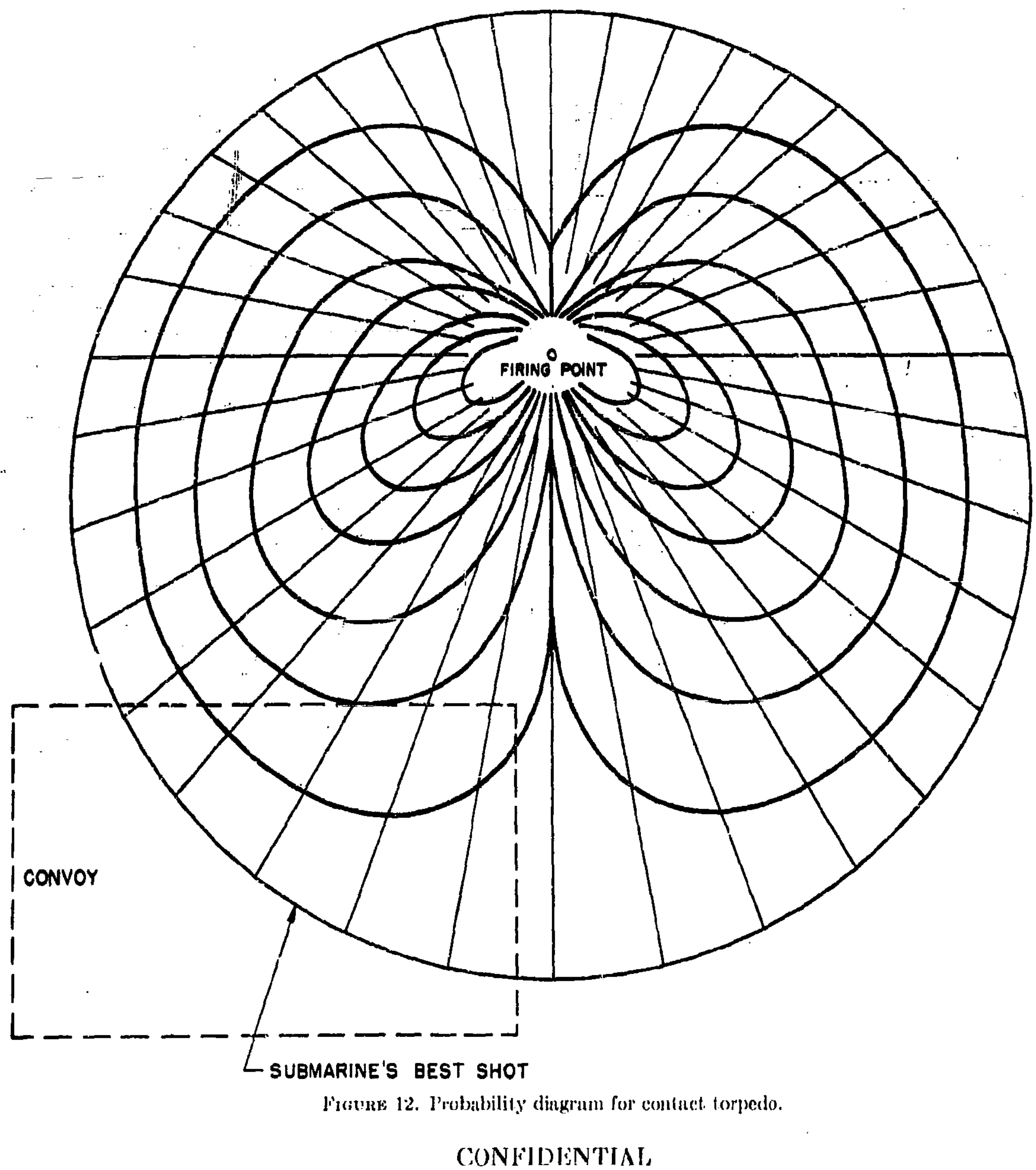


FIGURE 12. Probability diagram for contact torpedo.

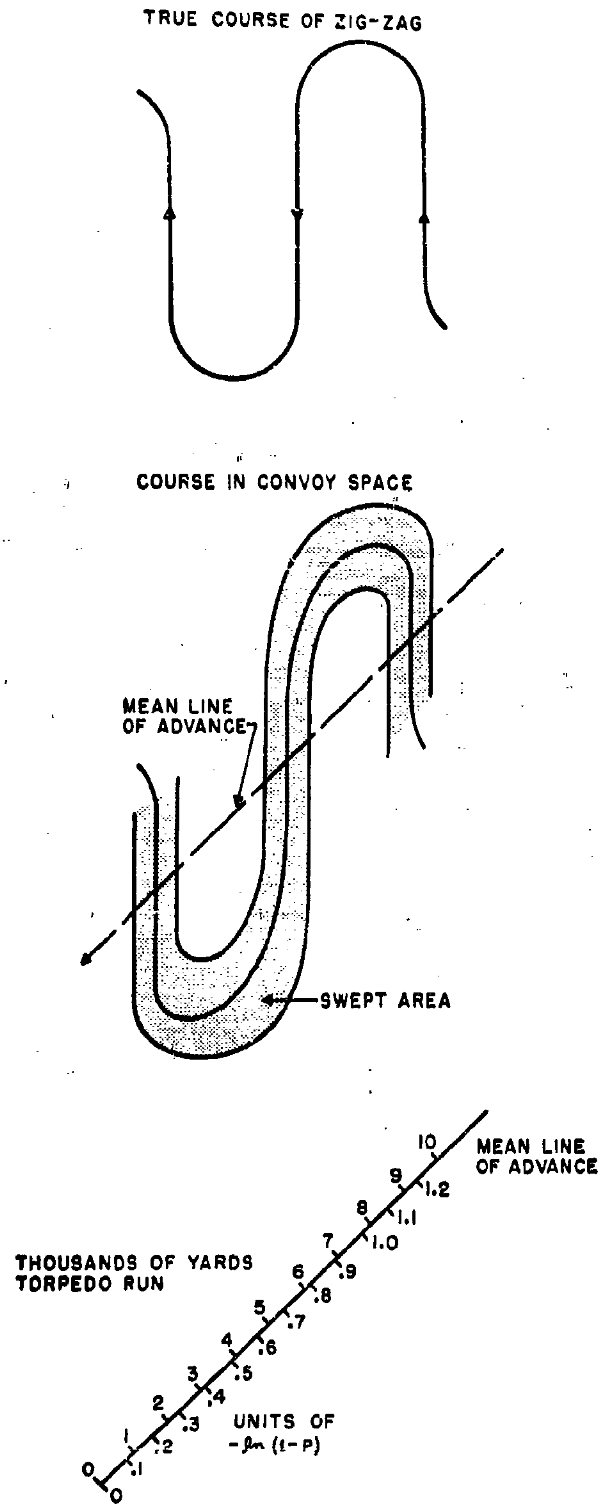


FIGURE 13. Sweep rate for zigzag torpedo.

ure 11 we can read off the value of $-\log(1 - P)$, in this case about 0.32. The corresponding value of P is about 0.27. In order to carry out a number of such calculations for different positions about the convoy, we prepare a diagram which has torpedo paths at every 10 degrees, similarly scaled off, and draw in what might loosely be called *isoprobability contours*. Such a net or coordinate system is shown in Figure 12. By placing the convoy (drawn on transparent paper) over it we can tell at a glance what the submarine's best shot is from any point by picking the one along which the convoy has the greatest "length" measured in that coordinate system. That length also tells us what the submarine's chance of a hit is. One such coordinate system will do for all firing positions, torpedo track angles, and sizes and shapes of convoy. A new system must be drawn for each density, size, and speed of ship in convoy, and for each new type of torpedo.

For curly torpedoes the procedure is more complicated. If, however, the torpedo is one which has a periodic type of motion, such as zigzag or circling, an estimate of its effectiveness is not too difficult. We can consider the motion of the torpedo (relative to the convoy) as made up of a straight-line motion along the mean line of advance, with periodic excursions superimposed. By plotting in detail the course and area swept out for one cycle we can determine the area swept per unit distance made along the mean line of advance. This determines an effective sweep width, which we can then use as if the torpedo proceeded in a straight course along the mean line of advance. Thus we draw the torpedo's mean line of advance and scale off distances in units proportional to $1/p_s S'$, where S' is the effective sweep width. We must also scale off along it the distance that the torpedo actually runs so that its maximum range will be properly taken into account. These various steps are illustrated in Figure 13.

Such torpedoes will actually run straight for at least part of the run and may then start to circle or zigzag as the case may be. The probability diagram applying to a straight run is then used out to the appropriate distance, and circling or zigzagging diagram for the remaining yards to run. This is shown in Figure 14.

The chief advantage of this method is that it is quick and flexible. Once the required diagrams have been worked out for any particular type of torpedo, one can readily decide what the submarine's best shot is from a given position, and estimate his chance

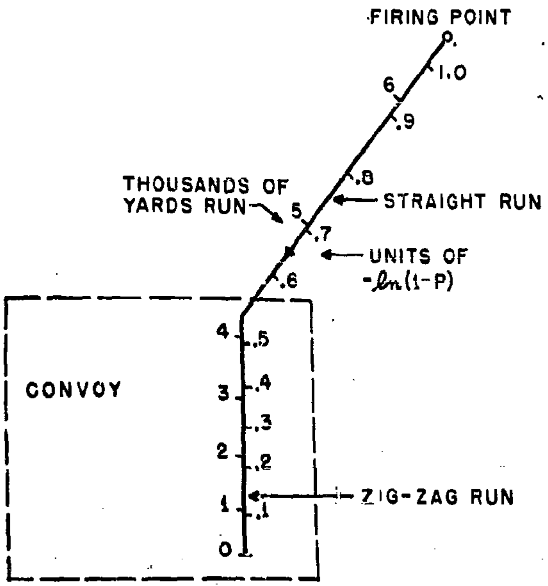


Figure 14. Torpedo which runs straight and then zigzags.

of scoring a hit. In a short time this probability function can be mapped for all points around the convoy. The approximations that are involved should be fairly good in most cases. When the range at which the torpedo is fired is very short (less than 1,000 yards), the actual chance of a hit may be higher than this method would indicate, since the submarine can single out an individual ship and aim the torpedo at it with a good chance of hitting that particular ship. When the range at which the torpedo is fired is very long (greater than 5,000 yards), the actual chance may be lower because an error in firing the torpedo might cause it to miss the convoy altogether. In the intermediate range, however, the probability calculated in this way should be fairly accurate.

It is possible to reduce the probability at long ranges to more reasonable values by introducing a very simple $\phi(a,a_0)$, which gives a rough approximation to the actual torpedo firing errors. For instance one can take ϕ as defined for three angles, the optimum and 10 degrees to either side of it, and equal to one-third at each angle. The probability of a torpedo hit for each of these three angles can readily be averaged to give a rough figure including the effect of firing errors.

Figures 15 to 17 are given as examples of the type of results which are obtained by calculations of this sort.

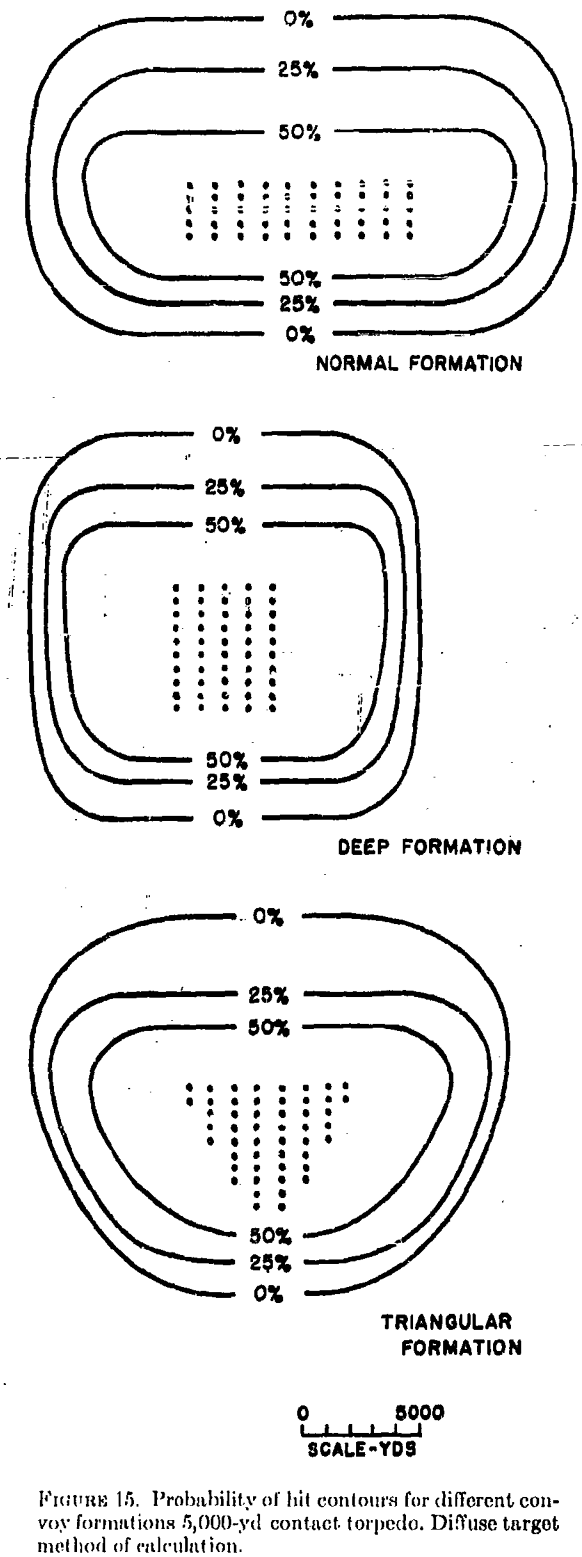


FIGURE 15. Probability of hit contours for different convoy formations 5,000-yd contact torpedo. Diffuse target method of calculation.

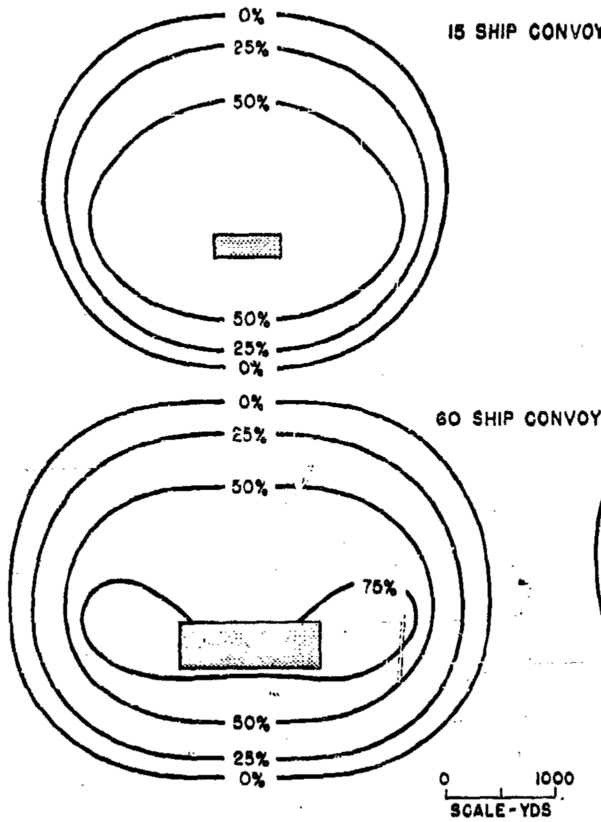


FIGURE 16. Probability of hit with 15,000-yd circling torpedo. Random sweep method of calculation: no error.

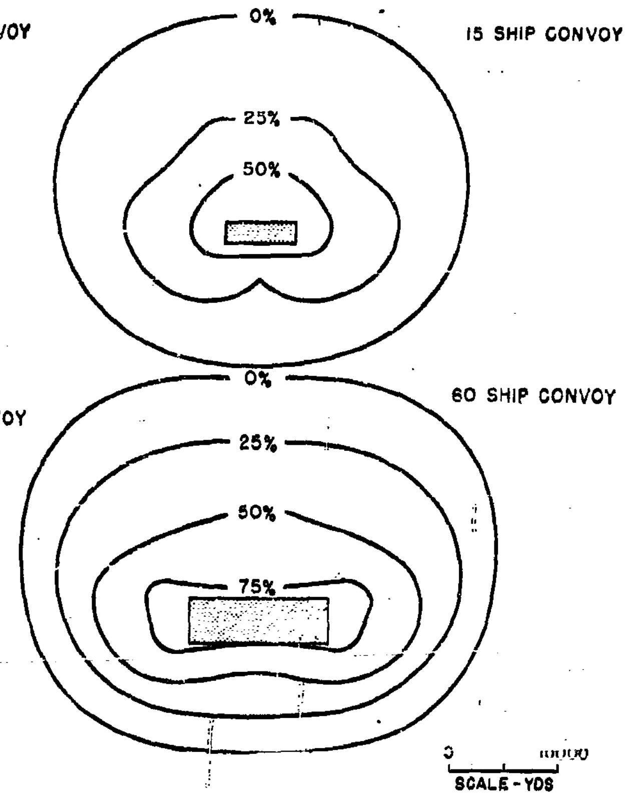


FIGURE 17. Probability of hit with 15,000-yd circling torpedo. Random sweep rate method of calculation: "10° error."

8.4 THE PROBABILITY OF HITTING A SINGLE SHIP OF A TASK FORCE

Ship formations with large spacing between ships, as in a task force, allow a relatively simple method of determining torpedo probability curves similar to those shown in Figures 15 to 17. A submarine firing torpedoes at such a formation would presumably fire at a specific target, and the increased probability of hitting due to neighboring ships would be small. The following method, which involves the probability of hitting a single ship, can be used:

$$\omega = \frac{S}{r} \text{ in radians,}$$

$$S = l \sin \theta,$$

$$\omega = \frac{180}{\pi} l \frac{\sin \theta}{r} \text{ in degrees.}$$

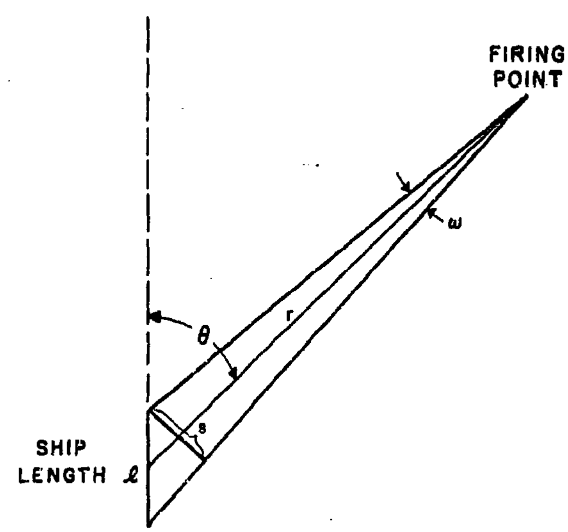


FIGURE 18. Angle subtended by ship.

Assuming a normal distribution of firing errors, let

$\sigma = \sigma(\theta)$ be the standard deviation of firing errors in degrees at bearing θ (relative space),

$p(r,\theta)$ = probability of obtaining a hit from range r at bearing θ.

Then

$$p(r,\theta) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\omega/2}^{\omega/2} e^{-x^2/2\sigma^2}dx$$

$$= \frac{2}{\sigma\sqrt{2\pi}}\int_0^{\omega/2} e^{-x^2/2\sigma^2}dx. \qquad (1)$$

Setting $y = \frac{x}{\sigma\sqrt{2}}$,

$$p(r,\theta) = \frac{2}{\sqrt{\pi}}\int_0^{\omega/2\sqrt{2}\sigma} e^{-y^2}dy$$

$$= \operatorname{erf}\frac{\omega}{2\sqrt{2}\sigma}.$$

Assume that it is desired to find the probability contour corresponding to the chance of hitting $p(r,\theta) = C$. To find the values of r,θ for $p(r,\theta) = C$, let erf $m = C$. Therefore,

$$\frac{\omega}{2\sqrt{2}\sigma} = m$$

$$\frac{180}{\pi}\, l\, \frac{\sin\theta}{r}\,\frac{1}{2\sqrt{2}\sigma} = m \qquad (2)$$

$$r = \frac{180 l}{2\pi m\sqrt{2}}\,\frac{\sin\theta}{\sigma}.$$

For example, for $C = 0.25$, $m = 0.225$, and if $l = 250$ yards,

$$r = 22{,}500\,\frac{\sin\theta}{\sigma}.$$

This reasoning is based on Figure 18 which is drawn in space moving with the velocity of the ship (the same as the task force) and thus r is the length of the relative track of the torpedo. In Figure 19, on the other hand, s (which has nothing to do with the S of Figure 18) is the length of the torpedo's track on the ocean. This figure shows the relation between the various angular errors, i.e., the standard deviations in angle. The two small lengths marked dx are parallel and equal, and perpendicular to the torpedo's ocean track s.

