Target Resolution: Capabilities of Modern Radar and Fundamental Limits

27 JULY 1964

Prepared by A. W. RIHACZEK
Electronics Research Laboratory

Prepared for BALLISTIC SYSTEMS AND SPACE SYSTEMS DIVISIONS

AIR FORCE SYSTEMS COMMAND

LOS ANGELES AIR FORCE STATION

Los Angeles, California

LABORATORY OPERATIONS • AEROSPACE CORPORATION

CONTRACT NO. AF 04(695)-269

SSD-TDR-64-107

Report No.
TDR-269(4250-43)-2

TARGET RESOLUTION: CAPABILITIES OF MODERN RADAR AND FUNDAMENTAL LIMITS

Prepared by

A. W. Rihaczek
Electronics Research Laboratory

Laboratory Operations
AEROSPACE CORPORATION
El Segundo, California

Contract No. AF 04(695)-269

27 July 1964

Prepared for

BALLISTIC SYSTEMS AND SPACE SYSTEMS DIVISIONS
AIR FORCE SYSTEMS COMMAND
LOS ANGELES AIR FORCE STATION
Los Angeles, California

SSD-TDR-64-107

Report No.
TDR-269(4250-43)-2

TARGET RESOLUTION: CAPABILITIES OF MODERN RADAR AND FUNDAMENTAL LIMITS

Prepared

A. W. Rihaczek

Approved

B. J. DuWaldt, Head
Electromagnetic Techniques Department

D. D. King, Director
Electronics Research Laboratory

This technical documentary report has been reviewed and is approved for publication and dissemination. The conclusions and findings contained herein do not necessarily represent an official Air Force position.

For Space Systems Division
Air Force Systems Command

Robert D. Eaglet
Captain, USAF
Space Technology Division

FOREWORD

This report was written for presentation at the Eighth Technical Meeting of the Avionics Panel of the Advisory Group for Aeronautical Research and Development (AGARD) of NATO, London, September 1964. The nature of the report was dictated by the purpose of the meeting: to familiarize the attendees with the state of the art and the problems in the field of radar.

SSD-TDR-64-107

Report No.
TDR-269(4250-43)-2

ABSTRACT

The growth in radar requirements, from crude short-range measurements to dense-target resolution at long ranges, is paralleled by a corresponding increase in the sophistication of radar systems: from simple constant-carrier pulses to post-detection integration, coherent integration of pulse trains, pulse compression signals, and the coherent processing of trains of such signals. The study starts with an interpretation of the radar uncertainty relation in its significance for target resolution, showing the role of waveform design as a means of achieving a match between the transmitted signal and the characteristics of the target environment. This provides a framework into which the various principles of high-resolution radar are fitted. In discussing the limitations on resolution performance, it is shown that achievable target resolution depends on the characteristics of the target environment in which the radar operates, the number of targets, and the size of the delay-Doppler space they occupy. These findings are applied to two practical examples: (1) the ground mapping radar using the synthetic aperture principle and (2) the case of extended target "clouds" consisting of a large number of discrete scatterers.

Preceding page blank

CONTENTS

Preceding page blank

FIGURES

I. INTRODUCTION

From the basic task of radar, the detection of a target and determination of its range, the step toward performing the same measurements on a group of targets is seemingly a small one; however, this step has so far taken the effort of the past two decades and is not yet completed. In the development of radar, multiple target resolution has turned out to be one of the most difficult problems to solve. The following discussion is an attempt to expose the fundamental nature of the resolution problem and, in putting the modern radar techniques into proper context, to present a unified view of the capabilities and limitations of radar for target resolution.

The trivial question of resolution when targets have an angular separation of more than a beamwidth will not be considered here. Moreover, it will be assumed that all targets actually illuminated by the beam are being illuminated with the same signal strength, so that no resolution on the basis of the beam pattern is possible. Thus we shall study the problem of target resolution on the basis of range and range rate measurements. Range acceleration and higher order range derivatives will be neglected, which fact is not too restrictive in radar practice as the effect of target acceleration, and more so of higher order range derivatives, usually is indeed negligible over the coherent signal processing time.