From Figure 19 the law of sines gives the following relationship:

$$\sin d = \frac{u}{v}\sin\theta$$

$$= k\sin\theta, \qquad (3)$$

where d is the the lead angle, θ the angle off the bow, u the ship speed and v the torpedo speed. It is seen that if v is assumed to be known, then the deviation of d from the true value will be due to errors in estimating u and θ. Let Δu and $\Delta\theta$ be the errors in estimating u and θ respectively, and let Δd be the corresponding error in the value of d as computed from the above formula. Then, to a first order of approximation,

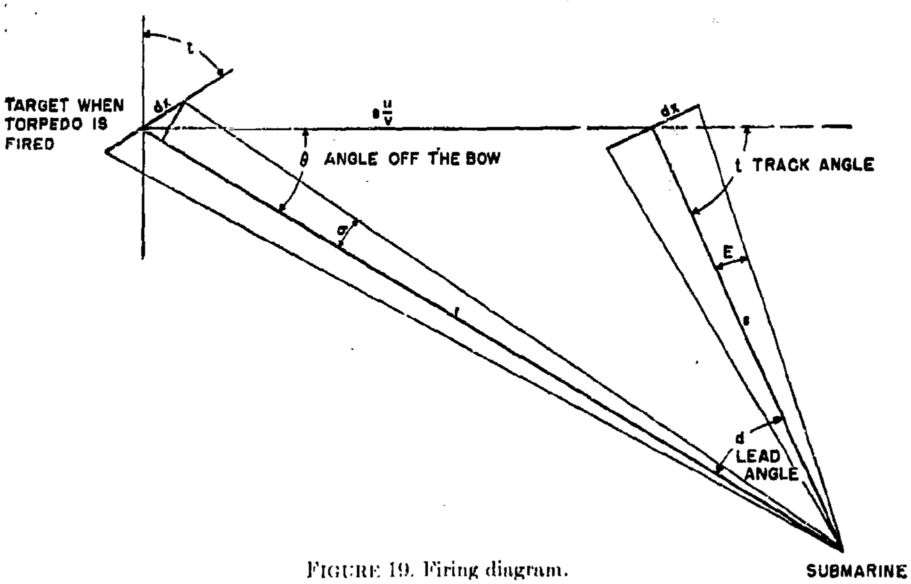


FIGURE 19. Firing diagram.

$$\Delta d \cos d = \frac{\Delta u}{v} \sin \theta + \frac{u}{v} \Delta\theta \cos \theta$$

$$= \frac{u}{v}\left(\frac{\Delta u}{u} \sin \theta + \Delta\theta \cos \theta\right).$$

Since

$$\cos d = \sqrt{1 - \sin^2 d} = \sqrt{1 - k^2 \sin^2 \theta},$$

we have

$$\Delta d = \frac{k\left(\frac{\Delta u}{u} \sin \theta + \Delta\theta \cos \theta\right)}{\sqrt{1 - k^2 \sin^2 \theta}}. \qquad (4)$$

In addition, we should add another term Δi to account for minor equipment and personnel inaccuracies. If the errors Δu, $\Delta\theta$, and Δi are independent and are distributed normally with standard deviations E_u, E_θ, and E_i respectively, then the standard deviation E of the normal distribution of d is given by the formula,

$$E^2 = \frac{k^2\left[\left(\frac{E_u}{u}\right)^2 \sin^2\theta + E_\theta^2 \cos^2 \theta\right]}{1 - k^2 \sin^2 \theta} + E_i^2.$$

Since it has been found that E_u/u often equals E_θ, the above formula may be written

$$E^2 = \frac{k^2 E_\theta^2}{1 - k^2 \sin^2 \theta} + E_i^2. \qquad (5)$$

Before equation (2) can be applied, however, the relationship between E and σ must be determined. From Figure 19,

$$dx \sin (\theta + 90 - t) = r\sigma,$$

$$sE(\sin \theta \sin t + \cos \theta \cos t) = \frac{s \sin t}{\sin \theta} \sigma,$$

$$E = \frac{\sigma \sin t}{\sin \theta\, |\cos (t - \theta)|},$$

$$= \frac{\sigma \sin t}{\sin \theta \cos d},$$

$$\sigma = \frac{\sin \theta \cos d}{\sin (\theta + d)} E, \qquad (6)$$

$$\sigma = \frac{\sin \theta \cos d}{\sin \theta \cos d + \cos \theta \sin d} E,$$

$$\sigma = \frac{E}{1 + \frac{\tan d}{\tan \theta}}.$$

Using equations (5) and (6) to determine values of σ for various values of θ, equation (2) may then be applied to determine values of r corresponding to the desired probability contour. Following through on the example to find the contour for $C = 0.25$, the following table can be computed:

θ	30°	60°	90°	120°	150°
E	7.3°	7.7°	7.9°	7.7°	7.3°
σ	5.4°	6.3°	7.9°	9.9°	11.6°
$r = 22{,}500 \frac{\sin \theta}{\sigma}$	2,080 yd	3,100 yd	2,850 yd	1,970 yd	980 yd

(u = 18 knots, v = 43 knots, E_θ = 17°, E_i = 0)

From the data in the table the curve $p(r,\theta) = 0.25$ can be drawn as in Figure 20.

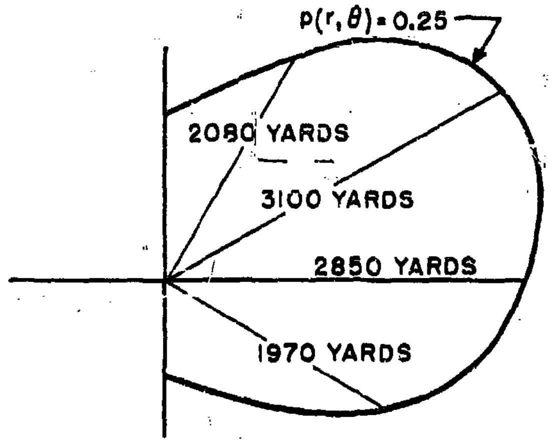


FIGURE 20. Danger zone template.

This curve represents the locus of points from which a submarine firing a 43-knot torpedo can expect a 25 per cent chance of hitting a ship 250 yards long, located at the origin and traveling 18 knots. A template, drawn to the proper scale and having the shape of this curve, can be used to determine the probability contour about a task force. *It must be remembered that this method does not consider the effect of the browning shot.* Figure 21 illustrates the use of such a template for a task force in a circular disposition.

The envelope of all the individual curves is the desired contour. (This method assumes that the submarine aims its torpedo at the optimum point, for the submarine, in the disposition and that there is a ship in the position aimed at.) The placement of the screen about the screened unit once the probability contours are known is described in the remaining pages of this chapter.

Only straight run torpedoes are envisaged in this treatment. Curly or homing torpedoes against ships

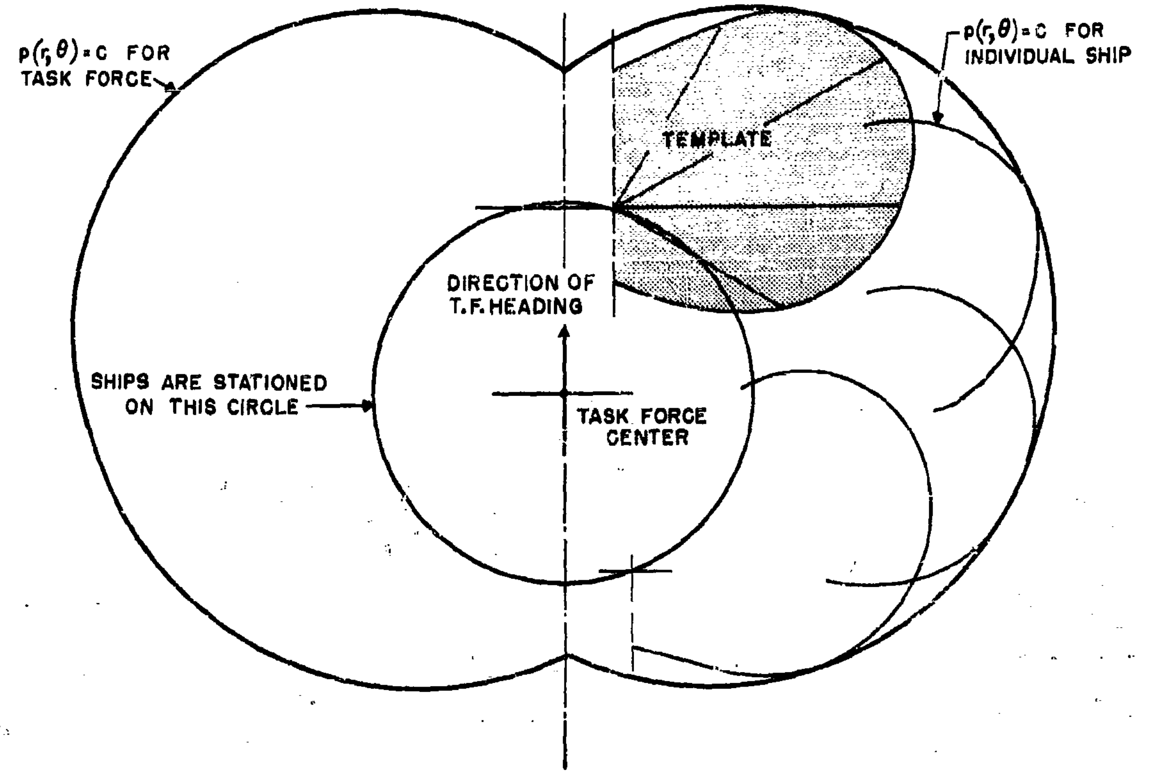


FIGURE 21. Use of template.

of a task force would require a combination of considerations, those of Section 8.3 and, possibly, detailed knowledge of the torpedo characteristics. Such latter fall outside the scope of the present work.

8.5 THE SUBMERGED APPROACH REGION

If a submarine is to have the chance C of scoring a hit with one torpedo, it must reach a point on the curve $p(r,\theta) = C$. Since this will be within visual or radar range of the convoy, it must make its approach to this curve submerged. Let its submerged speed be z, the speed of the convoy being u, and assume that,[a] as in the case of a normal submarine, $z < u$. Then it is not necessarily possible for the submarine to attain the curve $p(r,\theta) = C$. The positions on the ocean from which this is possible fill out a region called *submerged approach region* R_C, constructed as follows: Draw two tangents to the curve $p(r,\theta) = C$, the one on the right making an angle of $\psi = \sin^{-1}(z/u)$ degrees to starboard (relative), that on the left the same relative angle to port, both measured from the convoy's course. The included forward region is the

[a] High submerged speed submarines require separate consideration, which would be beyond the scope of this work.

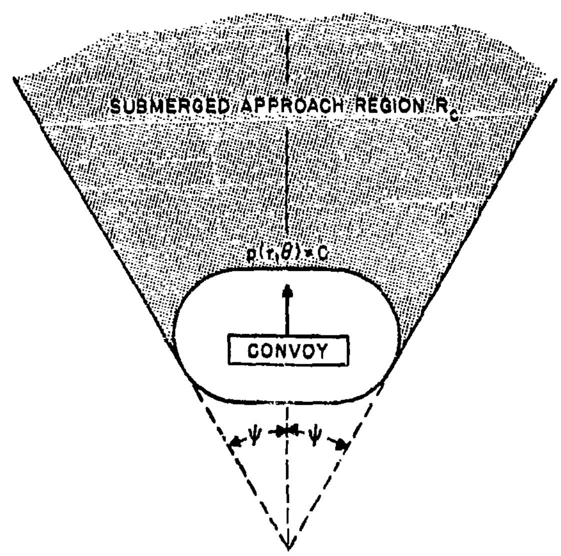


FIGURE 22. Submerged approach region.

submerged approach region (Figure 22). The angle and the tangents are the limiting (submerged) approach angle and lines, respectively. All this has been considered in Section 1.3.

8.6 THE PLACING OF THE SCREEN

Referring to Figure 22, it is obvious that the task which the sonar screen has to accomplish is *at most* to intercept submarines which come from the submerged approach region R_0 (here R_0 is drawn corresponding to $p(r,\theta) = 0$, the limit of the torpedo danger zone); for other submarines do not constitute an immediate danger, as long as they remain submerged. Inasmuch as the line efficiency of the screen is greater the more escorts there are per unit length, the best way of intercepting submarines entering the torpedo danger zone from R_0 is to dispose them in the shortest line connecting the limiting approach lines and lying outside the danger zone and inside R_0. Such a line is constructed by stretching a string around the forward part of the danger zone, its two ends being on the limiting approach lines and perpendicular to them (S_0S_0' in Figure 23).

Such a screen would give 100 per cent protection against submerged submarines if it were perfectly tight, i.e., if it had 100 per cent line efficiency. Unfortunately the distances involved require S_0S_0' to be so long that no normally available number of escorts can provide a screen with anything like such a high line efficiency. The efficiency might, for example, turn out to be only 15 per cent, which represents the chance of preventing the submarine from approaching to within easy torpedo range. Then consideration must be given to defending less of the torpedo danger zone with a shorter, and hence tighter, screen. Section 6.4 discusses the spacing between the screening units, and presents a method of determining the tightness of a screen having a given spacing. If, for example, C = 10 per cent, the level curve $p(r,\theta)$ = C = 0.1 being smaller than $p(r,\theta) = 0$, a screen along S_CS_C' would be shorter and hence tighter than S_0S_0'. If its line efficiency were, for example, 25 per cent,

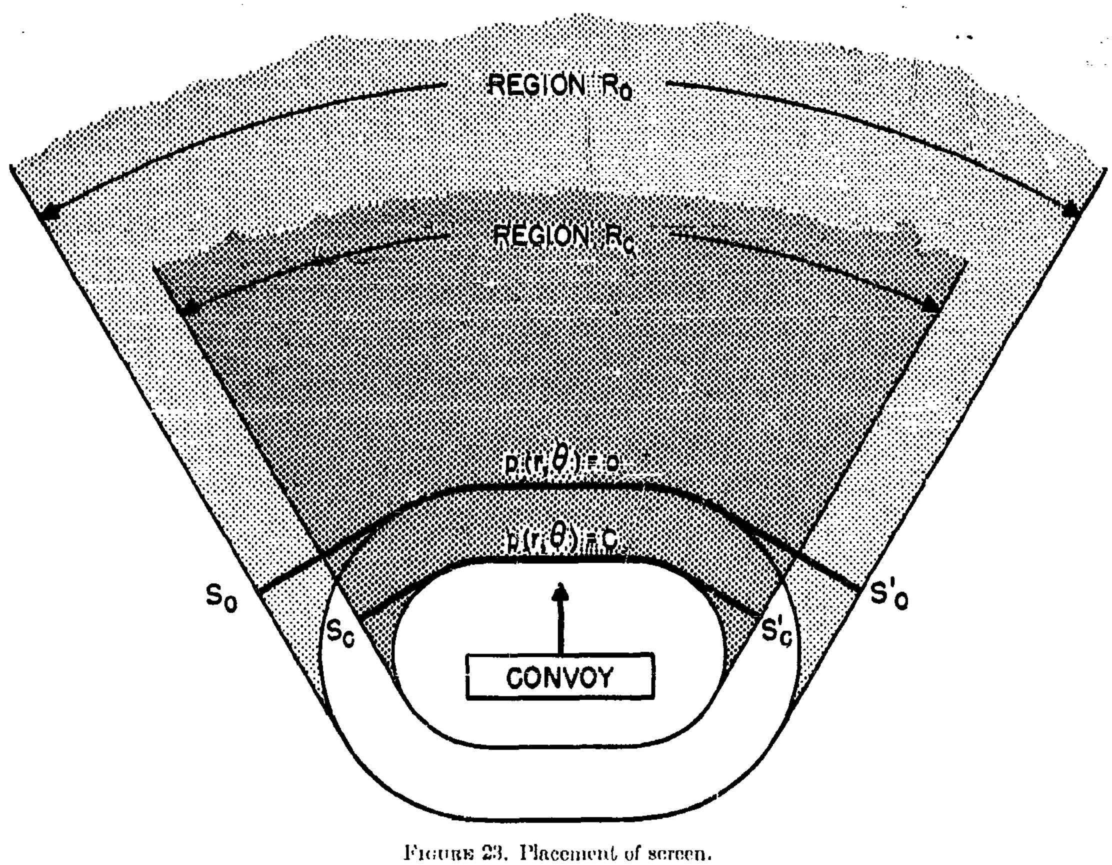


FIGURE 23. Placement of screen.

the submarine would have a *less* favorable chance of hitting the convoy than when the screen S_0S_0' was used. If it fires from outside the screen, its chance of a hit is, at most, 10 per cent; if it attempts to cross the screen so as to be able to fire from a very favorable position, its chance of success is possibly as great as 75 per cent. Obviously the submarine would do the latter (assuming that it is indifferent to its own safety and merely tries to maximize the chance of a hit). But in the previous case this course of action would have given an 85 per cent chance of success for the submarine. Thus S_CS_C' would be a better screen. This contraction of the screen must, however, not be carried too far. If $C = 80$ per cent, the level curve $p(r,\theta) = C = 0.8$ might well give a high (e.g., 90 per cent) line efficiency. The submarine would simply fire from outside the screen and not attempt to penetrate it and would have as much as an 80 per cent chance of making a hit. What value of C must be chosen to give optimum results?

Assuming always that the only consideration governing the submarine's behavior is the desire to make a hit and that it has no primary concern for its own safety, then its best course of action is to attempt to penetrate the screen S_CS_C' whenever $1 -$ (line efficiency of S_CS_C'), is *greater* than C, and will fire from just outside the screen if C is *greater* than $1 -$ (line efficiency of S_CS_C'). In either case its probability of scoring a hit (assuming that once it gets through the screen undetected it can certainly make a hit) is the *greater* of the two quantities $1 -$ (line efficiency of S_CS_C'), C. The situation is visualized by the graph of each of these quantities regarded as functions of C (Figure 24). The submarine's chance of hitting is evidently represented by the heavy line, i.e., $1 -$ (line efficiency S_CS_C') for values of C less than the intersection of the two curves, and C itself to the right of C_0. The optimum screen is the one corresponding with that C which gives the submarine the least chance of hitting, $S_{C_0}S_{C_0}'$. This leads to the principle:

To obtain the best screen, use a curve S_CS_C' of the type shown in Figure 23, and bring it in (i.e., increase C) until the chance of crossing it undetected just equals the chance of scoring a hit from a point just outside it.

There are several qualifications to be made before accepting the above result.

Firstly, submarines actually do give consideration to their own safety; thus, with the screen $S_{C_0}S_{C_0}'$ it would be more favorable to them to fire from outside the screen than to try to cross it. This would continue to be true even if $1 -$ (line efficiency S_CS_C') is somewhat greater than C. Hence, from this point of view, the "best screen" would be somewhat farther out than $S_{C_0}S_{C_0}'$, just how much farther is a difficult matter to estimate. Exactly the same reasoning can be made in different words, as follows: If we are going to have a certain chance of having one of our ships torpedoed, we would prefer to have a greater chance of getting the submarine; the best S_CS_C' should be a little farther out than $S_{C_0}S_{C_0}'$. In whichever form this reasoning is given, it reposes on the fact that a submarine firing outside the screen is less likely to be

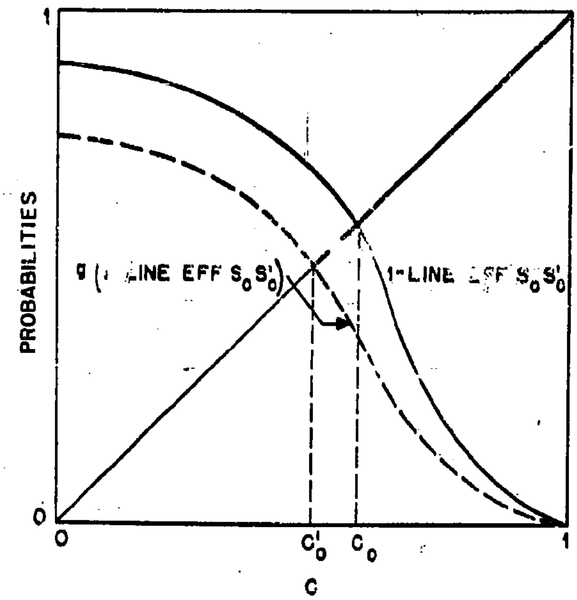


FIGURE 24. Optimum tactics diagram.

attacked than one which tries to cross the screen first.

Secondly, it is not strictly true that once a submarine has crossed the screen undetected it is sure to score a hit. Again this tends to make the "best screen" somewhat farther out than $S_{C_0}S_{C_0}'$. If the chance of a hit at close range is g, $(0 < g < 1)$, to carry through the preceding reasoning we must replace the graph of $1 -$ (line efficiency S_CS_C') by their product $g[1 -$ (line efficiency S_CS_C')] plotted in the dotted curve of Figure 24, and thus reach the best screen $S_{C_0'}S_{C_0}'$.

Thirdly, the submarine may be supposed to have more than one torpedo, whereas it has been assumed implicitly in the foregoing probabilities that only one torpedo was involved. Having n torpedoes would enlarge the level curves but also increase the damage done by a close-range submarine, and these two effects would operate in opposite directions.

Fourthly, it was assumed that a submarine encountering the screen does so with equal probability over the screen's entire length. Yet it is markedly advantageous to the submarine to operate near the sides of the screen so as to facilitate its escape after firing. And the normal tracking procedure used by submarines tends to bring them into contact nearer the front side than the front center of a convoy of any size. For both these reasons it is important to avoid lessening the screen's line efficiency near its ends, i.e., abeam.

Fifthly, in the case of screening a fast ship or task group which may be zigzagging or maneuvering radically, the limiting approach angle is increased and the limiting lines are spread out, and hence the screen must extend through a greater angle off either bow. But the principles developed earlier remain the same.

The extreme case is that of the fast carrier task force which must be prepared to change its course radically, even to backtrack at short notice (e.g., in order to launch planes into the wind, or to avoid or surprise the enemy, etc.). Circular screens around the whole task force are frequently used in such cases, either with equally spaced escorts, or, more efficiently from the antisubmarine point of view, with closer spacing in the forward parts of the circle.

Even when as radical a measure as a circular distribution is not necessary, all turning of the convoy into unswept waters must be avoided. The screen should be extended in the direction the convoy expects to turn, so as to detect any submarines possibly present therein. This is particularly important in view of the tendency of tracking submarines to accumulate along the flank: They are surfaced while tracking, but submerge and become a danger when the convoy turns.

Finally, it may be objected that the reasoning upon which the choice of $S_{C_0}S_{C_0}'$ was based appeared to assume that the submarine knew the values of the various probabilities involved, a thoroughly unrealistic assumption. Actually this does not invalidate the reasoning. We were merely calculating the chance of success of the submarine if it did the best thing from its point of view. Its ignorance can only result in its taking a, for it, less favorable course of action, i.e., it will diminish its chance of success. Thus our reasoning subsists but does not attempt to figure on the chance of the enemy's making a blunder. To figure in such a chance would be to carry the discussion to a higher order of tactical complexity, not the object of the present exposition.

In order to incorporate the considerations raised by these qualifications into the problem of fixing the optimum screen, a somewhat less artificial criterion of advantage would have to be used. We have concentrated our attention on the *probability of a hit* of one submarine with one torpedo, and shown how to minimize it. Actually one might want to minimize the number of ships sunk per submarine lost (submarines using salvos of torpedos and many sorts of torpedos, etc.), i.e., to posit a more realistic situation. But it does not appear that as far as the present general treatment is concerned such a more detailed and complicated study would materially alter the conclusions.

8.7 PATROLLING OF STATIONS

When the number of escorts is insufficient to provide even a moderately tight screen without closing to unduly short distances of the convoy, it is customary for the escorts to "patrol their stations," that is, to take a course which causes their position to oscillate about their station sometimes quite a distance (e.g., 500 yards) to the right and then to the left of the point (fixed relative to the convoy) which represents the assigned station. The reasoning upon which this process is based is the following.

When there are too few escorts, the distance between two adjacent ones will be such that a submarine has a very good chance of passing through the screen undetected provided it goes about midway between the escorts; in other words, the screen has "holes." There are two possibilities: Either the submarine knows where such holes are or else it does not. If it does, it can profit by their presence, and thus the strength of the screen will have to be judged by its *weakest* point. If it does not know, it will enter the screen at a randomly chosen point, and thus the strength of the screen would be measured by its *average* strength, i.e., the average of its probability of detection. Now the object of patrolling stations is to deprive the submarine of the possibility of utilizing the holes, since, when it is near the screen, it is proceeding submerged and without being able safely to use its periscope. Of course the patrolling must have an irregular or random character. Thus patrolling stations make the second of the above possibilities the actual one. The average tightness of the screen is the valid index of effectiveness, and, low though it be, it is much higher than the probability of detection in the hole.

Figure 25 shows the situation graphically. The ordinate represents the probability of detection at a point along the screen represented by the abscissa. Patrolling randomizes the situation with regard to the submarine, thus replacing the original curve (a) by the average ordinate horizontal line (b). The area under (a) equals that under (b).

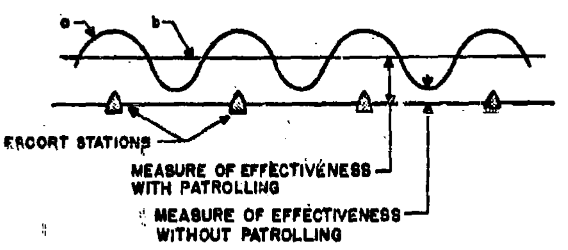


FIGURE 25. The effect of patrolling of stations.

8.8 PICKETS

The screen considered here is a line screen. One of a different sort and giving an effect of defense in depth (e.g., by alternate staggering of ships) would have the disadvantage of creating wake interference and difficulties of maneuvering. Nevertheless a line screen can be given some of the attributes of defense in depth by supplementing it with escorts stationed as *pickets* (1, 2, 3, 4), Figure 26. These pickets accomplish the following tasks.

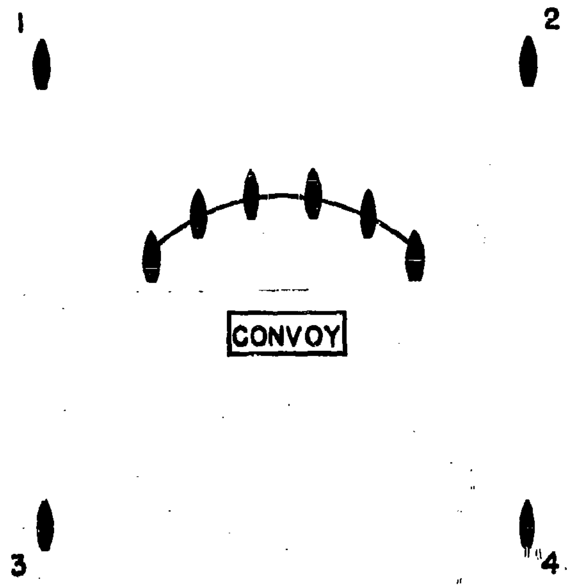


FIGURE 26. Stationing of pickets.

1. They act as visual or radar observers for surface submarines; this is particularly important when air coverage is lacking, or only present in forward sectors.
2. They investigate suspicious contacts and aid in making attacks upon submarines which may be detected by the screen or by their having attacked a ship in the convoy.

Chapter 9

AERIAL ESCORT

9.1 THE TACTICAL SITUATION

When a formation of ships (such as a merchant convoy in transit, a task force, or task group in cruising disposition) passes through submarine waters (i.e., waters possibly containing hostile submarines) safety can be increased by accompanying the formation with aircraft (carrier or land based) which perform systematic flights in its vicinity. Such flights are the subject of this chapter; they are called the *aerial escort*, the defended ships being termed in all cases the *convoy*. It is to be emphasized that the primary object here is *defensive*—to reduce the danger to the convoy—as contrasted with the primarily *offensive* purpose of flights such as the distant beam searches made by *Aircraft Carrier Escort* (CVE) killer groups, where the destruction of the maximum number of submarines is the primary aim. To put it in slightly different terms: If a submarine is present, the success of the aerial escort is measured by its ability to prevent the submarine from damaging the convoy; while the most satisfactory result is undoubtedly the sinking of the submarine, the escort must also be regarded as successful if its mere detection of the submarine permits the convoy to avoid it, or even if its presence induces the submarine to submerge and remain submerged in a region from which no submerged attack upon the convoy can be delivered.

In order to attack a convoy, a submarine must *first* detect the convoy; *second*, make an approach (usually with tracking) to within firing range of the convoy; *third*, fire its weapons; *fourth*, withdraw to safety. Aerial escort is chiefly instrumental in obstructing the first and second of these operations. Its first function (prevention of detection) shall be called the *scouting* effect; its second (prevention of approach), the *screening*.

In any definite situation there will be a maximum range R at which a submarine, using all its facilities, can detect the convoy. Visual detection of a nonsmoking convoy in ideal weather occurs at a distance limited only by the earth's curvature and atmospheric refraction, but it can be greatly lowered by adverse meteorological conditions and increased by convoy smoking. Radar detection of convoys is altogether dependent on the set but at present does not exceed good daylight visual detection. Hydrophone ranges depend on sound conditions and sea state, as well as size and speed of convoy; under ordinary conditions they are not much in excess of 15 miles, but under ideal conditions with a large fast task force they may attain 50 miles. R, which is the greatest of all these ranges measured from a reference point O fixed in the convoy and thus taking into account convoy size, has in most cases values between 15 and 30 miles. The circle of radius R and center at the reference point O is the *detection circle*. The detection circle may be described as the area within which submarines have an appreciable chance of detecting the convoy, and outside of which they have little.

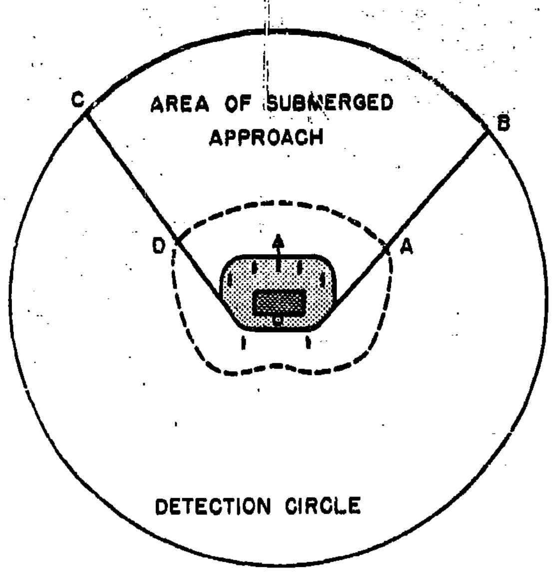


Figure 1. The tactical regions.

Figure 1, which is drawn relative to the convoy (in which, therefore, reference point O remains fixed), shows the tactically relevant features: The convoy and its surrounding *torpedo danger zone* (shaded), the *area of submerged approach* (see Chapter 8), and the detection circle. There is also shown (bounded by dotted lines) the area within which the ships of the convoy and its escort can detect a surfaced submarine. The convoy is heading up the page at the speed made good of v knots. The region $ABCD$ shall be called the *danger zone;* if a submarine has reached it, it is in contact with the convoy and can submerge at will and then make a submerged approach undetected by the aircraft; to reach the danger zone is a usual tactical objective of attacking submarines.

The figure is drawn on the assumption that the submarine's submerged speed is less than the convoy's speed. In the contrary case, the solid curve $CDAB$ may have to be drawn much farther back of the torpedo danger zone—corresponding to the locus within which the endurance of the submarine submerged just suffices to close the torpedo danger zone. Or the curve may be absent entirely, when this endurance is very great. In this case the danger zone will be the ring shaped area inside the detection circle and outside the dotted curve.

In order to have scouting effectiveness, i.e., in order to prevent the submarine from gaining contact with the convoy (so that it will either not know that the convoy is present, or, in case it has been directed to the convoy by means other than its own immediate detection, have no precise knowledge of the position or course of the latter), it is necessary that the aerial escort fly a sufficient amount directly outside the detection circle, particularly in the forward regions, which are those in which submarines are most likely to be encountered by the circle (Section 1.5). A method of evaluating numerically the scouting effectiveness of any given flight plan will be set forth in Section 9.2.

In order to have screening effectiveness, the escort flights must be so made that the tracking and approach procedures normally carried out by a surfaced submarine are impeded. Essentially this means that entrance to the danger zone must be barred, and that this area itself must be covered so that even if a submarine has entered it and submerged, it will have a material chance of being detected should it momentarily surface or otherwise show its presence. The importance of barring the arc BC is due to the fact that a submarine in the path of the convoy but unaware of the latter's presence is likely to be picked up by this arc as it moves over the water with the convoy, by which time the submarine will detect the convoy and then be in a favorable position to submerge and deliver an attack. The importance of barring AB and CD is due to the fact that these will normally be crossed by submarines which have tracked the convoy on the surface; such submarines cannot enter the area of submerged approach nearer to the convoy than A or D since they would be detected by surface units; and an attempt to enter beyond B or C would require their loss of contact with the convoy. A precise method of evaluating the screening effectiveness of a plan will be found in Sections 9.3, 9.4 and 9.5.

When aerial escort accomplishing these results is maintained day and night, maximum protection is obtained. Should it be necessary to discontinue the flights at night, it is important that predark flights of a scouting character be made at greater distances around the convoy in order to clear the waters through which the convoy will pass in the night as well as to detect submarines which may be stalking the convoy, intending to close it at night. The characteristics of such flights will be examined in Section 9.7.

The only practicable flight plans for regular aerial escort (all those which are flown continuously during a length of time, and excluding predark flights) are *periodic* ones. After the period of T hours the plan of flight is repeated, the paths flown during the next T hours being identical (as seen relative to the convoy) to those of the preceding T hours; this continues as long as the plan is flown. The period T should be of the order of one or two hours. Thus the plot of every regular plan relative to the convoy consists of one or more closed circuits, the whole being flown every T hours. The geographic plot is not closed, but is of an advancing recurrent nature, the length of the recurrent figure flown by each aircraft being T times the aircraft's ground speed (or average ground speed, in case of variations in the latter due to wind or other causes).

One assumption is presupposed in the quantitative probability reasoning concerning any aerial escort plan, that *the submarine does not know the plan.* Otherwise any plan using a normally restricted number of aircraft would be ineffective. Actually, it is difficult to imagine that the enemy could ever obtain sufficiently detailed information concerning the plan to be enabled thereby to penetrate undetected.

9.2 THE SCOUTING EFFECTIVENESS

A measure of the effectiveness of any aircraft escort plan as a method of scouting is the *probability* which it affords of detecting (visually or by radar) submarines which are so moving that they will eventually enter the detection circle, detecting them, that is, before they enter the circle.[a] It is assumed that these submarines are proceeding surfaced (or schnorchelling) in a straight course at speed u. The probability will depend on this speed, as well as on the position of the submarine's path. It should be noted that the word "scouting" is used in the present study to indicate a primarily defensive operation, in contrast to a frequent and more offensive connotation of the term.

At this stage it is important to emphasize that submarines outside the detection circle are to be regarded as unalerted and either ignorant of the convoy's existence or insufficiently informed of its position to be enabled to make a systematic approach. Consequently they may be regarded as *random submarines*; all courses are equally likely. And when attention is confined to submarines on courses leading eventually to their entry into the detection circle, they continue to be random submarines but with appropriately altered distribution of courses.

For expediency of computation, the probabilities of interception before entry into the detection circle are evaluated for targets of each *velocity* class (see Section 1.5). In practice this means that, after having estimated the probable cruising speed u of an unalerted surfaced (or schnorchelling, as the case may be) submarine, a track angle ϕ is selected, and attention is fixed on the class C_ϕ of surfaced submarines of speed u which (1) proceed at the selected track angle ϕ, (2) have tracks relative to the convoy which enter the detection circle. Then the probability P_ϕ that such submarines be detected by the aircraft is calculated. This is essentially a problem of barrier patrol evaluation but pertains to the case of barriers in which the aircraft flights, though periodic, are very irregular. The treatment has been given in Section 1.6, and Chapter 7, and we confine ourselves here merely to setting forth a simple method which will usually give sufficient accuracy.

Step 1. Draw the escort plan to scale relative to the convoy, and draw the detection circle.

Step 2. From ϕ, u, and the convoy speed, find the relative course θ and speed w with respect to the convoy (see Section 1.2). Draw the two tangents to the detection circle making the angle θ with the convoy course.

Step 3. Measure the total length L_ϕ of that part of the aircraft path which lies outside the detection circle and between the two tangents drawn in step 2, and is situated on that side of the circle corresponding to entrance into the circle by submarines of class C_ϕ.

The required probability is then

$$P_\phi = 1 - e^{-WL_\phi/2RwT},$$

where W = effective search width for the detection of surfaced (or schnorchelling) submarines by the aircraft (see Section 2.6).

This calculation of P_ϕ is carried out for a set of angles covering the full circle. These may be taken equally spaced either in ϕ or in θ, at for example 15 degree intervals (or larger if less complete information suffices). They are conveniently exhibited by marking each P_ϕ off radially along a line starting at the origin and making an angle $\theta - \pi$ (and hence directed *toward* the incoming submarine) with the direction of the convoy. When the resulting points are joined by a smooth curve, a *polar diagram* (to be called the *scouting diagram*) is obtained. A corresponding polar diagram of P_ϕ plotted against ϕ or $\phi - \pi$ can be drawn, but being less directly related to the relative picture of the flight plan it gives a less clear indication of its scouting tightness.

It is noted that in the case of fast convoys whose speed v is greater than the surfaced (or schnorchelling) cruising speed u of the submarine, the polar curve is a closed loop lying ahead of the polar origin O. This corresponds with the fact that along a given θ (represented as the radial line at angle $\theta - \pi$ with the convoy's course), there are two velocity vector angles ϕ, one corresponding with submarines proceeding toward the convoy, the other to those headed away from it but overtaken by it. Thus the loop is cut in two points by the above radial line. When $v < u$, there is only one point of intersection, the loop enclosing O. When $v = u$, O is on the loop.

Actually, the chief use of the scouting diagram is in obtaining the *scouting coefficient*, now to be defined. In appraising the scouting effectiveness of a plan, consideration must be given to the total detected fraction f of all the submarines whose motion will lead them into the detection circle. The number of submarines of C_ϕ entering the circle per unit time is equal to the area $2Rw$ in relative space which con-

[a] Another measure of effectiveness is the ability of the aircraft to force the submarine to submerge outside the detection circle, and thus greatly to reduce its chance of detecting the convoy. This does not require the aircraft to detect the submarine but, rather, the submarine to detect the aircraft. Since this depends on the amount of flying in the various areas without the circle, in much the same manner as the aircraft's probability of detecting the submarine does, the probability of the latter affords a significant norm of evaluation for the plan. Therefore this will not be considered further.

tains them, times the density of members of C_ϕ (by the randomness, "independent" of ϕ). Thus their number is proportional to $w[w = w(\phi)]$. Of these, a number proportional to wP_ϕ is detected on the average. Hence, the total number detected is proportional to the integral

$$\frac{1}{2\pi}\int_0^{2\pi} w(\phi)P_\phi d\phi,$$

which is conveniently found by taking the arithmetic mean of the numbers $w(\phi)P_\phi$ found as above for equally spaced angles ϕ. If every submarine were detected ($P_\phi = 1$), the above would reduce to

$$\frac{1}{2\pi}\int_0^{2\pi} w(\phi)d\phi.$$

To obtain the required fraction, the first expression is divided by the second, giving

$$f = \frac{\int_0^{2\pi} w(\phi)P_\phi d\phi}{\int_0^{2\pi} w(\phi)d\phi};$$

a weighted mean of the quantities P_ϕ. This can be computed by taking the equally spaced angles ϕ_1, ϕ_2, $\cdots$ and obtaining

$$f = \frac{w(\phi_1)P_{\phi_1} + w(\phi_2)P_{\phi_2} + \cdots}{w(\phi_1) + w(\phi_2) + \cdots},$$

or by using previous results [Chapter 1, equation (4)] to evaluate the denominator. In the notation of the reference, this gives

$$f = \frac{\text{arithmetic mean of } w(\phi_1)P_{\phi_1}, w(\phi_2)P_{\phi_2}, \cdots}{(u + v)\ (2/\pi)E(\sigma)},$$

where v is the convoy's speed.

The number f, called the *scouting coefficient*, is in the nature of a figure of merit of the scouting of the aerial escort plan.[b] The values are reasonably insensitive to possible misestimations of u but depend very materially upon the value of R.

Certain modifications of the above procedure may be useful in special situations. Thus in the case of certain submarines having extra batteries, a high submerged speed may make the submerged overtaking of the convoy possible, although only at relatively short distances, on account of limited battery endurance at such speeds. This means that the danger zone may be bounded in the rear by a curve passing a few miles, e.g., ten, to the rear of the convoy and gently concave in its direction, and intersecting the detection circle at points near each beam. Then an evaluation of the scouting diagram and coefficient, with the detection circle replaced by the composite curve consisting of the forward part of this circle joined to the above rear limiting approach curve, may be useful. Such an evaluation proceeds along similar lines to the earlier one, except that the distance D between the tangents now depends on ϕ: $D = D_\phi$ rather than $D = 2R$ must be used in the denominator of the exponent in the formula for P_ϕ. This procedure has a certain inaccuracy, inasmuch as it counts submarines inside the detection circle but behind the rear limiting approach curve as random, whereas they can be expected to have gained enough knowledge of the convoy's location to change from an accidental to an intended approach. There are cases, however, where this inaccuracy may be ignored.

9.3 THE SCREENING EFFECTIVENESS

A measure of the effectiveness of an aircraft escort plan as a screen is the probability which it affords of detecting surfaced submarines[c] which are making a systematic approach to the convoy. Such submarines are not only alerted but they have a fairly precise knowledge of the convoy's location, course, and speed. Their approach may include a tracking procedure, and some form of closing (e.g., along a normal approach course or a collision course) to a favorable position for submerging and thence making a submerged close approach and attack. Their surfaced approach course may therefore be quite complicated, and since the probability of detection depends on its geometrical shape, as well as upon the phase of the aircraft's motion (i.e., its position on its circuit when the submarine is at a given distance from O), the evaluation of all such probabilities is not feasible. Hence the importance of the following line of argument, intended to establish as a useful norm of screening effectiveness the probabilities of detection calculated for *collision courses only*. With the

[b] Theoretically, a plan could be designed so that f is maximized (subject to the flying restrictions). The direct determination of a plan on the basis of this principle seems to be less profitable than the indirect method of drawing up a practicable plan and then checking its scouting value as above, possibly altering the plan (without sacrificing other important features) for scouting improvement.

[c] Or submarines which by schnorchelling or using their periscope are affording the possibility of aerial detection.

reference point O in the convoy as center, a circle is drawn so large that it includes the whole aircraft escort track so far within it that there is no probability that any of the aircraft detect a submarine outside it. This circle, in general much greater than the detection circle, will be called the *reference circle*. Any approaching submarine will cross the reference circle at some (last) point A at the bearing θ relative to the convoy's course. This is shown in Figure 2,

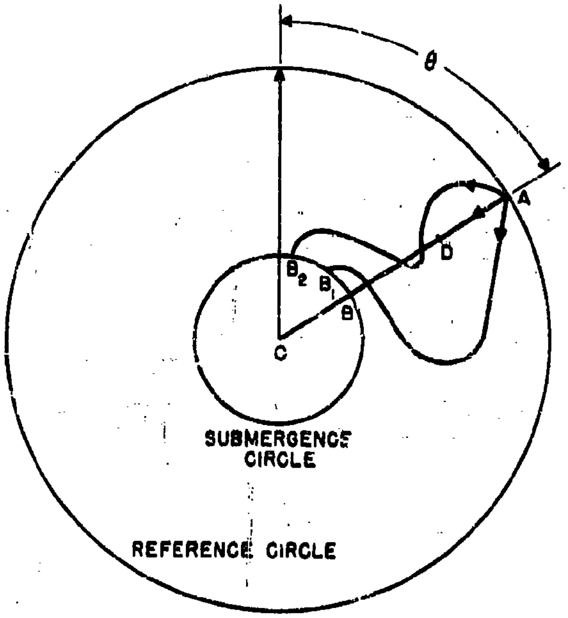


FIGURE 2. The reference figures in an approach.

drawn relative to the convoy. One characteristic of the submarine's path is θ; another is the distance from the convoy at which it submerges. Figure 2 shows three paths, AB (the collision approach), AB_1, and AB_2, all corresponding to the same θ and same distance from convoy of point of submergence (thus B, B_1, B_2 are all on the same "submergence circle").[d] How do such paths compare in their probabilities of detection? Again it must be emphasized that the submarine is assumed not to know the aerial escort plan: the most it knows is that aerial escort is being flown. Otherwise aerial escort with any normal number of aircraft could not be assumed to be useful, inasmuch as the submarine could be presumed to know how to get in undetected. Consequently, the paths AB, AB_1, etc. are randomly situated with respect to the aircraft paths. Now of all the submarine paths AB, AB_1, etc., the *collision course AB is the one which takes the least time for the submarine to get from A to the submergence circle.* If therefore, the density of air coverage is substantially the same throughout the regions through which AB, AB_1, etc. can be expected to pass, the probability of detection for AB is *less* than that for AB_1, AB_2, etc. On account of the intention of tracking the convoy before closing, and then of entering the submerged approach region (Figure 1), many submarine paths may be expected to be of the type of AB_1, hence, if the probability of detection of AB is to represent a genuine conservative estimate (lower limit) of the probabilities of detection of AB_1, AB_2, etc., *the flight plan must provide adequate barring of crossing the line AB of* Figure 1.