It will be further assumed that the radar uses a matched-filter or correlation receiver, that is, a receiver which gives optimum detection performance in additive, white Gaussian noise. An obvious justification for this assumption is the fact that most target detection radars use this type of receiver or at least approximate it. However, even though target resolution appears to be the problem of signal detection in a signal background rather than a background of thermal noise, it is not likely that a more efficient receiver than the matched-filter type can be found. A justification of this contention is beyond the scope of this paper, but can be indicated by pointing out that the essence of target resolution is in concentrating the individual returns in delay and Doppler so that they can be recognized as separate returns. Allowing a suitable modification of the transmitted signal waveform, this goal can be best achieved with a matched-filter system, and the problem becomes that of signal design for optimum resolution performance.

The basis for analyzing combined range and range rate resolution in a matched-filter radar was laid by Woodward (Ref. 1) through his study of the combined filter response in delay and Doppler or so-called ambiguity function. Work along the same lines was continued by Siebert (Ref. 2), and an application of these concepts to target resolution in clutter was given by Westerfield, et al. (Ref. 3), and to resolution in a dense-target environment by Fowle, et al. (Ref. 4). The purpose of the present study is to give a general exposition of the resolution problem, so as to show what high-performance radar can and cannot do.

II. AMBIGUITY FUNCTION AND RADAR UNCERTAINTY RELATION

First, we summarize briefly the notation and basic concepts to be used in the following analysis of target resolution. Signals will be written in the complex notation as introduced by Gabor (Ref. 5),

$$\psi(t) = \mu(t)\, e^{j2\pi f_0 t} = a(t)\, e^{j\phi(t)}\, e^{j2\pi f_0 t} \quad , \qquad (1)$$

where $\mu(t)$ is the complex modulation function of the real signal, f_0 is the carrier frequency, $a(t)$ is the amplitude modulation function or real signal envelope, and $\phi(t)$ is the phase modulation function. The frequency modulation function is the time-derivative of the phase modulation function, $\phi'(t)$. The real signal is the real part of the complex signal, and the complex signal is obtained from the real signal by omitting the negative frequencies and doubling the amplitude of the positive frequencies. In the case of radar, the signal bandwidth is usually small compared to the carrier frequency, and the modulation functions defined above have practical meaning.

The notation will be simplified without any restrictions on the generality of the results by normalizing the signal energy such that

$$\int_{-\infty}^{+\infty} |\mu(t)|^2\, dt = \int_{-\infty}^{+\infty} |M(f)|^2\, df = 1 \quad , \qquad (2)$$

where M(f) is the Fourier transform of the complex modulation function $\mu(t)$, that is, the frequency spectrum of the complex modulation function. A further simplification is achieved by defining the carrier frequency, in agreement with common usage of the term, as the normalized (because of Eq. 2) first moment of the energy density spectrum of the signal,

$$f_0 = \int_0^{\infty} f\, |\Psi(f)|^2\, df \quad , \qquad (3)$$

and choosing the origin of time such that the normalized first moment of the squared signal envelope is zero,

$$\overline{t} = \int_{-\infty}^{+\infty} t\,|\mu(t)|^2\,dt = 0 \qquad . \tag{4}$$

With the above conventions, the definitions for three very useful signal parameters take relatively simple forms. The first parameter, the effective signal bandwidth β_0, plays an important role for measurement precision and resolution in range. It is defined as the rms deviation of the energy density spectrum from its mean or β_0^2, equivalently, as the second moment of the energy density spectrum of the modulation function:

$$\beta_0^2 = (2\pi)^2 \int_{-\infty}^{+\infty} f^2\,|M(f)|^2\,df \qquad . \tag{5}$$

Analogously, measurement precision and resolution in range rate depend on the effective signal duration t_0, defined as the second moment of the squared signal envelope:

$$t_0^2 = (2\pi)^2 \int_{-\infty}^{+\infty} t^2\,|\mu(t)|^2\,dt \qquad . \tag{6}$$

The third parameter is an average measure for the internal phase structure of the signal and, as will become clear later, relates to the combined precision in range and range rate. This parameter, α, will be referred to as the effective phase constant of the signal. It is defined as

$$\alpha = 2\pi \int_{-\infty}^{+\infty} t\phi'(t)\,|\mu(t)|^2\,dt \qquad , \tag{7}$$

and hence measures the linear FM content in the frequency modulation function $\phi'(t)$ of the signal. It can be shown that among all possible types of modulation and for fixed β_0 and t_0, α reaches its maximum value for linear FM.