Consider now two collision courses: AB and AO (Figure 2). The probability of detection of AB is obviously less than that of AO (at most equal to AO, in the case when all flying is done well outside the submergence circle). If the density of flight were essentially the same at all the places through which the collision course of bearing angle θ passes, out to a point D, and then fell to zero, the probability of detection of AB would be given by multiplying that for AO by the factor BD/OD. If there were more flying farther out along OD than at points nearer to O, the factor would be *greater* than BD/OD; if less flying far than near, the factor would be *less* than BD/OD, and, indeed, the probability of detection of AO gives little information about that of AB when the flying is highly concentrated near the convoy; but such plans are bad plans (since they duplicate the sighting operations of the surface escorts and since they cover regions where the submarine can be expected to be already submerged) and need not concern us here. Their exclusion is part of the task of the design of a plan (Section 9.6 below). Thus, with reasonable escort plans, the probability of detection of AO is a satisfactory index of the tightness of the screen along the line of bearing θ. We may add that the probability for AO is the same (on account of the way in which the reference circle was drawn) as that for a collision path of bearing θ but extending to infinity instead of to A.

In conclusion it is here posited:

Under the proviso that (1) the entrance to the submerged approach region be adequately guarded

[d] We are assuming that the boundary of the submergence region is a circle, thus considering only the simplest case. With a tactical situation known in more detail, some other region may be more appropriate. For example, the danger zone has been used in analyzing the plan of Figure 17, but the principles of the reasoning remain the same.

(so that probability of detection of AB_1 is not less than AB), and (2) the flying be not unduly concentrated near the convoy, the probability $p(\theta)$ of detection of a collision course submarine coming from infinity to O along the bearing θ calculated for various representative values of θ is a norm of the screening effectiveness. The polar diagram $r = p(\theta)$ is called the screening polar diagram. Its construction is the subject of Section 9.4.

A modified procedure is to take the polar diagram in which probabilities of detecting collision course submarines outside the danger zone are plotted, instead of probabilities of detection all the way in to O. This occasionally gives useful indications, but its value is mainly in scrutinizing poor plans, i.e., those in which an undue amount of flying is done close in. In a well-designed plan, the diagram as it has been considered earlier should not be too different from the modified case here mentioned.

9.4 THE SCREENING POLAR DIAGRAM—GENERAL CONSIDERATIONS

As in Section 9.2, the period T of the flight plan is ascertained, and the diagram of the plan is drawn in convoy space. The problem is to find the probability $p(\theta)$ of detecting a surfaced submarine approaching on the collision course which in convoy space is the radial line drawn out from O at the angle θ with the convoy course. In order to see precisely what is involved, we shall give the exact formulation of $p(\theta)$, and afterwards study methods for its approximate evaluation.

Let t denote the time (hours) before a particular submarine reaches O (at O, $t = 0$). Marking an x axis in the direction of the convoy's heading (the reference point O in the convoy being the origin) and a y axis in the starboard direction (Figure 3), the submarine's position at any time t is (x_t,y_t), and if $w(\theta)$ is its relative speed, we have

$$x_t = w(\theta)t \cos \theta$$
$$y_t = w(\theta)t \sin \theta,$$

which are, in fact, its equations of motion.

If (X_t^i, Y_t^i) are the coordinates of the ith aircraft at time t, this aircraft's equations of motion are of the form

$$X_t^i = X^i(t),$$
$$Y_t^i = Y^i(t),$$

where the two functions $X^i(t)$ and $Y^i(t)$ are each periodic and have T as a period. Indeed, T is the smallest *common* period of all the functions $X^i(t)$, $Y^i(t)$ for all values of i (i.e., the least common multiple of the periods of all aircraft circuits of the plan). Of course some of the aircraft may repeat their circuit more often than others, so smaller numbers, such as $T/2$, may be periods of some of the pairs of functions $X^i(t)$, $Y^i(t)$. It must be realized that to give the escort plan is to give these functions, and vice versa.

Let

$$\gamma\left(\sqrt{(x - X)^2 + (y - Y)^2}\right)dt$$

be the instantaneous probability of sighting a target (surfaced submarine) at (x,y) by an observer (aircraft) at (X,Y) (see Section 2.2).[c] As in Chapter 2,

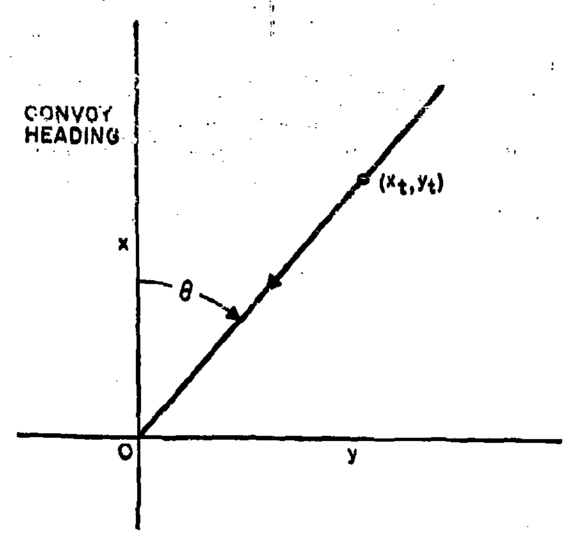


FIGURE 3. Collision approach: the coordinate system.

the probability of sighting the submarine on its entire collision course by the ith aircraft is

$$1 - \exp\left\{-\int_0^\infty \gamma\left[\sqrt{(x_t - X_t^i)^2 + (y_t - Y_t^i)^2}\right]dt\right\},$$

and the probability that at least one of the aircraft sight it is

$$1 - \exp\left\{-\sum_i\int_0^\infty \gamma\left[\sqrt{(x_t - X_t^i)^2 + (y_t - Y_t^i)^2}\right]dt\right\},$$

[c] Effects of target aspect and bearing on this probability are being ignored in thus assuming that γ is a function of the distance between observer and target.

the summation in the exponential being over all values of i which occur (e.g., from 1 to 3 if there are three aircraft flying the plan).

This is not yet the value of $p(\theta)$.

Suppose that a second submarine is following the first at a distance $w(\theta)\tau$ behind it (i.e., τ hours later). If $\tau = T, 2T, 3T$, etc., its chance of detection is the same as before, because of the periodicity of the plan. But if $0 < \tau < T$, it will arrive at the various points of its path when the aircraft are at a different phase in their circuits from those of the previous case, and thus the probability of detection will be different. To find its value, we have but to write the equations of motion of the second submarine (x_t', y_t'),

$$x_t' = w(\theta)(t + \tau) \cos\theta = x_{t+\tau},$$

$$y_t' = w(\theta)(t + \tau) \cos\theta = y_{t+\tau},$$

and substitute these expressions into the above formula in place of (x_t, y_t), i.e., replace (x_t, y_t) by $(x_{t+\tau}, y_{t+\tau})$. Now the value of $p(\theta)$ is the probability of detecting a submarine chosen at random on the line of bearing θ. Its position is characterized by the value of τ, which is uniformly distributed in the interval $0 \leqq \tau < T$. Thus to obtain $p(\theta)$ we must average the *probabilities* obtained as before for each τ; we have then finally

$$p(\theta) = \frac{1}{T}\int_0^T \left\{1 - \exp\left[-\sum_i \int_0^\infty \gamma(r_i)\, dt\right]\right\} d\tau,$$

$$r_i = \sqrt{(x_{t+\tau} - X_t^i)^2 + (y_{t+\tau} - Y_t^i)^2}.$$

The exact computation of $p(\theta)$ would be a formidable task and is not warranted, in view of the fact that any explicit expressions for γ and $w(\theta)$ are at best only very approximate, while at the same time there is bound to be considerable navigational inaccuracy in flying the plan. Fortunately, approximate methods exist whereby $p(\theta)$ can be evaluated. The earliest method was to use a definite range law of sighting and treat the problem graphically. This method, except as giving very crude indications, has had to be abandoned. For it not only gives an unsatisfactory degree of approximation, but, more important, it frequently yields a value $p(\theta) = 1$, thus giving the impression that a perfect barrier could be established (corresponding with the idea of "sweeping the ocean about the convoy clean"). This is an illusion: Every plan gives the submarine a finite (though possibly small) chance of an undetected surfaced passage through the aircraft screen, provided the velocities make such an approach kinematically possible. A much better method, for the visual case at least, is the one based on the inverse cube law of sighting. It is set forth in detail with the aid of examples in Section 9.5. Here we merely outline the general principles of the approximation to $p(\theta)$, whatever be the nature of γ, and thus applying equally to any law of visual or of radar detection.

As the first step in the approximation, suppose that the submarine crosses[f] two partial paths C_1,

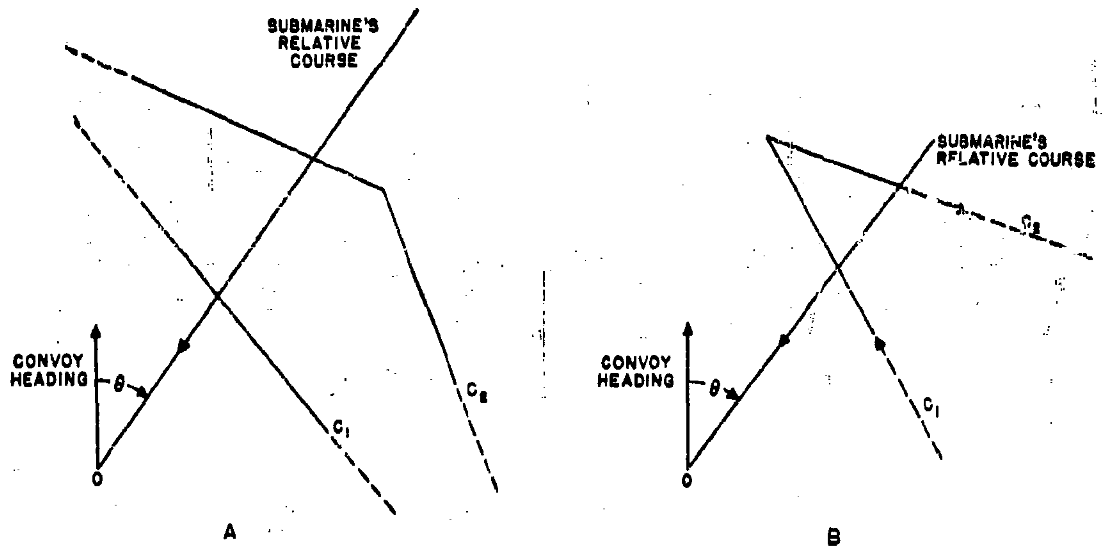


FIGURE 4. Types of paths crossed.

[f] I.e., either actually intersect, or pass in close enough proximity to afford an appreciable chance of a sighting.

C_2, of the plan (and possibly others). This may occur, as in Figure 4A or 4B.

The distinction is that in (A) the paths are either flown by different planes or by the same plane at considerably different epochs, and therefore, because of the irregularities in the flights and in the submarine's motion, etc., *the detection by the aircraft on C_1 and the detection on C_2 may be regarded as statistically independent events*; whereas in the case (B) *such independence may not be assumed*. Furthermore, detection on any of these paths will not occur (i.e., occur with but a negligible probability) at a considerable distance from the submarine's path, for example on the dotted portions of C_1 and C_2 and beyond. Thus we may at our convenience either suppress all this more distant part of the paths or alternatively (when it is mathematically simpler) produce the straight lines to infinity away from the submarine's course. It then becomes a much simpler problem to calculate the probability of detection by C_1 and then by C_2 separately in (A), and by the combination $C_1 + C_2$ in (B). And we are thus led to the following first simplification in the computation of $p(\theta)$.

For each given θ, separate the aircraft's paths into coherent pieces [C_1 and C_2 separately in (A), $C_1 + C_2$ joined in (B)], and regard each piece simplified at its distant parts either by suppressing them or producing them to infinity as straight lines. Then compute the probability of detection for each part by the methods to be developed below. Lastly, combine the probabilities of the different parts as independent probabilities. Thus, in (A), if $p_1(\theta)$ and $p_2(\theta)$ are the probabilities of detection by C_1 and C_2, and if no other paths are crossed, we have

$$p(\theta) = 1 - [1 - p_1(\theta)]\,[1 - p_2(\theta)].$$

The second step in the approximate computation of $p(\theta)$ concerns the evaluation of the probability of detection $p'(\theta)$ for a single coherent part of track ($p' = p_1$ or p_2 in the previous example). Figure 5 shows the situation schematically. The coherent part of track is C (which may also be thought of as bent back, as in Figure 4B, where $C = C_1 + C_2$).

Since the aircraft flying C has a much greater speed than the submarine, the latter will move only very slightly during the time it takes the aircraft to traverse that part of C (the solid line) which is close enough to the submarine's path to give any appreciable chance of detection. Hence the approximation: *Regard the submarine as stationary* [e.g., at the point (x,y)] *while the aircraft makes a given flight of C.* The formula considered earlier then gives for the probability of detection during this particular flight

$$1 - \exp\left\{-\int \gamma\left[\sqrt{(X(t) - x)^2 + (Y(t) - y)^2}\right]dt\right\}$$

$$= 1 - \exp\left\{-\int_C \gamma\left[\sqrt{\left(X\left(\frac{s}{v}\right) - x\right)^2 + \left(Y\left(\frac{s}{v}\right) - y\right)^2}\right]\frac{ds}{v}\right\}$$

$$= 1 - \exp\left\{-\int_C f(r)\,ds\right\},$$

where v is the aircraft speed, s is the distance along its path C measured from any fixed reference

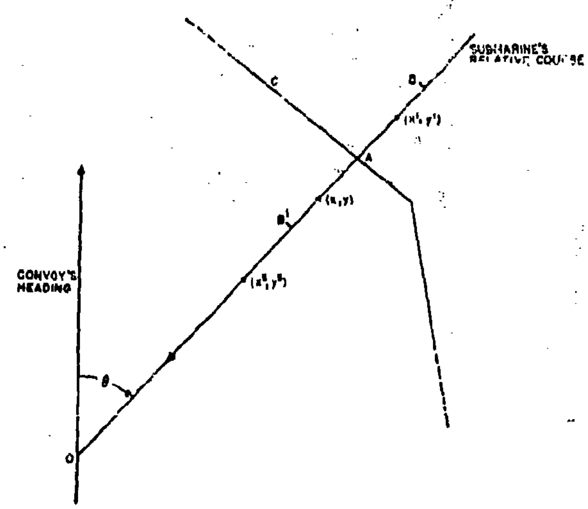


FIGURE 5. Successive positions seen during approach.

point, and $\int_C$ denotes integration over the whole of C. $\int_C f(r)ds$ is the contact potential (Section 2.3), r being the distance between the target at the fixed point (x,y) and the observer (aircraft) at the moving point (X,Y). A further approximation is involved here, in that the average ground speed v is taken as the aircraft speed in convoy space. Now C is flown regularly every T' hours, where $T' = T$ or an integral submultiple thereof (e.g., $T' = T/2$). Hence if at one epoch (time when the aircraft is at a fixed reference point on its path C such as A, Figure 5) the submarine is at (x,y), T' hours earlier it was at

(x',y'), and T' hours later it will be at (x'',y''), where

$$x' - x = x - x'' = T'w(\theta)\cos\theta,$$

$$y' - y = y - y'' = T'w(\theta)\sin\theta;$$

and there will be other earlier and later positions of the submarine corresponding with multiples of T'. The probability of detection in at least one of these positions is

$$1 - \exp\left\{-\sum_n \int_C f(r_n)\,ds\right\},$$

where

$$r_n = \sqrt{(X - x_n)^2 + (Y - y_n)^2},$$

(x_n,y_n) being the nth position of the submarine, e.g., $(x_n,y_n) = (x,y)\ (x',y')$, (x'',y''), etc., and $X = X(s/v)$, $Y = Y(s/v)$, the coordinates of the (moving) aircraft. As before, this must be averaged over a complete set of representative positions of (x,y), e.g., over all positions specified by

$$[x + \tau w(\theta)\cos\theta,\ y + \tau w(\theta)\sin\theta],\ 0 \leqq \tau < T'.$$

Thus

$$p'(\theta) = \frac{1}{T'}\int_0^{T'}\left\{1 - \exp\left[-\sum_n \int_C f(r_n)ds\right]\right\}d\tau.$$

The third step in the approximate evaluation of $p(\theta)$ aims at simplifying the summation Σ_n in this formula for $p'(\theta)$. To this end, select a distance D having the following property. At a lateral range greater than D, an aircraft on a straight course has a negligible chance of sighting the submarine, whereas at less than D it begins to have a chance of sighting which must be taken into account. Thus, if $D = 2E$ (E = effective visibility), the chance of sighting at lateral range D is of the order of 2 per cent; and for definiteness, this value of D shall be used herein. The decision to neglect points of the submarine's path farther from C than the distance D leads to the construction of the segment $B'B$ of the submarine's path, where $B'B$ is the locus of points not farther than D from C (Figure 5). It may happen that $B' = O$; but in all cases the notation is such that B' is nearer to O than is B. Figure 6 shows a scale attached to the submarine's path: B' has the coordinate zero, B has b ($b = B'B$), a typical position of the submarine is ξ, etc. And on this scale the points wT', $2wT'$, etc., have been marked [$w = w(\theta)$]. Figure 6 (A), (B), and (C) show three typical cases.

In case (A) there is at most one opportunity for an aircraft flying C every T' hours to detect the submarine. With the probability b/wT', the submarine will be on $B'B$ for some flight of C, and having moved at least wT' miles between flights, it will not be on $B'B$ at any other flights of C. Hence we have

$$p'(\theta) = \frac{b}{wT'}\,\frac{1}{b}\int_0^b\left\{1 - \exp\left[-\int_C f(r_\xi)ds\right]\right\}d\xi$$

$$= \frac{1}{wT'}\int_0^b\left\{1 - \exp\left[-\int_C f(r_\xi)ds\right]\right\}d\xi,$$

where

$$r_\xi = \sqrt{\left[x_\xi - X\left(\frac{s}{v}\right)\right]^2 - \left[y_\xi - Y\left(\frac{s}{v}\right)\right]^2}$$

is the distance between the aircraft at the point on its path corresponding with s, and the submarine's position (x_ξ,y_ξ) corresponding with ξ. Time averaging used previously has been replaced by the equivalent process of length averaging, and all terms but one in the summation Σ_n are absent.

In case (B), there will surely be at least one opportunity for the aircraft to sight the submarine, and surely not more than two such. The case of just one occurs when the submarine's position ξ is in the interval (b,wT'), an event of probability $(2wT' - b)/wT'$. If this occurs, the chance of detection is

$$\frac{1}{2wT' - b}\int_{b-wT'}^{wT'}\left\{1 - \exp\left[-\int_C f(r_\xi)ds\right]\right\}d\xi.$$

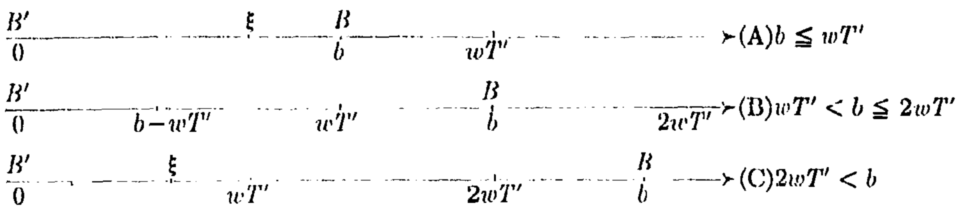


FIGURE 6.

The case of just two opportunities occurs when the (one and only one) ξ in $(0,T'w)$ falls in the interval $(0,b - wT')$, an event of probability $(b - wT')/wT'$. Then the second and last opportunity occurs when the submarine is at $\xi + wT'$, which will necessarily be in (wT',b). The chance of detection in this case is

$$\frac{1}{b - wT'}\int_0^{b-wT'}\left\{1 - \exp\left[-\int_C\left\{f(r_\xi) + f(r_{\xi+wT'})\right\} ds\right]\right\} d\xi.$$

When these expressions are multiplied by $(2wT' - b)/wT'$ and $(b - wT')/wT'$ respectively, and the products added (corresponding to total probability of two mutually exclusive events) the final expression for $p'(\theta)$ is obtained.

This assumes an absolute accuracy and regularity of flights and submarine motion lasting through the very appreciable time T', an assumption never realized in practice. It is more realistic, as well as mathematically simpler, to regard the two opportunities of detection as independent events and to combine them as such. If their individual probabilities are, respectively, p_1 and p_2, so that

$$p_1 = \frac{1}{b-wT'}\int_0^{b-wT'}\left\{1 - \exp\left[-\int_C f(r_\xi)\,ds\right]\right\}d\xi,$$

$$p_2 = \frac{1}{b-wT'}\int_0^{b-wT'}\left\{1 - \exp\left[-\int_C f(r_{\xi+wT'})ds\right]\right\}d\xi$$

$$= \frac{1}{b-wT'}\int_{wT'}^{b}\left\{1 - \exp\left[-\int_C f(r_\xi)ds\right]\right\}d\xi,$$

then their combination is $p_1 + p_2 - p_1p_2$.

In case (C), there are many positions of the submarine, differing successively by wT', at which the aircraft may be in sighting range. When C (Figure 5) is not close to 0, it is usually sufficiently accurate to regard the summation Σ_n as an infinite sum, i.e., to employ formulas for infinitely many congruent and equally distant sweeps (in target space). When C is close to O, this method must be modified to take account of the fact that the sweeps occur essentially on one side of the target (in target space), e.g., the probability p given by the sweep formula is replaced by $1 - \sqrt{1-p}$ to obtain $p'(\theta)$. Intermediate cases would offer greater complications in a direct treatment, but in practice it is usually sufficient to compute the two extreme values cited above and accept an intermediate value, such as their arithmetic mean.

It does not appear profitable to carry the discussion further in general terms, i.e., without assuming a definite form for γ or $f(r)$.

9.5 PLOTTING THE SCREENING POLAR DIAGRAM: INVERSE CUBE LAW

The practical problem of plotting the screening polar diagram when the escort plan and various relevant quantities are given is a task of approximate numerical computation of the probabilities $p(\theta)$ discussed theoretically in the last section. It can be carried out only when the law of detection is given, either as an equation, or a table or graph of the function $f(r)$. Chapters 4 and 5 have studied the law in detail in the cases of vision and radar and represent the best information existing at the time of publication. Their results are mainly in tabular or graphical form, and their application to the present problem involves a rather lengthy set of numerical computations. Partly to illustrate the principles of procedure in as simple a form as possible and partly to show how the computation was actually done in the case of many plans which have become fleet doctrine and had to be designed before the more detailed information of Chapters 4 and 5 was available, the present section will be based on the inverse cube law of detection (see Section 2.2). But it must be borne in mind that this is at best approximate and is not applicable in the case of any considerable amount of haze (low meteorological visibility). However, it is doubtful if when this law shows that a certain plan is better than another a more accurate law would reverse the comparison.

The inverse cube law finds its expression in the equations

$$f(r)ds = \frac{m ds}{r^3}, \quad m = 0.046E^2, \qquad (1)$$

where r is the range and E the effective visibility in nautical miles (see Chapter 2, equations (10) and (46), with dt replaced by $ds = dt/v$ in the former. From these, the process of Section 2.3 of Chapter 2 as illustrated by Figure 5 of that chapter leads to equation (23), as shown therein. It is expedient to rewrite these results in a slightly different form, giving Figure 7 below, which is drawn relative to the *target* (rather than the observer, as Figure 5 of Chapter 2). In other words, the target is fixed in

Figure 7, and the observer (aircraft) is flying with the relative velocity w. The angles ω', ω'' of the earlier treatment are replaced by γ', γ'' (not to be confused with the instantaneous sighting probability density). Thus we have

$$p = 1 - e^{-\int_{y'}^{y''} f(r)dy}, \tag{2}$$

$$m\int_{y'}^{y''} \frac{dy}{r^3} = m\int_{y'}^{y''} \frac{dy}{(x^2 + y^2)^{\frac{3}{2}}} = \frac{m}{x^2}\left(\frac{y''}{r''} - \frac{y'}{r'}\right). \tag{3}$$

It must be remembered that x in the above equations is the distance along the normal to the aircraft path from the target. Of equations (4), (5), and (6), only equation (6) is simple enough to be put through the averaging process which is necessary when the submarine position is not fixed. Thus, it will be found necessary to substitute infinite aircraft tracks for the finite ones found in the actual escort plan to be evaluated.

Often the aircraft will fly a path such as is indicated in Figure 8A and B. This path is made up of two semi-infinite straight lines joined at the turning point O. Using equation (5), since the contact potentials are additive (see Section 2.2), it is possible to derive a convenient form for the combined contact potential.

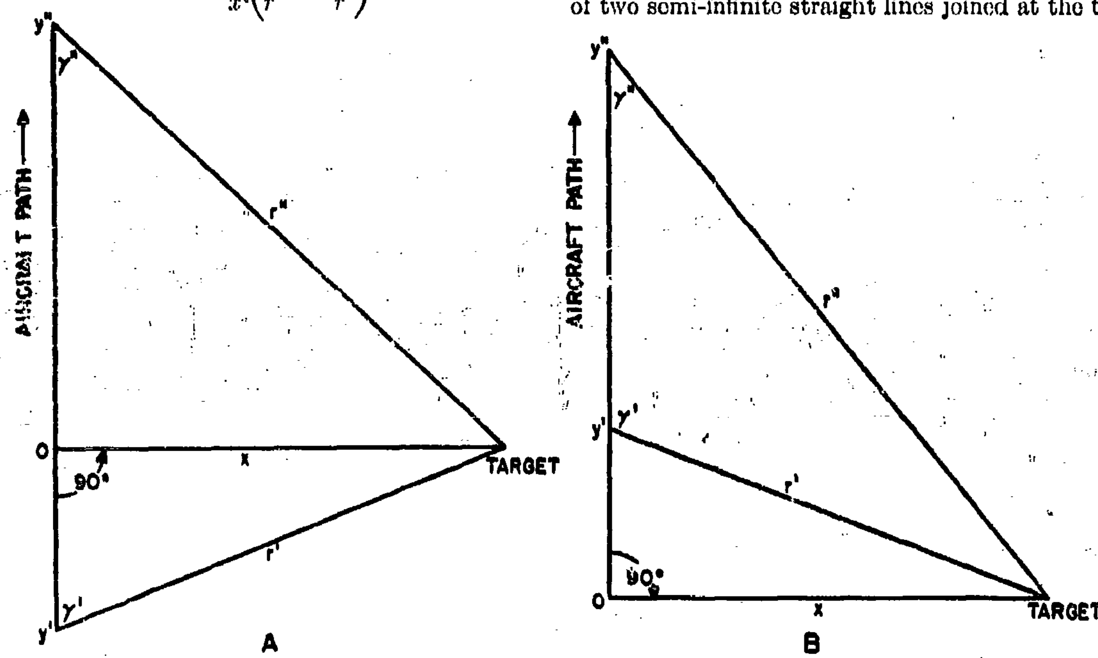


FIGURE 7. Detection at fixed speed and course, drawn relative to the target.

The quantities r' and r'' are the ranges from the respective limits of the aircraft track, y' and y'', to the target. These extreme range lines make the angles γ' and γ'' with the aircraft track. Thus, substituting in (3):

$$m\int_{y'}^{y''} \frac{dy}{r^3} = \frac{m}{x^2}(\cos\gamma'' - \cos\gamma'). \tag{4}$$

Should y' go to $-\infty$ or y'' to $+\infty$, the above equation, substituted in (2) above, becomes

$$p(x) = 1 - e^{-m/x^2(1+\cos\gamma)}, \tag{5}$$

where γ is the appropriate one of γ', γ''. Similarly, for a flight from $-\infty$ to $+\infty$,

$$p(x) = 1 - e^{-2m/x^2}. \tag{6}$$

$$\int_{A/C\text{ path}} f(r)dy = \frac{m}{(x\sin\gamma_1)^2}(1+\cos\gamma_1) + \frac{m}{(x\sin\gamma_2)^2}(1+\cos\gamma_2), \tag{7}$$

$$= \frac{m}{x^2}\frac{1}{2}\left(\csc^2\frac{\gamma_1}{2} + \csc^2\frac{\gamma_2}{2}\right).$$

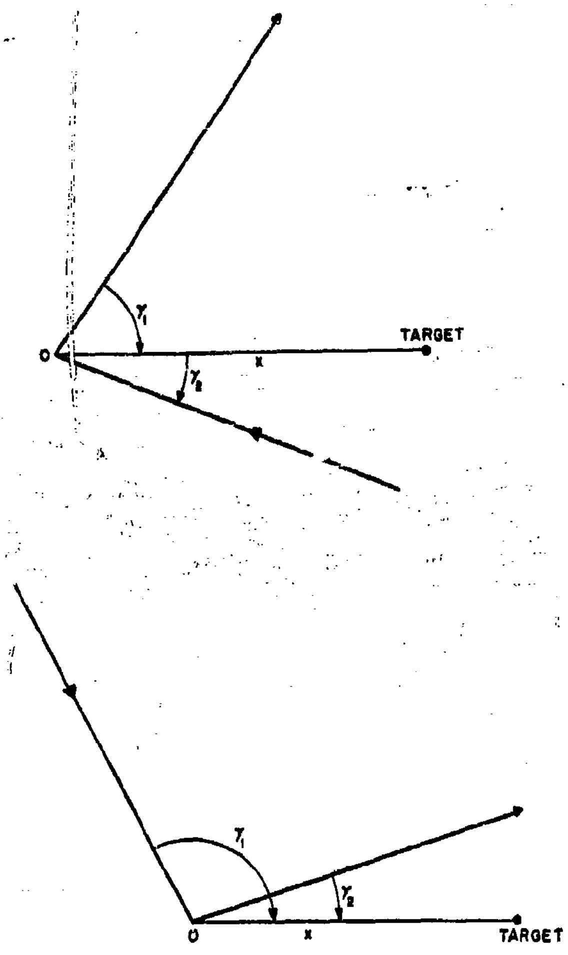
By setting $m' = 0.046(E')^2$

$$= 0.046\left(\frac{E}{2}\right)^2\left(\csc^2\frac{\gamma_1}{2} + \csc^2\frac{\gamma_2}{2}\right), \tag{8}$$

the form of equation (6) may be obtained:

$$p(x) = 1 - e^{-2m'/x^2}. \qquad (9)$$

Target known to be on a straight-line interval. If, instead of a stationary target, we have a moving one,



B

FIGURE 8. Submarine locus passing through angle of aircraft path.

the above formulas can still be applied to solve the problem. It is first necessary to consider cases where the submarine position at the times of the aircraft pass is known to be in a given straight-line interval and equally likely at any point therein. The case in which the aircraft flies along a straight infinite path which meets the submarine path at right angles at O is shown in Figure 9. The submarine is taken to be "uniformly distributed" in the interval from O to A (length a). Since the chance that the target be between x and $x + dx$ is dx/a, the probability of detecting the submarine is:

$$p(a) = \frac{1}{a}\int_0^a (1 - e^{-2m/x^2})\,dx. \qquad (10)$$

In the more general case where the angle α in Figure 9 may have any value, it is necessary to use as lateral range $x \sin \alpha$. The following equation is obtained.

$$p(a) = \frac{1}{a}\int_0^A (1 - e^{-2m/(x \sin \alpha)^2})\,dx. \qquad (11)$$

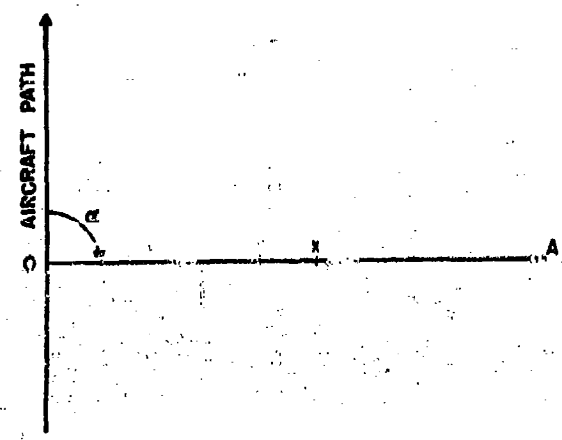


FIGURE 9. Straight aircraft path: averaging process. ($a = \overline{OA}$.)

Similarly, the case of the bent aircraft path can be expressed in this form. Figure 10 illustrates the situation. (The submarine may be at any place between O and A.) The expression for the probability involves the quantity m' defined in equation (8):

$$p(a) = \frac{1}{a}\int_0^a (1 - e^{-2m'/x^2})\,dx. \qquad (12)$$

Equations (10), (11), and (12) can all be integrated to yield

$$p(a) = g\left(\frac{E'}{a}\right) = 1 - e^{-X^2} + X\sqrt{\pi}\,(1 - \operatorname{erf} X), \qquad (13)$$

where

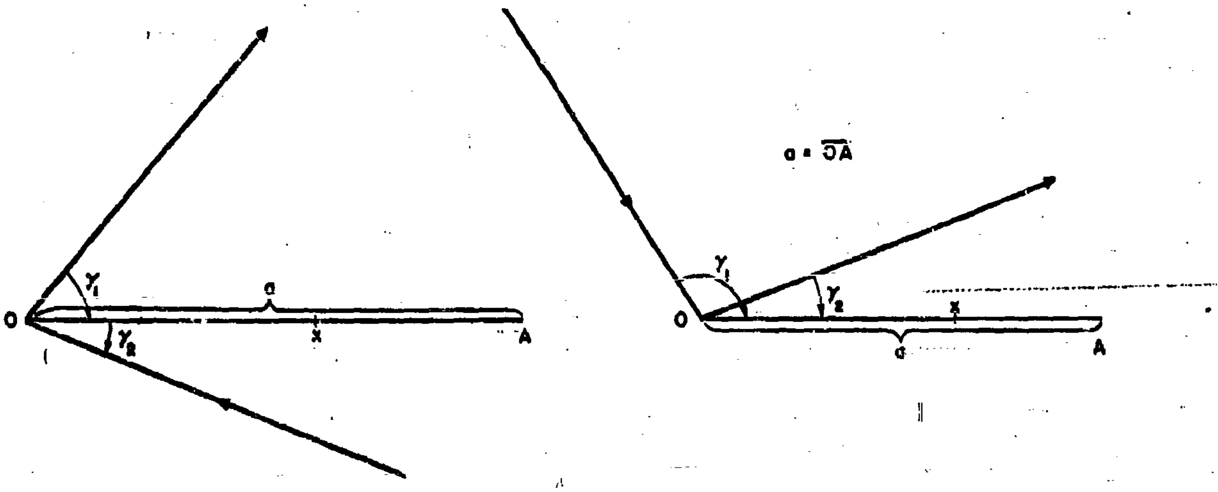
$$X = 0.303\,\frac{E'}{a}. \qquad (14)$$

For straight aircraft paths (Figure 9),

$$E' = E \csc \alpha; \qquad (15)$$

for bent aircraft paths (Figure 10),

$$E' = \frac{E}{2}\sqrt{\csc^2 \frac{\gamma_1}{2} + \csc^2 \frac{\gamma_2}{2}}. \qquad (16)$$



A

Figure 10. Aircraft on a bent path.

The principal curve on Plate I plots $g(E'/a)$ against E'/a (for this curve the abscissa E'/wT', is equal to E'/a).

The formulas heretofore derived can be applied to a more general case, namely, that in which the submarine is equally likely to be at any position in any straight line interval B to C; if the aircraft turns, its turning point must lie either on BC or on BC extended. The first case to be considered is that in which B and C are on the same side of the aircraft path, as shown in Figure 11.

For purposes of derivation, let us assume that there is a submarine in the interval O to B and that it is equally likely to be at any position in the interval. The probability of detecting that submarine is p_B, and can be calculated with the aid of equation (13). It can also be calculated as shown below in terms of p_C and p_{BC}, which are defined as follows:

p_C = the probability of detecting the submarine if it is given in the interval $\overline{OC}$.

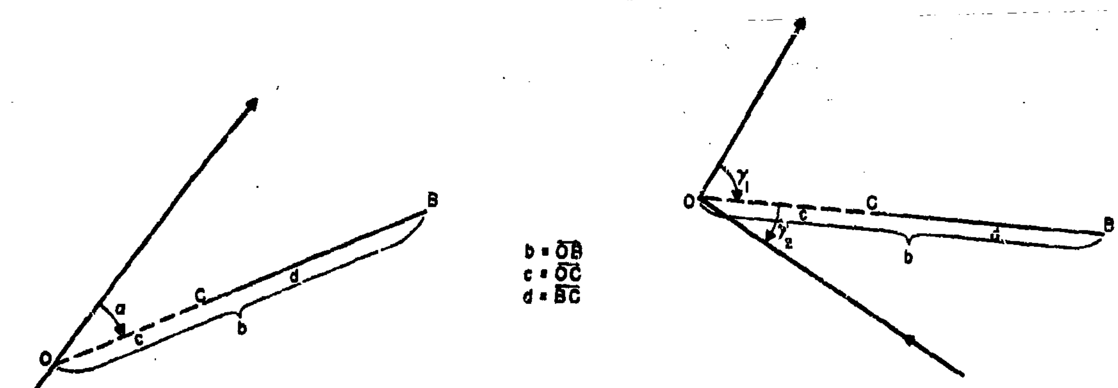
p_{BC} = the probability of detecting the submarine if it is given in the interval $\overline{BC}$.

[p_C, as well as p_B, can be calculated by equation (13).]

c/b = the probability that the submarine is in $\overline{OC}$.

d/b = the probability that the submarine is in $\overline{BC}$.

Thus, with the submarine anywhere between O and B, the product $c/b(p_C)$ is the chance of detecting the submarine between O and C, while $d/b(p_{BC})$ is



A STRAIGHT AIRCRAFT PATH B BENT AIRCRAFT PATH

Figure 11. General disposition of target with respect to aircraft path.

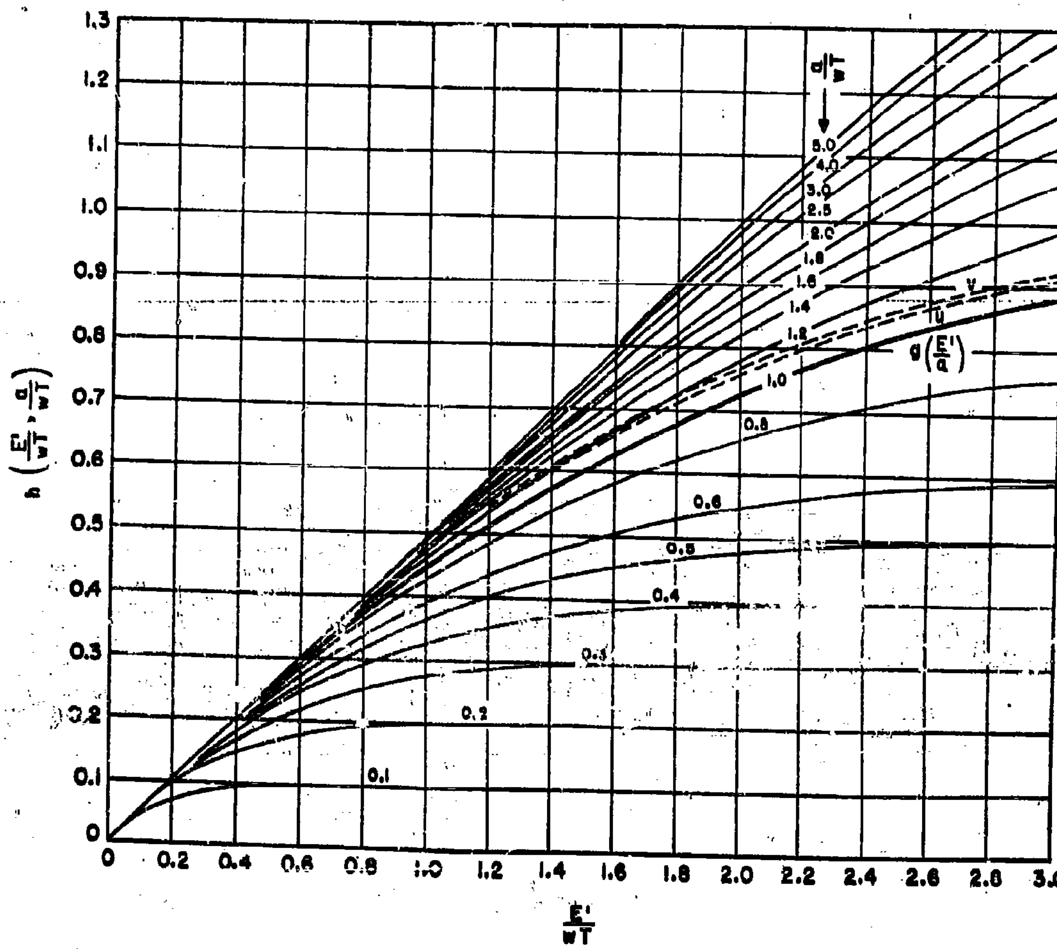


PLATE I. Curves for the computation of detection probabilities.

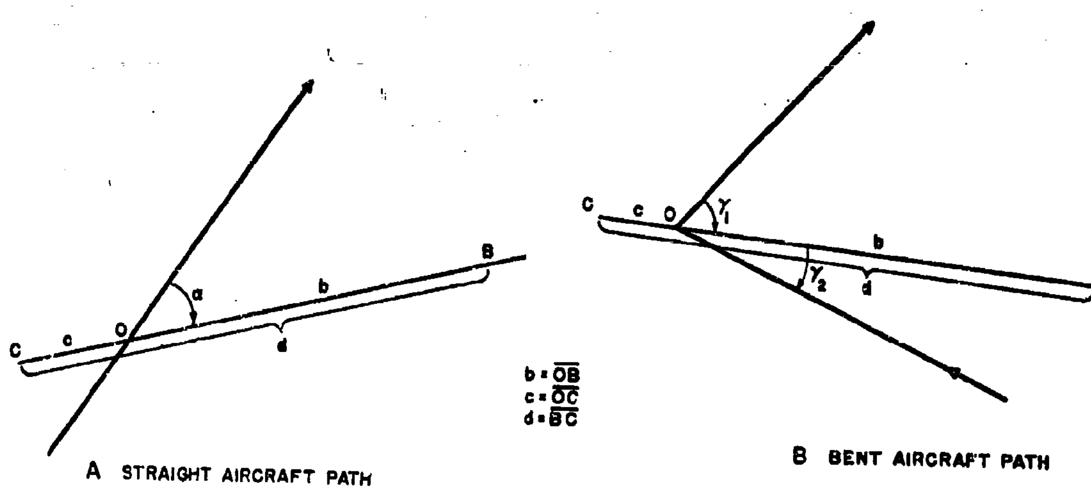


FIGURE 12. Case in which target can be on either side of aircraft path.