In the presence of returns from more than a single target, the matched-filter response will be the superposition of responses to signals having different amplitudes, range delays, and Doppler shifts. Thus the tool for analyzing the resolution problem is the two-dimensional matched-filter response in delay and Doppler as used by Woodward (Ref. 1). Since it is usual practice to destroy the fine phase information by envelope-detecting the matched-filter output, the actual quantity of interest is the complex envelope of the response and its absolute value, the real envelope. The complex envelope of the matched filter response can be written in the two alternative forms (Ref. 1),

$$\chi(\tau, \nu) = \int_{-\infty}^{+\infty} \mu(t)\mu^*(t - \tau)\, e^{j2\pi\nu t}\, dt \tag{8}$$

$$= \int_{-\infty}^{+\infty} M^*(f)M(f - \nu)\, e^{j2\pi f\tau}\, df \quad , \tag{9}$$

where τ and ν are the differential range delay and Doppler shift, respectively, with the origin $\tau = \nu = 0$ chosen at the peak output for a perfectly matched filter. In other words, τ is the running time variable for the matched filter response and ν is the Doppler mismatch between return signal and "matched" filter.

The complex envelope of the matched-filter response, $\chi(\tau, \nu)$, is commonly known as the ambiguity function, alluding to the fact that subsidiary peaks of the response may, in the presence of noise, introduce an uncertainty as to the presence and location of a target. The ambiguity function is, of course, simply the filter response to a signal which is matched to the filter except for an arbitrary Doppler shift ν. The plot of $|\chi(\tau, \nu)|^2$ as a surface above the τ, ν-plane will be referred to as the ambiguity surface, and a vertical cut through this surface for constant Doppler shift ν_0 gives the power response of the matched filter for that particular value of Doppler mismatch. This permits us to indicate an ambiguity surface through a series of filter responses as shown in Fig. 1.

To summarize the important properties of the ambiguity surface, application of the Schwartz Inequality to $|\chi(\tau, \nu)|^2$ in Eq. (8) leads to

$$|\chi(\tau, \nu)| \leq \int_{-\infty}^{+\infty} |\mu(t)|^2 \, dt = \chi(0, 0) = 1 \quad . \tag{10}$$

In words, no subsidiary peak of the ambiguity surface can exceed the central peak in height, and this peak has a height of unity, regardless of the signal waveform. Furthermore, as already pointed out by Woodward, the total volume under the ambiguity surface equals unity,

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\chi(\tau, \nu)|^2 \, d\tau \, d\nu = 1 \quad . \tag{11}$$

The essence of Eqs. (10) and (11) is that, for constant signal energy, the volume under the ambiguity surface and the measurement interference it represents can be shifted around in the τ, ν-plane by changing the signal waveform, but cannot be reduced. Since changing the signal energy affects all target returns in the same manner, an increase in signal energy improves detection performance in noise but does not improve target resolution.

The volume constraint results from the definition of time and frequency as a pair of Fourier transforms, which prevents a signal from being sharply confined in both time and frequency domain simultaneously. This fact leads to the uncertainty principle of quantum mechanics and, in radar systems, to constraints on the combined resolution in range and range rate. Lacking a better terminology, we shall refer to this phenomenon as the radar uncertainty