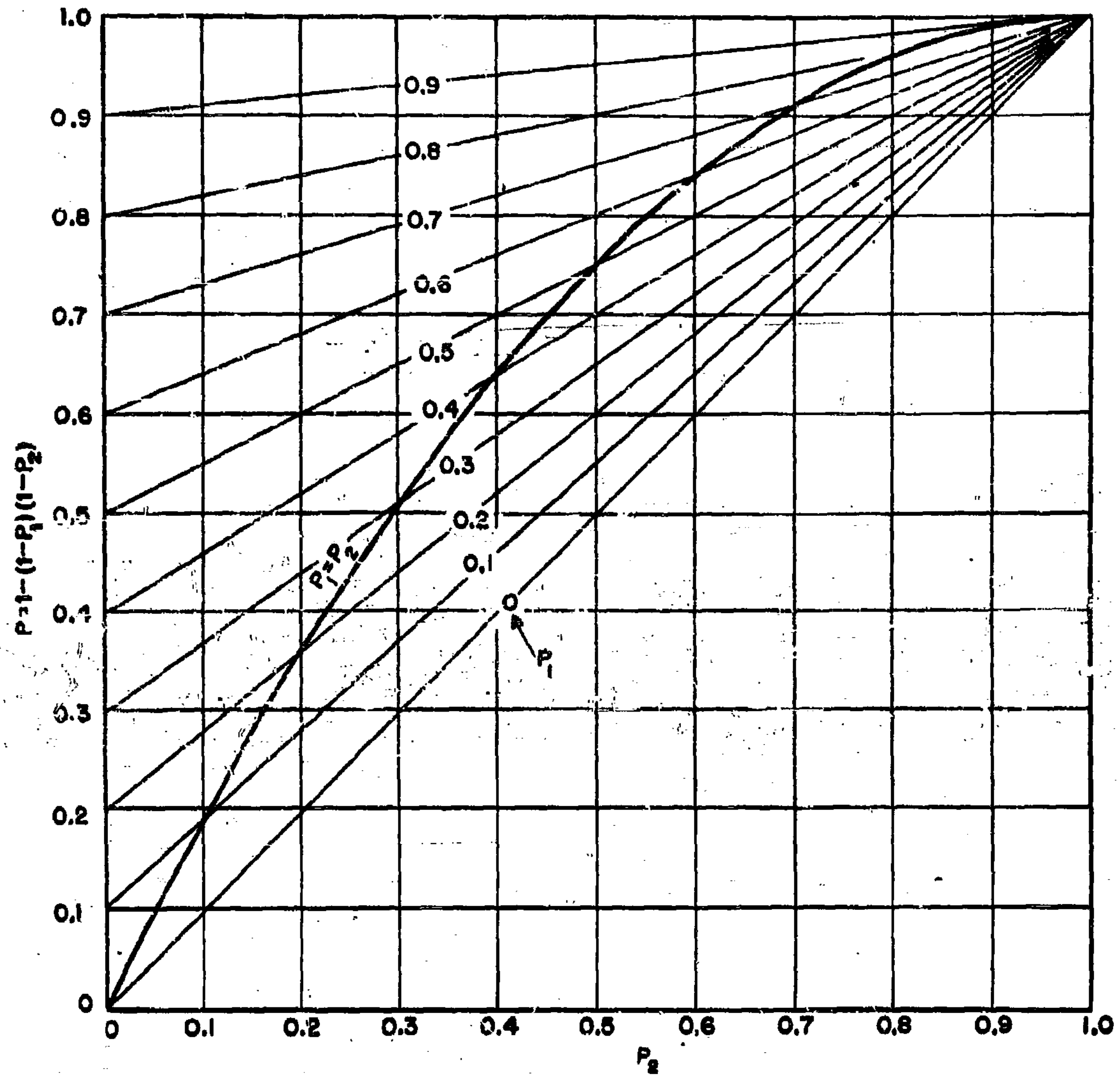


PLATE II. Curves for combining independent probabilities.

the chance of detecting it between B and C. Adding the two together gives

$$p_B = \frac{c}{b} p_C + \frac{d}{b} p_{BC}, \tag{17}$$

and, transposing, we obtain the desired probability,

$$p_{BC} = \frac{b}{d} p_B - \frac{c}{d} p_C. \tag{18}$$

The second case, illustrated in Figure 12, occurs when O falls inside the interval BC. Following the reasoning of the previous case, the obvious conclusion is:

$$p_{BC} = \frac{c}{d} p_C + \frac{b}{d} p_B, \tag{19}$$

where both p_C and p_B are calculated with the aid of equation (13). When the aircraft's path is not straight, as in Figure 12B, the correct angles must be used to calculate p_C and p_B. While γ_1 and γ_2 apply to p_B, *their supplements* are used in computing p_C. Thus for p_C equation (20) becomes:

$$E' = \frac{E}{2}\sqrt{\csc^2\frac{\pi-\gamma_1}{2} + \csc^2\frac{\pi-\gamma_2}{2}}$$

$$= \frac{E}{2}\sqrt{\sec^2\frac{\gamma_1}{2} + \sec^2\frac{\gamma_2}{2}}. \tag{20}$$

The aircraft, in some situations, will fly parallel or nearly parallel to the locus of possible submarine positions, as in Figure 13. Referring to equation (6) for an infinite aircraft flight past a stationary target, it is seen that the probability of detection at any

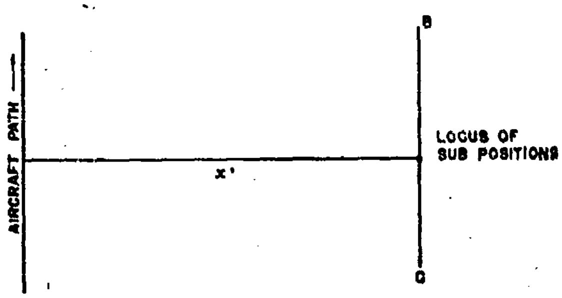


FIGURE 13. Case of parallel aircraft path.

one submarine position, and, therefore, along the entire locus, is given by the expression:

$$p_{\text{parallel}} = 1 - e^{-2m/x^2} \tag{21}$$
$$= 1 - e^{-0.092(E/x)^2}$$

A plot of equation (21) versus the quantity E/x is shown in Figure 14. If the submarine is assumed to be in the interval BC and its course is parallel to, or makes a very small angle with, the aircraft's course,

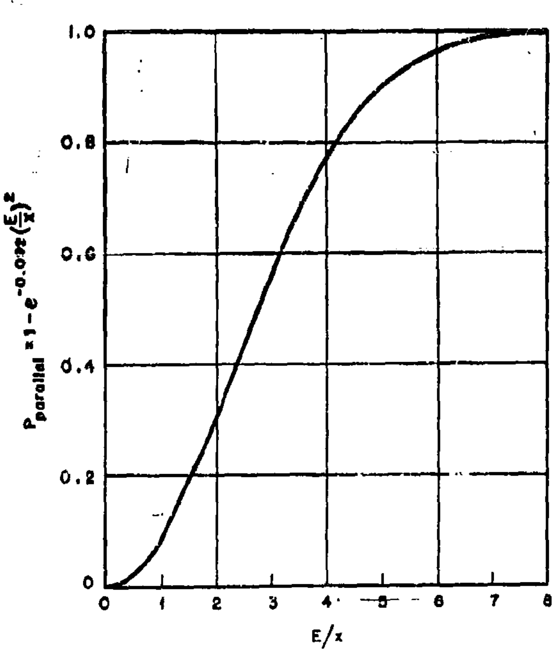


FIGURE 14. Probability of detection in parallel flight.

the probability of its detection on a single infinite flight is given by the ordinate corresponding to the appropriate value of E/x. For nearly parallel courses the average value of x may be used.

Target moving on a straight path. Let us reconsider the foregoing cases (Figures 11 and 12), assuming now that the plane repeats its "infinite" path every T hours (the period of the actual screening plan). It is desired to find the probability of contacting a submarine while it is moving from point B to point C with the velocity w. This w is small enough compared with the aircraft speed so that the submarine motion during a particular aircraft pass is negligible. Between passes the submarine covers the distance wT. The time at which the submarine reaches point B is completely at random. However, the situation repeats itself every T hours. A segment of length wT may be chosen arbitrarily on the submarine's (straight) path. The submarine will be somewhere in this wT section during one of the aircraft passes. During the preceding and succeeding passes the submarine will be in the adjacent wT-length segments. This set of nonoverlapping wT segments should be extended to include at least all of BC. Since the rigorous solution of the problem is independent of the position of the set of wT segments relative to BC, we shall in each case choose a set for which the errors due to various approximations will be minimized.

If d, the distance between B and C, is less than or equal to wT, the submarine is exposed to only one pass, for it will be outside BC at the time of any other pass. Choosing a wT section which encloses BC, the submarine is equally likely to be anywhere in this section at the time of the associated aircraft pass. d/wT is the chance that the submarine is between B and C and thus exposed to the detection. The earlier paragraphs have given the chance (p_{BC}) of contacting a submarine which is definitely in BC. Therefore, the desired probability (p) is $(d/wT)p_{BC}$. Then when $d < wT$, equations (18) and (19) become, respectively:

$$p = \frac{b}{wT}p_B - \frac{c}{wT}p_C = h\left(\frac{E'}{wT}, \frac{b}{wT}\right) - h\left(\frac{E'}{wT}, \frac{c}{wT}\right), \tag{22}$$

$$p = \frac{b}{wT}p_B + \frac{c}{wT}p_C = h\left(\frac{E'}{wT}, \frac{b}{wT}\right) + h\left(\frac{E'}{wT}, \frac{c}{wT}\right), \tag{23}$$

$$p = \frac{d}{wT}p_{\text{parallel}}, \tag{24}$$

where, in each case,

$$h\left(\frac{E'}{wT}, \frac{a}{wT}\right) = \frac{a}{wT}\,g\left(\frac{E'}{a}\right) = \frac{a}{wT}\,p_A. \qquad (25)$$

In equation (25) above, a is the distance along the submarine path from the aircraft track to the extreme positions of the submarine (b or c, as the case may be). In equation (22), a (i.e., b or c) may be greater than wT and the function h may be greater than unity. Nevertheless, as long as d is less than wT, subtracting these h's will give the correct probability. Working from the $g(E'/a)$ curve, $h[(E'/wT), (a/wT)]$ was calculated for various a/wT and plotted against E'/wT on Plate I. Obtaining p is just a matter of reading, adding or subtracting values from the curves.

If d is greater than wT, a given submarine may be exposed to more than one aircraft pass. In the simplest such case the submarine path extends exactly to the point where it intersects the aircraft track and $d = 2wT$. For the rigorous treatment of this case, a submarine which is at distance $x < wT$ from O on one pass must have been exposed at the position $x + wT$ on the previous pass. Equation (12) then becomes,

$$p_{2wT} = \frac{1}{wT}\int_0^{wT}\left\{1 - e^{-2m'\left[\frac{1}{x^2} + \frac{1}{(x + wT)^2}\right]}\right\}dx. \qquad (26)$$

This is plotted as curve u on Plate I.

There are no general formulas giving rigorous solutions when $d > wT$. Each situation would have to be treated separately and usually with more difficult integration than in equation (26).

The following approximate procedure has been found sufficiently accurate. A submarine's position in one wT segment is assumed to be independent of its position in the adjacent one at the time of the previous aircraft pass. The submarine path BC can then be broken into sections s_i of length wT or less; the probability of contact can be found for each section using equations (22) or (23). Finally, these probabilities, say $p_1, p_2 \cdots p_n$, can be combined as independent events, to give p_{BC}. This may be expressed by:

$$p = 1 - (1 - p_1)(1 - p_2)(1 - p_3) \cdots (1 - p_n) \qquad (27)$$

Plate II gives a rapid solution of equation (27) for two p's; for more sections Plate II may be used repetitively.

The series of wT sections is placed with respect to BC in such a way that one of the wT sections has as high a chance of contact as possible. Then the contribution of the other sections is minimized and so is the effect of the independence assumption. For Figure 11 the principal section, S_1, should be taken from C to $C + wT$. In the special case of Figures 9 or 10, S_1 would be from O to wT. In Figure 12A, if b and c are greater than $1/2wT$, S_1 should extend from $-1/2wT$ to $1/2wT$. If $c < 1/2wT$, S_1 should extend from $-c$ to $wT - c$. If as in Figure 12B the aircraft turns at O, the larger part of S_1 should be on the side of O having the best chance of sighting, that is, the side associated with the larger E'; usually S_1 was taken from $0.2wT$ on the weak side to $0.8wT$ on the strong side.

To find the order of magnitude of the error involved, consider the case treated rigorously in equation (26) in which the submarine may be sighted between $x = 0$ and $x = 2wT$. Curves u and v of Plate I give, respectively, the rigorous and the approximate solutions. These differ by only about one per cent, which is a smaller error than those to be discussed in the following paragraphs. The possibility of fluctuations in the values of w and T introduces doubts with regard to the rigorous method and makes the approximate method even more excusable.

Now that we can handle any interval of submarine track, let us begin to consider the effectiveness of an actual screening plan against a submarine on a specific collision course. Unless detected, the submarine proceeds from well beyond detection range to an end point E at which it either submerges or fires at the convoy. Also relative to the convoy, the aircraft will repeat every T hours a closed path composed of straight legs. None of these legs goes to infinity. Nevertheless, the *only* practical way of treating a screening plan is to divide the aircraft path into sections each of which can be treated as an approximation to one of the cases already discussed. In each of these cases the aircraft path was an infinite straight line coming up to the submarine track and the same or another line going off to infinity.

Consider the illustrative aircraft path $FGHIJF$ and the submarine path in to E as shown in Figure 15. In this case four hypothetical aircraft paths must be considered; three of them are made by extending the legs FG, GH, and JF to infinity in both directions; on the fourth path the aircraft approaches from a great distance on a straight path through H to I where it turns and goes out to infinity on a straight path through J. Each infinite path is to be treated by one of the methods described earlier to give the probability of contact if the plane were searching only during the corresponding segment, FG, GH, HIJ, or JF, of the screening plan. (These individual probabilities will later be combined to give the effectiveness of the entire plan.) However, if with each infinite path the *entire* submarine track in to E were considered, the individual probabilities found would be too high, as a result of the finite chance of contact while the aircraft is on the imaginary extensions. This error is to be eliminated roughly by limiting in each case the section of the submarine track in which a contact can be made. Thus an infinite aircraft track and a limited submarine track shall be substituted for a limited segment of aircraft track and an infinite submarine track (see Figure 16). This substitution is required because only infinite aircraft tracks can be handled at all simply.

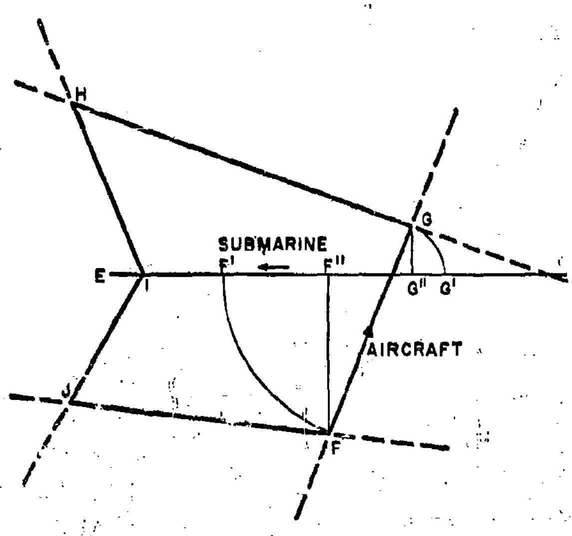


FIGURE 15. Example of path and track approximations.

The following scheme is employed in limiting the submarine track. The submarine is never allowed beyond the limits set up in the original problem, i.e., to the left of E in Figure 15.

Let b' and c' be the distances from O (the intersection of the aircraft and submarine paths or their extensions) to the limits of the segment of aircraft track. Whenever possible b and c, the distances from O to the extremes of the submarine path, are taken to equal b' and c', respectively, and are chosen so that the angles between b and b' and between c and c' are each acute. For example, in Figure 15 the submarine intervals $F'G'$, EG'', and EF'' are used for legs FG, GH, and JF, respectively. The principal exception occurs when, as with HIJ, the side of the submarine track makes an obtuse angle with both legs of the aircraft track; then that side of the sub-

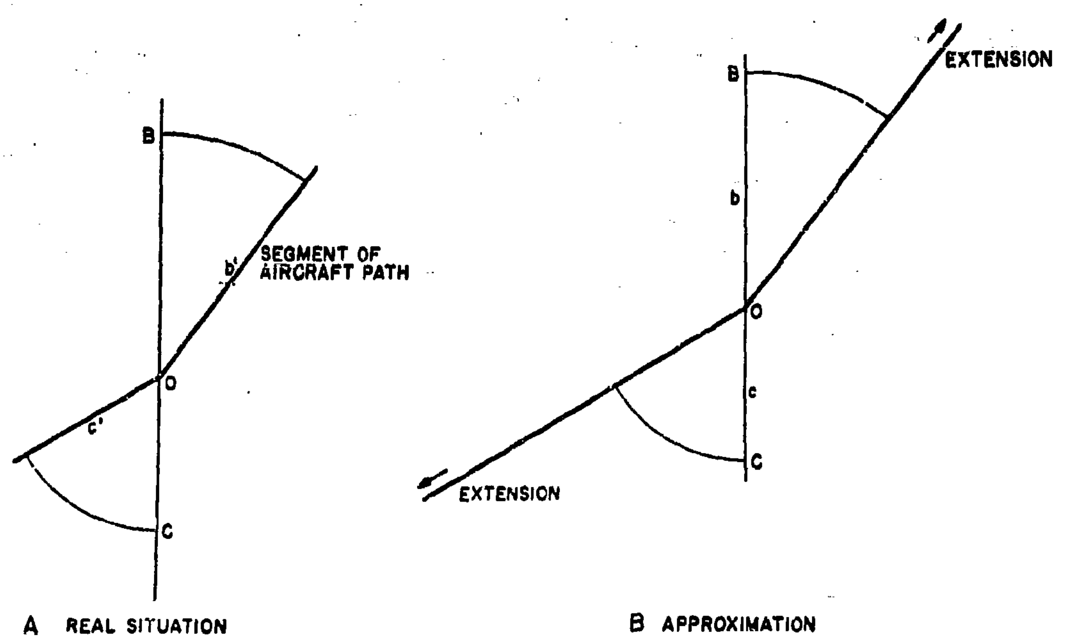


FIGURE 16. The geometrical principles of path and track approximation.

marine track is not limited, while the other side must be limited in length to a weighted average of the lengths of the aircraft legs. In Figure 15 this questionable averaging is avoided and the entire submarine path (to the right of E) can be used, since IE is less than either IH or IJ.

The last paragraph may be justified as follows: In the real situation with the aircraft on a limited segment of track, most of the chance of contact accumulates while the submarine is on the part of its track to which it is limited in the approximation, because in most cases this is the part of the infinite submarine track closest to the segment of aircraft track. This large portion of the desired probability is correctly obtained by the approximate method, for it is the part contributed by the portion of the infinite aircraft track which lies within the real segment. The remaining so-called "end effects" are in the real case the chance of the aircraft on its segment contacting a submarine outside BC (in Figure 16A) and in the approximation the chance of the aircraft on the infinite extensions of its segment contacting a submarine in BC (Figure 16B). The symmetry about O between the real and the approximate cases makes the end effects about equal, since all the submarine to aircraft distances are duplicated. A systematic error (a pessimistic one) enters only when the aircraft segment is much shorter than wT. The error due to terminating the submarine track at a point E within BC enters only into the end effects and is usually small. In any case, these errors are small compared with those which could arise from the inaccuracy of the inverse cube law.

The chances of contact on the individual segments of aircraft track are combined as *independent* probabilities to yield the desired probability of detection by the entire plan. Plate II may be used in making this computation. The error resulting from the independence assumption is seldom over four per cent. Usually one of the individual probabilities predominates and so the method of including the others is unimportant.

In initially choosing the submarine collision courses (i.e., the θ for which $p(\theta)$ is computed), it is advisable to take one through each corner of the aircraft track and then avoid others close to these corners. Those through the corners can be handled well by the bent track formulas given above. On the other hand, with neighboring tracks, the two legs (such as FG and GH at G in Figure 15) have to be extended to infinity. The farther the intersection is from the submarine track, the less important will be the end effects due to these extensions. Also the independent combination of the probabilities on these two legs is a dubious procedure since both legs have their best chance of contact against the same part of the submarine track at the same time. The farther G is from the submarine track the better.

9.6 PRACTICAL INSTRUCTIONS FOR OBTAINING A POLAR DIAGRAM[g]

The foregoing will now be summarized in a set of practical rules for obtaining a screening polar diagram.[g] Illustrating these rules, Plate III contains the calculations leading to the polar diagram (Figure 18) for the air escort plan of Figure 17.

1. Starting with the aircraft track drawn relative to the convoy as shown in Figure 17, choose the submarine paths to be studied in evaluating the plan.
   a). Each submarine path starts well outside sighting range and comes in on a collision course until it reaches the danger zone (submerged or firing area).
   b). Each path is associated with the angle θ which it makes with the convoy's heading.
   c). There should be a sufficient number of paths so that no important variations in the polar diagram will be missed.
   d). The submarine paths through each corner of the aircraft path should be included.
   e). Other paths close to corners should be avoided.
   f). For escort plans symmetrical about the convoy's line of advance, only θ's up to 180 degrees need be considered.
2. Make out a table with column headings 1 to 20, as in Plate III.
3. Record the θ's and corresponding values of w, the submarine's speed relative to the convoy.
4. Using T, the time (hours) of repetition of the escort plan, calculate and record the values of wT.
   a). In the absence of a true space diagram, it is sufficiently accurate to calculate T by dividing the length of the relative aircraft track by its true speed.
5. For each submarine path divide the aircraft

[g] In the method given here, the probabilities $p(\theta)$ are computed for paths ending not at the convoy reference point O, as in the general discussion of Section 9.4, but at the boundary of the submergence region (here, the danger zone).

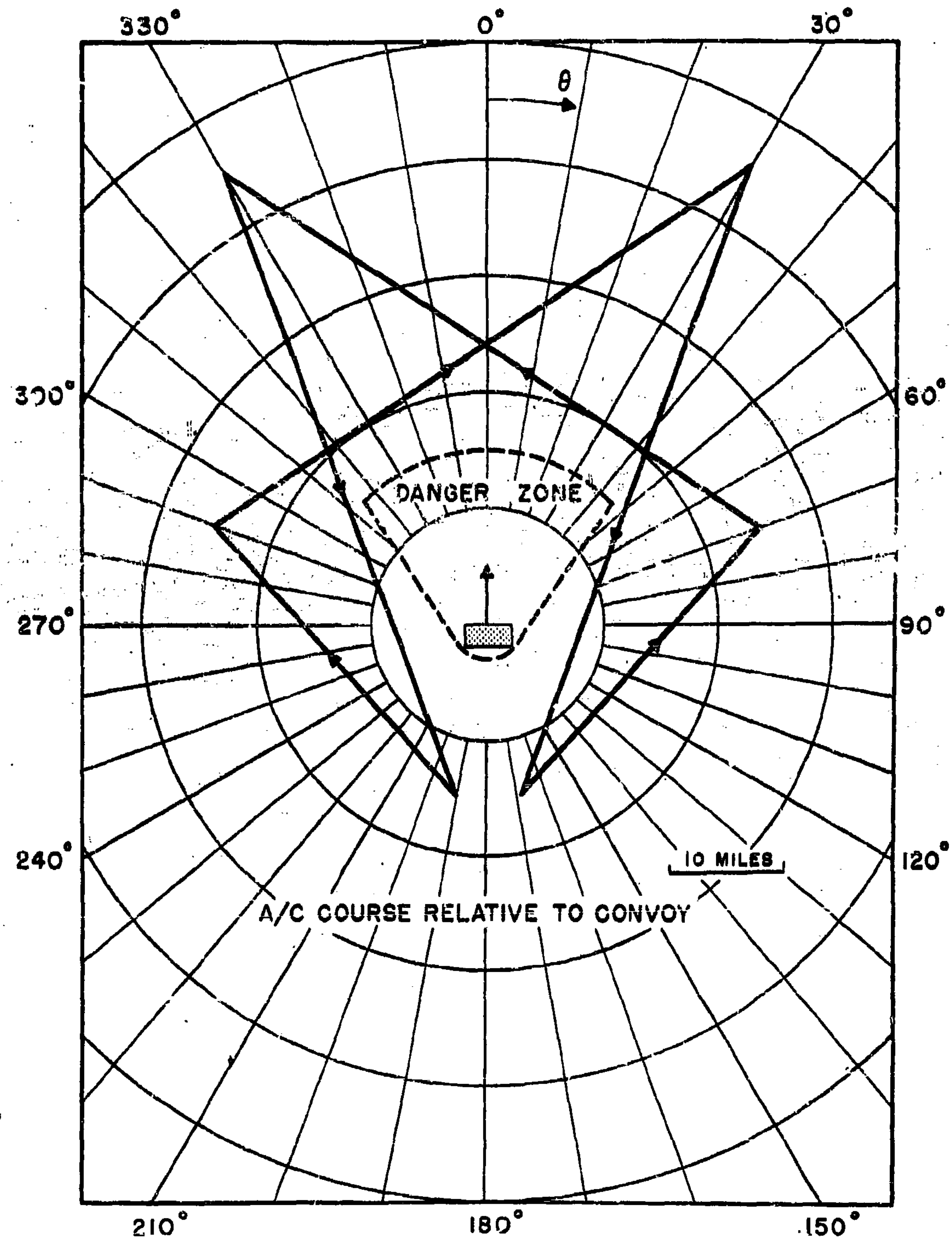


FIGURE 17. A possible aircraft escort plan.

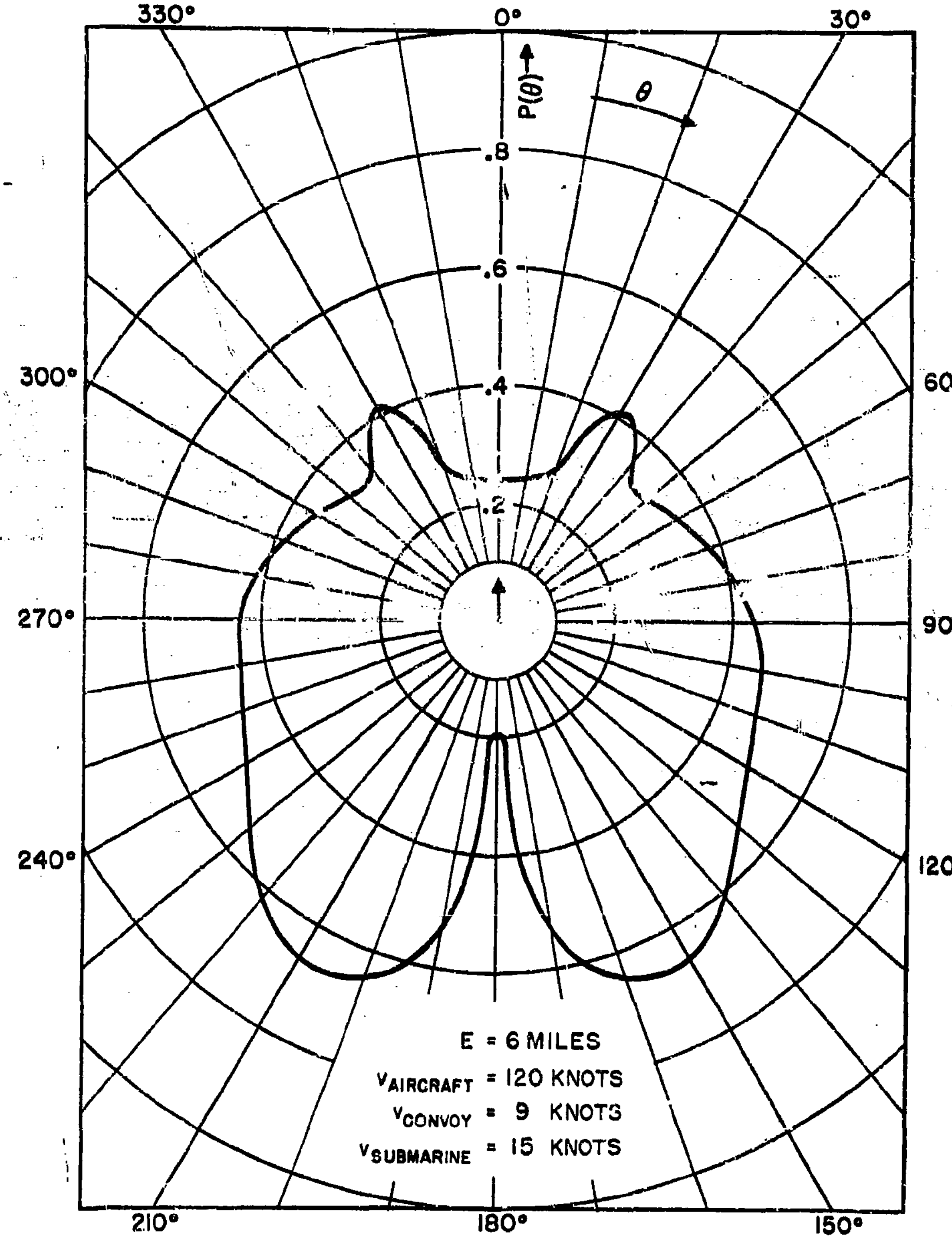


FIGURE 18. Screening polar diagram.

track into segments each of which is either a single straight leg or two legs joined by a turn on the submarine's path.

a). A row of the table should be devoted to each of these segments.

b). Usually there are segments located so far from the submarine path that their contribution is negligible and they can be eliminated by inspection.

6. Columns 4 to 8 deal with pertinent data measured from the relative diagram and are best filled out row by row (see below).

7. Record in column 7, for each straight segment, the angle α between the segment and the submarine path or their extensions.

8. Record in column 8, for each bent segment, the angles γ_1 and γ_2 between the submarine path and each leg of the segment.

a). Both γ_1 and γ_2 are measured from the same side of the submarine track from the intersection.

9. For each segment of aircraft tr[illegible] the submarine path is to be limited, in order t[illegible]ate for substituting an infinite aircraft track [illegible] the actual segment.

a). The distances b and c (as used in columns 4 and 5) are measured from O, the intersection of the aircraft and submarine paths or their extensions, to the extremes, B and C, of the limited submarine path.

b). With the exception noted below, b and c are equal to the distances from O to the extremes of the aircraft segment; B and C are each found by rotating the segment about O through an acute angle to coincide with the submarine path. *Exception.* When the aircraft turns so that the γ's are either both acute or both obtuse, the part of the submarine path making obtuse angles with both legs is not limited. The other part is limited to an average of the two legs (weighted by the cosecants of the γ's).

10. In columns 4 and 5 record b/wT and c/wT.

a). If BC is entirely on one side of O, the smaller of b/wT and c/wT is recorded in column 5 and given a negative sign, by convention.

11. Whenever the aircraft segment and submarine path are approximately parallel, B and C are found as in 9 above, where in this case the rotation is about a point O at a great distance.

a). The length BC of the limited submarine path is called d and divided by wT for listing in column 6.

b). Also listed in column 6 is x, the average distance between BC and the aircraft segment.

12. Record in columns 9 and 10 the required trigonometric functions.

13. Record the following in columns 11 and 19.

a). For straight segments,

$$E'/E = \csc \alpha$$

[in accordance with equation (15)].

b). For bent segments,

$$\frac{E'}{E} = \tfrac{1}{2}\sqrt{\csc^2 \gamma_1/2 + \csc^2 \gamma_2/2}$$

and

$$\frac{E'}{E} = \tfrac{1}{2}\sqrt{\sec^2 \gamma_1/2 + \sec^2 \gamma_2/2},$$

corresponding to the two parts of the submarine path separated by the intersection point O. The former applies to the part of the path from which the γ's are measured. [Equations (16) and (20)].

c). Only one of the equations in b) above is used when the limited submarine path falls entirely on one side of O.

14. Columns 13 to 20 involve the value of E, the effective visibility, and must be recalculated for each such value.

15. For flights parallel to the submarine path,

a). Record in column 13 the values of E/x, using the entry that has been made in column 6.

b). Obtain p_{parallel} from Figure 14 by using E/x.

c). Whenever $d/wT \leq 1$, obtain p_{BC} (column 19) by multiplying p_{parallel} by d/wT. (p_{BC} is the probability associated with this segment of the aircraft track against the submarine path.)

d). Whenever $d/wT = n + f$, where n is an integer and $0 \leq f < 1$, multiply p_{parallel} by f to obtain p_f. Using equation (27) or Plate II, obtain p_{BC} by combining p_{parallel} with itself n times and with p_f all as independent probabilities.

16. Record the values of E'/wT in columns 14 and 15. These are obtained by multiplying E/wT by the entries, E'/E, in columns 11 and 12.

PLATE III. CALCULATION OF THE SCREEN[illegible] [illegible]OLAR DIAGRAM FOR PLAN K.

v = 120 knots v_c = 9 knots v_{sub} = 15 knots $T = \frac{286}{120}$ = 2.38 hours Assuming E = 6 miles

1	2	3	4	5	6		7	8		9		10		11	12	13	14	15	16	17	18	19	20
θ (degrees)	x (knots)	xT (miles)	b	c/xT	d/xT (For parallel paths)	x miles	α (degrees)	γ_1 (degrees)	γ_2	$\csc\frac{\gamma_1}{2}$	$\csc\frac{\gamma_2}{2}$	$\sec\frac{\gamma_1}{2}$	$\sec\frac{\gamma_2}{2}$	$E'/E = \frac{1}{2}\sqrt{\csc^2\gamma_1/2+\csc^2\gamma_2/2}$	$E'/E = \frac{1}{2}\sqrt{\sec^2\gamma_1/2+\sec^2\gamma_2/2}$	E/x	E'/xT	E'/xT	h_b (For principle xT)	h_c	Additional contributions from other xT's	p_{BC}	$p(\theta)$
0	24	57.1	0.173 .14	0.175 −0.65			57 20							1.1[illegible] 2.92			0.125 0.306		0.062 0.155	0.062 −0.150		0.124 0.005 double	0.245
10	23.8	56.6	0.[illegible]7 0.5[illegible]	0.221 0.12	0.53	22	47 67							1.[illegible] 1.[illegible]		0.273	0.165 0.115		0.72 0.057	0.72 0.050		0.144 0.107 0.015	0.[illegible]
20	23.2	56.2	0.236 0.643	0.335 0.10	0.54	8	37 77							1.66 1.027		0.75	0.180 0.111		0.090 0.055	0.090 0.050		0.180 0.[illegible]95 0.04	0.27
30	22.1	52.[illegible]	0.572 0.[illegible]	∞ 0.095			87	10	26	11.47	4.45	1.003	1.027	6.15 1.001	[illegible].72		5.704 0.114	0.082	0.300 0.057	0.042 0.05		0.342 0.107	0.41
45	19.7	46.9	[illegible]50 0.26	0.065 0.11	0.6 0.13	14.5 11	25 78							2.36 1.021		0.414 0.545	0.302 0.131		0.142 0.065	0.06 0.05		0.202 0.125 0.024 0.006	0.32
55	18.2	43.3	0.76 [illegible].16	0.11 0.29	0.35 0.7	8 19	35 69							1.74 1.07		0.75 0.315	0.241 0.146		0.115 0.06	0.030 0.074		0.195 0.134 0.025 0.071	0.34
70	15.[illegible]	[illegible]	0.53 1.[illegible]	∞ 0.12			50	23	53	4.13	2.24	1.030	1.118	2.34 1.304	[illegible]		0.362 0.212	0.124	0.175 0.106	0.062 0.080	0	0.237 0.168	0.37
90	12.0	25.[illegible]	[illegible] 1.4 1.14	0.42 0.14 −0.54			48 70 32							1.347 1.062 1.888			0.282 0.223 0.396		0.135 0.111 0.198	0.135 0.002 −0.180	0	0.270 0.203 0.018	0.44
120	8.3	19.8	1. 0.58	0.45 0.24			78 80							1.021 1.017			0.31 0.308		0.145 0.145	0.145 0.127	0.010	0.293 0.272	0.4[illegible]
150	6.6	15.7	0.33 0.43 3.5	0.61 0.45 −3.0	0.54	9	72 50 11							1.05 1.304 −5.24		0.667	0.402 0.51 0.290		0.175 0.22	0.200 0.22		0.375 0.44 0.035 0.000	0.66
170	6.1	14.5	0.84 2.55 1.35	∞ −1.8 −0.65			10 32	30	52	3.86	2.28	1.033	1.111	2.29 5.75 1.89	0.76		0.95 2.38 0.79	0.10	0.415 1.096 0.375	0.05 −0.995 −0.355	0	0.455 0.101 0.040	0.54
180	6.0	14.3	1.05 1.43	−0.28 −0.63			42 29							1.49 2.92			0.63 1.22		0.303 0.562	−0.213 −0.45		0.090 0.112	0.19

17. When the sum of the values in columns 4 and 5 is less than or equal to 1,

a). These values together with those in columns 14 and 15 may be used together with Plate I to yield h_b and h_c (columns 16 and 17).

b). Giving column 17 the same sign as column 5, add algebraically h_b and h_c to obtain p_{BC} (column 19).

18. When the sum of the values in columns 4 and 5 is greater than 1,

a). Choose the section S (length wT) of submarine path BC, which has the greatest chance of contact. Within the limits imposed by columns 4 and 5, the center of S should be located as follows:

(1) For straight aircraft segments the center should be as close as possible to the intersection point O.

(2) For bent aircraft segments, the center should be as close as possible to a point $0.3wT$ from O on the side having the greater E'/E.

b). Using the a/wT's corresponding to the extremes of S and using the E'/wT's from columns 14 and 15, find the corresponding h's from Plate I, and enter them in columns 16 and 17.

c). For each of the remaining sections (length wT or less) of BC, find the chances of contact (by a subtraction of two points having the same abscissa on Plate I) and enter these separately in column 18.

d). Obtain p_{BC} (column 19) by combining the algebraic sum of columns 16 and 17 with these smaller probabilities (column 18) by the independence procedure [equation (27) or Plate II].

19. Column 19 now has the contribution of each aircraft segment against each submarine path. For each submarine path combine these p_{BC}'s as independent probabilities [using equation (27) or Plate II] to give $p(\theta)$, the chance of contacting the submarine on collision course θ when the whole aircraft plan is considered.

20. Plot $p(\theta)$ against θ to give the desired polar diagram (Figure 18).

9.7 THE DESIGN OF A PLAN

As in all questions of this nature, there are two possible viewpoints. Either we may lay down the tactical results to be achieved (scouting and screening) and then find a plan which achieves them with the least expenditure of effort (number of aircraft and flying time), or else we may fix the total amount of effort available and seek the plan which maximizes the tactical results. Now the solution of the second problem furnishes that of the first, since it automatically informs us of what can be achieved with one aircraft, with two aircraft, etc.; and it remains only to pick the first plan in this sequence which gives the required result. It shall accordingly be from this viewpoint that the problem is approached here.

Let the number n of aircraft be given together with their capabilities of speed v, endurance, and detection. The convoy's speed v_0 (in the sense of mean course made good) is also supposed to be known. Before there can be any question of the "best" plan, a decision must be made as to how much relative importance must be attached to *scouting* as compared with *screening*. Now this is a purely tactical question. It must be settled on the basis of presumed enemy tactics and submarine capacities, as well as of our own defensive capabilities and our vulnerability. For example, against a submarine capable of high submerged speed and endurance, screening could be expected to be less effective than scouting, whereas the reverse might be true if the submarine were of the older type without these capabilities. When this decision has been reached and the relevant speeds and distances have been estimated, five conditions must be satisfied.

1. The entrance to the submerged approach region must be adequately guarded.

2. Flying must not be unduly concentrated about the convoy.

3. The circuits must be closed in convoy space, i.e., the aircraft must automatically meet the convoy.

4. The time between successive meetings of the convoy must not be excessive (never more than two hours; one hour is much better than two).

5. The plan must be navigable with reasonable ease. This means that one involving many turns must be avoided.

Now obviously it is not feasible to deduce an exact plan from these data and the requirement of maximizing scouting and screening. It is necessary *first* to invent plans and *afterwards* to test, modify, and select until a satisfactory one is obtained, and to exhaust all visible possibilities of improvement. Such a procedure is an art quite as much as a science. It

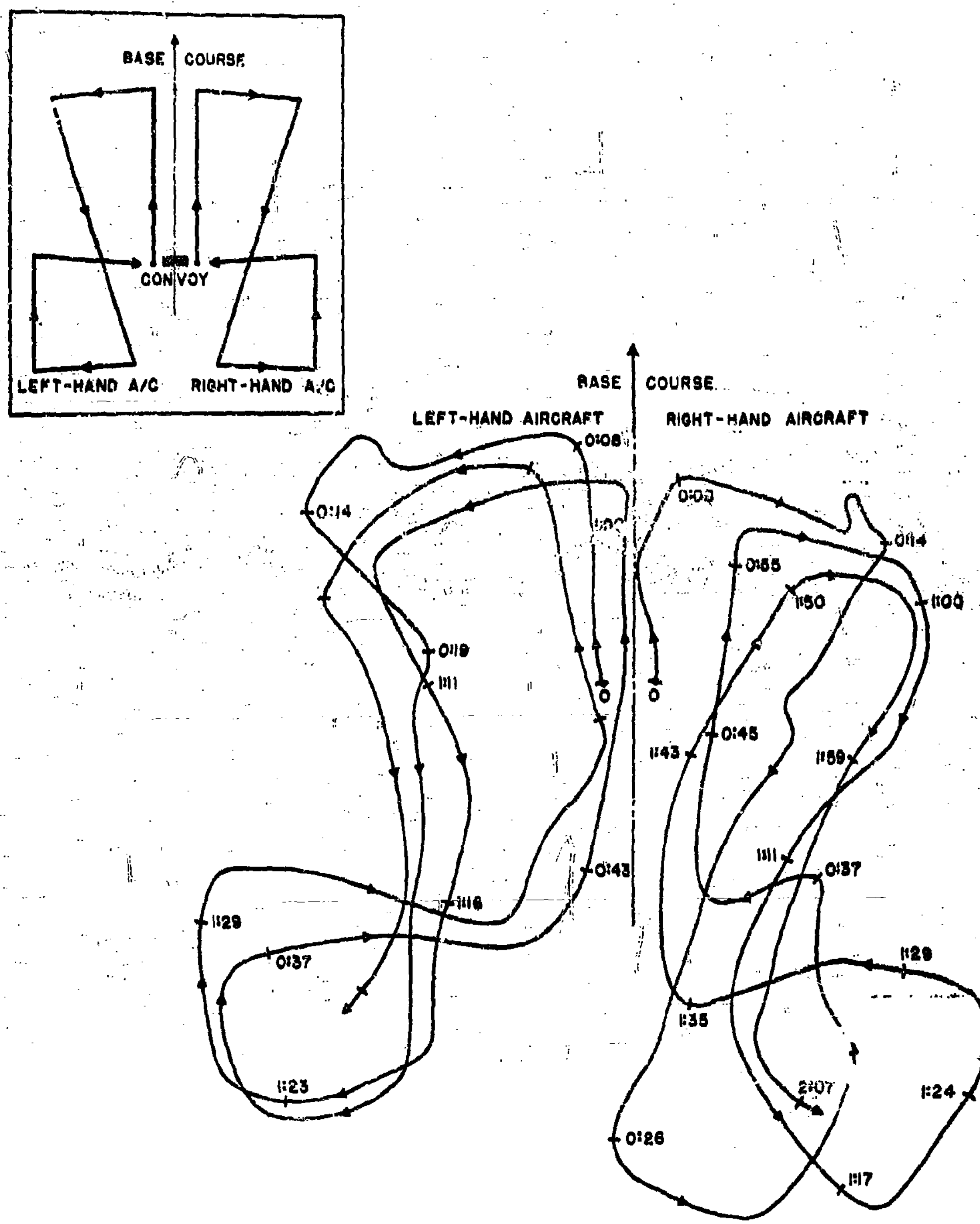


PLATE IV. Actual tracks of aircraft in operational flight. Inset shows search plan being used.

is advisable to make the designs first relative to the convoy rather than in space fixed with respect to the ocean, since the former picture is simpler and more direct in its illustration of relevant features; but eventually it will be transferred to geographic space, where the last refinements can be made.

When a satisfactory plan has been designed for one speed ratio $r = v/v_C$, it is usually desirable to extend it without changing its fundamental character, so that it will be applicable to broader ranges of r. This should be done without changing the angles of the courses in the original plan, if possible, but rather by varying certain leg lengths. It is simply the kinematic problem of bringing all planes back to the convoy periodically every T hours, given the new values of r. It is solved by simple trigonometry as follows: Draw a full period of the plan in geographic space, i.e., the figure that repeats itself every T hours. Consider the part of the plan flown by one aircraft. Let it be desired to vary two particular leg lengths, x and y, maintaining the others constant (values given). Let z be the distance the convoy moves in the time T. Then the three variables x, y, z are related by three equations of the first degree: first, the requirement that the length of the flown path (vT) divided by z ($v_C T$) be equal to $v/v_C = r$; second, that the algebraic sum of the projections of the legs on the convoy path shall be z; and third, that the algebraic sum of the projections on a normal to the convoy path be zero. Solving these equations, x, y, z are obtained as functions of r and, of course, T. Tables of the results are furnished with the plan. Of course, it is necessary to re-examine the plan (calculating scouting and screening diagrams anew, at least at critical places) for the extreme values of r used, in order to be sure that the necessary protection is maintained.

Since the effectiveness of a plan depends on the conditions of detection (visual or radar), reasonable and conservative estimates of visibility should be made in making all these calculations. It may be necessary to repeat such calculations for other visibilities. However, it is unlikely that a plan which is better than another at one visibility shall be worse than it at a different visibility, unless the change of visibility entails an essential change in the tactical situation.

It may be remarked that the diagram in Figure 18 exhibits an imperfection in the plan of Figure 17, namely, insufficient protection in the important forward regions. Better plans are actually used.

An important point of a practical nature must be kept in mind in designing and evaluating any air escort plan. It is that aircraft are not apt to fly a given plan to a very great degree of accuracy. Just how far they may deviate from the plan will depend on wind, visibility, radar and other navigational equipment, as well as on the temperament and experience of the pilot. It will depend also on his finding it necessary to investigate presumed contacts, which very often turn out to be false contacts. The chart shown on Plate IV gives the actual flights (drawn relative to the convoy) of two aircraft which were flying the simple straight-line plan shown in the insert (a plan which is now obsolete). The aircraft were actually tracked by radar from the carrier during an operation in World War II and their positions marked at the successive epochs indicated on the chart. This provides an object lesson on the difference between theory and practice. Incidentally, it explains why it is so often realistic to apply the formula of random flights of Chapter 2, equation (40). It has even been felt that, after all, the chief value of having a systematic plan is to get the (perforce, random) flights out where they would be effective, rather than to leave the pilot to follow his own devices, which has generally led to flights bunched up in the wrong places, usually too close to the convoy.

0.8 FINAL SWEEPS

As long as the convoy is in submarine waters, maximum protection is obtained by flying the escort plan without letup. But it is often necessary to discontinue the flights for a protracted period of time, for example, during the hours of darkness. The danger incident to this can be minimized by flying *final sweeps* at a greater distance than the normal flights, immediately after the latter are discontinued. These final sweeps are essentially of a scouting character; they aim at detecting submarines which might constitute a menace to the convoy at a later period, after aerial escort has been discontinued.

The first thing to realize is that it is not normally possible to detect by flights made at the time ($t = 0$) of discontinuance of aerial escort, all the submarines which could possibly close the convoy during the next H hours of escortless travel. For it would be necessary to sweep the area of the ocean in which such submarines must be; this, in the case of $u > v_C$, is the shaded circular region of Figure 19, obtained by multiplying all lengths of the circular diagram of relative speeds (vector **w**) by H. This means that an area of $\pi u^2 H^2$ would have to be swept. For example, if $u = 12$ knots and $H = 8$ hours, the area is 29,000 square miles. With an aircraft of 130 knots and sweep width 10 miles, a time of 29,000/1,300 = 22.3 hours

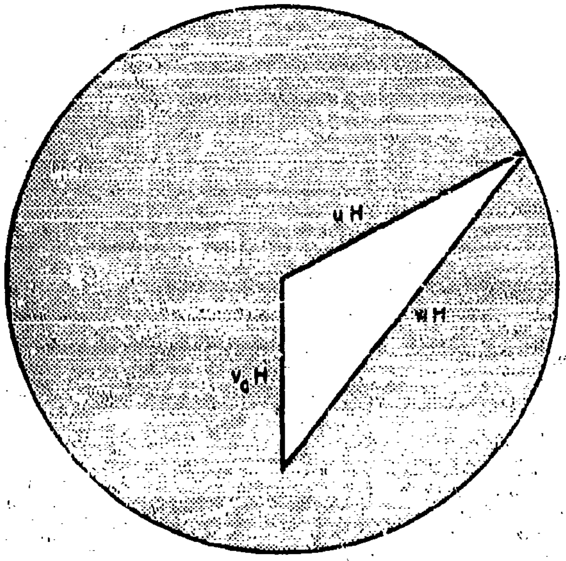


FIGURE 10. Circle of possible future contacts.

would be required. Even with four planes, 5.6 hours would be needed; and more, if the difficult track problem of covering the area is taken into account, with the resulting loss of efficiency. If less than this coverage is available, what should be covered and what neglected? This is essentially the question treated in Chapter 3 (specifically, the corollary at the end of Section 3.3).

Taking a system of rectangular coordinates with origin fixed at the convoy reference point O (and hence moving with the convoy), and axis of abscissas in the direction of convoy motion, we must evaluate the function $p(x,y)$, where $p(x,y)\,dxdy$ is the probability of there being a "dangerous" surfaced submarine in a $dxdy$ region at the point (x,y) (at the initial epoch $t = 0$), where by a "dangerous" submarine is meant one which will surely come in contact with the convoy (i.e., enter the detection circle of radius R, center O) during the subsequent H hours. Next, we must evaluate the amount of coverage or searching effort Φ available. Thus if we have the use of n aircraft for h hours each and if their speed is v and effective search width is W, $\Phi = nvhW$. According to the reasoning of Section 3.3, we are warranted only in searching the region of ocean A characterized by the two following properties:

1. A_b is the locus of points (x,y) for which $p(x,y) \geq b$; this determines the *shape* of A_b.

2. The value of b and therewith the choice of the region A_b are fixed by the requirement of equation (22) of Section 3.3. According to Chapter 3, A_b should then be searched in the manner specified by equation (23) of that chapter. A geometric construction is given therein. But in practice, such refinement is not apt to be possible, and the best procedure would be to use any practicable search, of a more or less uniform sort, in the estimated region A_b. Of course the whole of the detection circle itself must be included in the search to detect submarines already in contact with the convoy, particularly trailing behind and tracking on the flanks.

The estimation of $p(x,y)$ is difficult, not because of the mathematics involved (a crude mathematical formulation is quite sufficient) but because it depends on the correct appraisal of the tactical situation regarding the submarine. There are two extreme cases which may be considered as examples.

Case 1. The submarines are uniformly distributed over the part of the ocean of interest and are cruising at an estimated speed of u knots in any direction (i.e., in uniformly distributed random directions; see Section 1.4). This is the case in which the submarines are picked up purely by chance and not as the result of any systematic patrolling on their part. The problem of finding $p(x,y)$ in this case has been solved in Chapter 1 [see Section 1.5, problem 5, together with Figures 10, 11, 12, and 13; the $\Phi(r,\beta)$ of this reference is not to be confused with the "amount of searching effort" Φ used above]. In this reference, the probability $P(r,\beta)$ is found [(r,β) are the polar coordinates of (x,y)]; unlike $p(x,y)$, this is not a probability or expected value *density* but a *probability*, namely, the probability that the target given to be at (r,β) [or (x,y)] when $t = 0$ shall subsequently enter the detection circle. But if the mean density of submarines is N, there will be $N dxdy$ submarines *in all* in the $dxdy$ region at (x,y), of which the fraction $P(r,\beta)$ will enter the circle; hence the relevant expected number of submarines in $NP(r,\beta)\,dxdy = NP(r,\beta)\,rdrd\beta$. Now as in the proof of corollary at the end of Section 3.3, the constant of proportionality N is immaterial in the final results. Hence equations (22) and (23) as well as the geometrical construction given in that chapter apply with the former $p(x,y)$ replaced by $P(r,\beta)$. Finally, in the problem of Chapter 1 the time before contact was not limited, whereas we are restricting it to H hours. But with the searching effort normally available, only regions within an H hour submarine run of the detection circle can be searched, so that the time restriction applies automatically.

Let us consider as example the case $v_C/u = k = \frac{2}{3}$, e.g., submarine cruising at 15 knots, convoy at 10 knots. The level curves of constant probability $P(r,\beta)$ (i.e., the boundaries of A_b for different values of b) are given in Figure 11 of Chapter 1. Let the detection range of the submarine on the convoy be $R = 20$ miles. This gives the scale in Figure 11, Chapter 1, of 20 miles to the inch. With two 130-knot aircraft available for two hours and assuming the search width $W = 10$ miles, we have $\Phi = 2 \times 2 \times 130 \times 10 = 5{,}200$ square miles of searching effort available. To determine the probability contour of Figure 11, within which the search must take place, we have to invoke (22) of Chapter 3, which becomes

$$\iint_{A_b} \log P(r,\beta) r dr d\beta - A_b \log b = \Phi = 5{,}200.$$

The values of $P(r,\beta)$ for $k = \frac{2}{3}$ would have to be drawn from Section 1.5, Chapter 1, and the integration and calculation of the area A_b performed for different values of b, the arc finally chosen. Then a plan of search would be devised as much as possible in accordance with (23) of Chapter 3.

The following rough graphical version of this process illustrates the principles of the method in a form which can be carried out without unreasonable trouble in practice and which provides about all the accuracy which our rather dubious tactical and numerical assumptions would appear to warrant.

With a planimeter (or an approximating ellipse) calculate the areas of the various A_b regions of Figure 11 (Chapter 1), remembering that $b = 0.1$ for the 10 per cent region, $b = 0.15$ for the 15 per cent, $\cdots$, $b = 1$ for the detection circle. We find approximately for the A areas, and the ΔA rings between successive regions:

$$\begin{aligned}
A_1 &= 1{,}260, & \Delta A_1 &= A_1 - A_{0.4} = 150,\\
A_{0.4} &= 1{,}410, & \Delta A_{0.4} &= A_{0.4} - A_{0.25} = 1{,}710,\\
A_{0.25} &= 3{,}120, & \Delta A_{0.25} &= A_{0.25} - A_{0.15} = 3{,}030,\\
A_{0.15} &= 7{,}050, & \Delta A_{0.15} &= A_{0.15} - A_{0.10} = 8{,}950,\\
A_{0.1} &= 16{,}000.
\end{aligned}$$

To find the graph of the function

$$f(b) = \iint_{A_b} \log P(r,\beta) r dr d\beta$$

against b out to $b = 0.1$, we may multiply each ring area by the arithmetic mean of the (natural) logarithms of the probabilities of the two bounding level curves (except initially, the area of the detection circle is multiplied by $\log 1 = 0$), and then add the results out to the ring bounded by the value of b in question. The curve for $f(b)$ is then drawn through the resulting five points. Next, the function

$$F(b) = 5{,}200 + A_b \log b$$

is graphed, taking the five values of A_b given above and multiplying by $\log b$, b being the value corresponding to the boundary; a smooth curve is then passed through the resulting points. These curves are found to intersect at about the point $b = 0.13$, i.e., $f(0.13) = F(0.13)$; but this is equation (22) of Chapter 3.

TABLE 1.

b	$\log b$	Avg log b	A ($A_1 = 1{,}260$)	Avg log $b \times \Delta A_b$	$f(b)$	$F(b)$
1.0	0	0		0	0	5,200
0.4	−0.91	−0.45	150	−68	−68	3,910
0.25	−1.40	−1.20	1,710	−2,060	−2,128	550
0.15	−1.90	−1.65	3,030	−6,650	−8,778	−8,100
0.10	−2.30	−2.04	8,950	−18,300	−27,078	−30,000

Now since one cannot hope to search in accordance with equation (23), Chapter 3, the best practical recommendation is to cover a region extending a little beyond the 15 per cent curve of Figure 11 (Chapter 3), for example, by a sector search. The probability of detecting the target by a random search of the 15 per cent curve, is, by (40) of Chapter 2,

$$p = 1 - e^{-LW/A} = 1 - e^{-\Phi/A_{0.15}}$$
$$= 1 - e^{-5{,}200/7{,}050} = 0.52.$$

Thus, there is a little over half a chance that a given submarine in the 15 per cent region will be detected. Therefore, the search is decidedly useful, but by no means as good as having regular escort flown through the night.

Case 2. The submarines are mounting guard along the route of the convoy, as they estimate it. If there are several submarines acting as a coordinated group, they may form a line patrol across the convoy path of a type illustrated by the following example: Submarines are placed in a line across and at right angles with the convoy path patrolling stations 20 miles apart; each submarine proceeds at a low speed ($u = 10$ knots) and is never more than half an hour's run from its station. [This was a German plan described in the British Monthly Anti-Submarine Report (secret), November 1942.] Submarines acting alone may patrol at greater distances about their stations, but in all cases the distance is limited by the objective of intercepting convoys moving along the line.

This case is easier to treat than the first, since the only dangerous submarines here lie in a band whose middle line is the convoy's future course and whose total width is $2(R + d)$, where d is the presumed distance of lateral patrol of the submarines, while R is (as before) the submarines' detection range on the convoy during the time of discontinuance of the normal aerial escort (e.g., at night). For a submarine at lateral range more than $R + d$ on either side of the convoy's course would be unlikely to constitute a threat; $p(x,y) = 0$ outside the band, $p(x,y) =$ constant > 0 within the forward length of the band of Hv_C miles. Thus, this whole piece of band must be swept. This is a feasible operation. With $R = 20$, $d = 10$, $v = 8$, $H = 8$, the rectangular band's area is 3,840 square miles; one plane of 130 knots and 10-mile search width would require 3,840/1,300 = 2.95 hours. With two or three planes, not only this area but a somewhat wider one, including the daytime detection circle, could easily be swept in a very reasonable time, as the last hour or so of daylight.

But in setting up this search, it is essential to know in advance what deceptive steering is intended during the night. Convoys usually make a deceptive change of course shortly after dark; the band may have to be a bent one under such circumstances.

GLOSSARY

A/C. Aircraft.

ACOUSTIC TORPEDO. Homing torpedo guided to its target by means of echo ranging or listening.

ANOMALOUS PROPAGATION. In sonar, pronounced and rapid variations in echo strength caused by large and rapid local fluctuations in propagation conditions.

ASB RADAR. A 60-cm Navy radar for surface search by carrier-based aircraft.

ASDIC. British echo-ranging equipment; letters are derived from "Anti-Submarine Development Investigation Committee."

ASG RADAR. AN/APS-2, a 9-cm ASV and search radar.

ASV RADAR. A radar system for detecting and homing on a surface vessel from the air.

ASYMMETRY FACTOR. Ratio of target length to width.

BARRIER LINE. Mathematical reference line across and perpendicular to a channel.

BEARING. Angular position of target with respect to own ship (relative bearing) or to true north (true bearing).

BLIP. In radar, echo trace on indicator screen.

BRIGHTNESS CONTRAST. The difference between target and immediate background brightness expressed in units of effective background brightness.

BROWNING SHOT. A shot aimed at a general area containing targets, on the chance of a random hit.

B-SCOPE. A scope which presents a rectangular plot of range versus bearing. Spot brightness indicates echo intensity.

CAVITATION. The formation of vapor or gas cavities in water, caused by sharp reduction in local pressure.

CHANNEL. Strip of ocean through which it is known targets must pass.

COLLISION COURSE. Course steered by an attacking craft to intercept the attacked craft's course so that a coincidence will occur between the two crafts.

CONTACT. An instance of detection of an enemy unit.

CONTACT PROBLEM. Study of the capabilities of the detection agency.

CONVOY. A group of merchant ships sailing together, usually defended by naval craft escorts.

CURLY TORPEDO. A torpedo which, after closing the convoy, steers a sinuous course to increase its chance of a random hit.

CVE. Navy designation for Aircraft Escort Carrier, a small aircraft carrier.

DEFINITE RANGE LAW OF DETECTION. The assumption that detection is sure and immediate within a given critical range and impossible beyond that range.

DENSITY. Number of objects per unit area.

DE. Destroyer escort, a large, high-speed antisubmarine ship.

DF. Radio Direction Finding; use of directional receivers to estimate enemy position by intercepting and obtaining the bearings of enemy transmissions.

ECHO RANGING. Method of locating underwater objects by sending sound pulses into the water and receiving an echo reflected from the target. Target range is derived by measuring total transit time of the sound pulse.

EFFECTIVE RADAR CROSS SECTION. A measure of the reflecting ability of the target; it indicates the ratio of total power reflected from the target to incident power density impinging from the radar.

EFFECTIVE SEARCH WIDTH. Twice the range of an equivalent definite range law of detection (equivalent, with respect to detecting a uniform distribution of targets).

EFFECTIVE VISIBILITY, E. In parallel sweeping, half the sweep spacing for which the probability of detection is 0.5.

EPOCH. Any given instant of time.

ERF. Error function or probability integral:

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} dx.$$

EVASIVE ROUTING. Routing to avoid known submarine positions.

FIX. Presumptive position of target as determined from a single observation.

FIXATION. The short stationary periods of actual seeing between jumps of the eye in visual scanning motion.

FOVEA. Central part of retina, region of maximum acuity for daylight vision.

GLIMPSE PROBABILITY, g. Probability that a target be sighted at a single glimpse.

HOMING TORPEDO. Torpedo self-guided to its target by some property of the target.

KNOWN RANDOM DISTRIBUTION. A distribution in which the probabilities are known in advance.

LATERAL RANGE. The minimum distance (at closest approach) between target and observer. When the target is at rest, it is the perpendicular distance between the target and the observer's track.

LATERAL RANGE CURVE. A curve which gives the probability of detection of an object by an observer proceeding along a straight course as a function of distance of direct approach.

LEAD ANGLE. In this volume, the angle between the barrier line and line of flight of the observer aircraft.

LIMITING APPROACH ANGLE. Angle whose sine is the ratio of target velocity to observer velocity. This angle determines the slope of the lateral boundaries of the region of approach of target to observer.

LIMITING SUBMERGED APPROACH ANGLE. Limiting approach angle for a submerged submarine.

LISTENING. Use of sonar to detect sonic and supersonic sounds generated by the target itself.

LOOKING. Trying to detect with any of the means considered: visual, radar, sonar, etc.

MAXIMUM SIGHTING RANGE. The range at which the target contrast reaches the foveal threshold.

ONE-PING PROBABILITY. The probability of detection by using a single ping.

PING. Acoustic pulse signal projected by echo-ranging transducer.

PPI. Plan Position Indicator.

RADAR. Generic term applied to methods and apparatus that use radio for detection and ranging.

RECOGNITION DIFFERENTIAL. The number of decibels by which a signal level must exceed the background level in order to be recognized 50 per cent of the time.

REGION OF APPROACH. Area in which a craft, moving with velocity less than that of another craft, must have its starting point to be able to intercept the latter.

RELATIVE BEARING. Target bearing relative to own ship bearing.

RELATIVE COURSE, θ. Angle between the observer's vector velocity and the target's relative velocity.

RELATIVE VELOCITY, **w**. The target's vector velocity with respect to the observer.

REVERBERATION. Sound scattered diffusely back toward the source, principally from the surface or bottom and from small scattering sources such as bubbles of air and suspended solid matter.

SCANNING LINE. The locus of points on the ocean surface at which the detecting instrument is directed.

SCHNORCHEL. U-Boat combination air intake and exhaust tube that permits submerged diesel operation.

SEARCH RATE. See sweep rate.

SEARCH WIDTH. See sweep width.

SEA RETURN AREA. Ocean surface area which reflects radar pulse back to the transmitter.

SONAR. Generic term applied to methods and apparatus that use sound for navigation and ranging.

SWEEP RATE. A measure of the searching craft's effectiveness in covering an area. It is the number of contacts made per hour per unit of target density by the searching craft, expressed in square miles per hour.

SWEEP SPACING, S. Distance between track lines in parallel and equally spaced search.

SWEEP WIDTH. Sweep rate divided by speed (approximately).

TARGET. Object of search.

TARGET ASPECT. Orientation of the target as seen from own ship.

TARGET STRENGTH. A measure of the reflecting power of the target. Ratio in decibels of the target echo to the echo from a 6-ft diameter perfectly reflecting sphere at the same range and depth.

THRESHOLD CONTRAST. Just perceptible contrast of target against its background.

TIME OF FIX. Time or epoch at which information about the point of fix is given.

TORPEDO DANGER ZONE. Area around a ship, and moving with the ship, within which a torpedo must be fired if it is to have any chance of scoring a hit.

TRACK ANGLE, ϕ. Angle between the vector velocities of observer and target.

TRACK PROBLEM. Selection of path and motion of the observer for assumed position and motion of the target.

TRANSDUCER. Any device for converting energy from one form to another (electrical, mechanical, or acoustical). In sonar, usually combines the functions of a hydrophone and a projector.

u-MOVING SPACE. Barrier flight plan as seen from a target moving with speed u.

X-BAND. Band of radar frequencies with wavelength approximately 3 cm.

ZIGZAG. Change course frequently to make attack by submarines difficult.

INDEX

The subject indexes of all STR volumes are combined in a master index printed in a separate volume. For access to the index volume consult the Army or Navy Agency listed on the reverse of the half-title page.