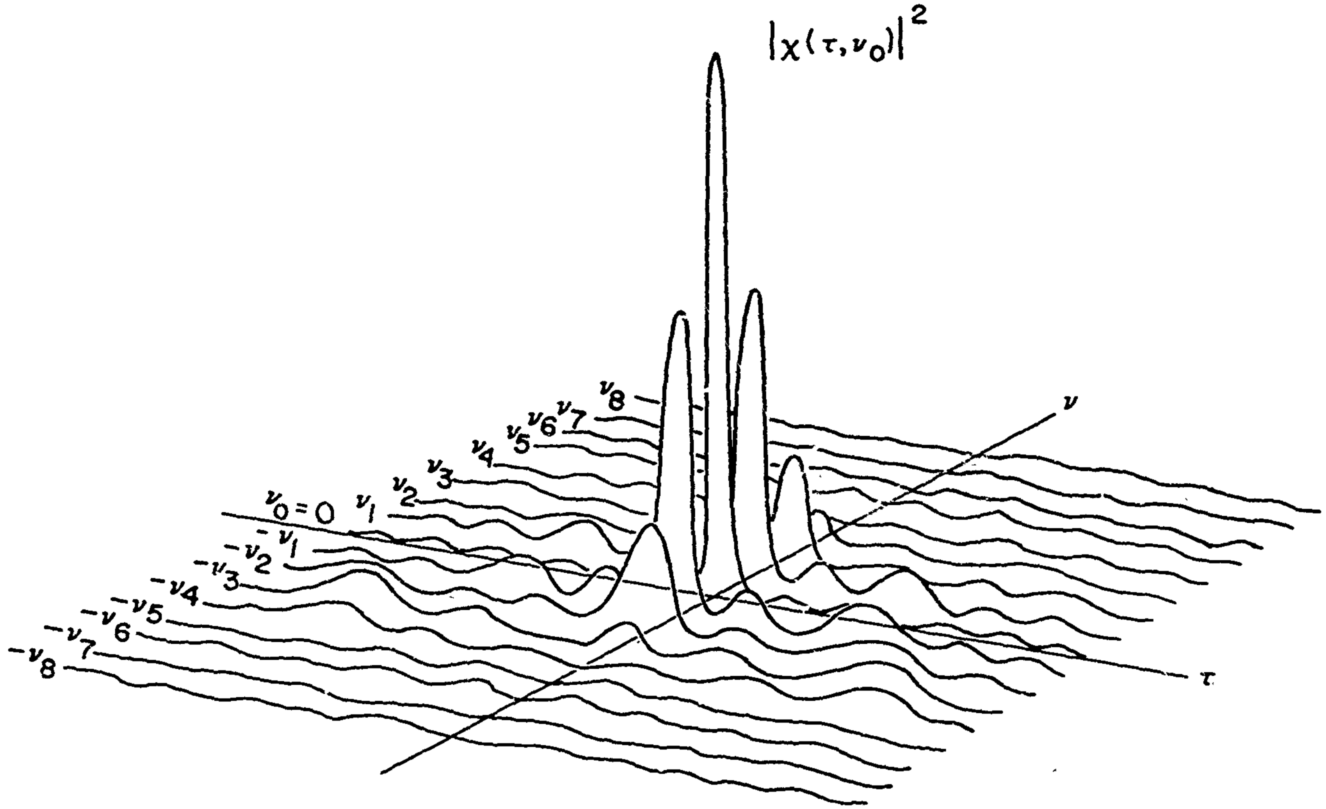

Fig. 1. Ambiguity Surface as the Sum of Matched-Filter Responses to Signals Mismatched in Doppler

relation. For simple constant-carrier pulses, the uncertainty relation means that squeezing the matched-filter response in the delay domain to achieve better definition of the response in range must result in a corresponding widening of the response in Doppler, implying a degradation in range rate resolution. For signals of arbitrary waveforms, the effects of the uncertainty relation can be put in evidence by integrating $|\chi(\tau, \nu)|^2$ in τ and, with the use of the Parseval theorem, deriving a relation given in Ref. 3:

$$\int_{-\infty}^{+\infty} |\chi(\tau, \nu)|^2 \, d\tau = \int_{-\infty}^{+\infty} |\chi(\tau, 0)|^2 \, e^{-j2\pi\nu\tau} \, d\tau \quad . \tag{12}$$

The left integral is the area under a vertical cut through the ambiguity surface for constant Doppler shift, and as such describes the total volume distribution in the Doppler domain. The right integral is the Fourier transform of the ambiguity surface on the delay axis. Hence, Eq. (12) states that the total volume distribution of the ambiguity surface in the Doppler domain is given by the value of the ambiguity surface on the delay axis through a Fourier transform relation. This means that if the ambiguity surface is narrowed along the delay axis in order to improve measurement precision and resolution in range, the volume must spread proportionately in the Doppler domain. This is the general form of the radar uncertainty relation. For a constant-carrier pulse, spreading of the volume in the Doppler domain means widening of the central response peak and degradation of proximal target resolution. Though this need not be true for an arbitrary waveform, the spreading of the integrated volume as dictated by the uncertainty relation may lead to an equally objectionable degradation of target resolution in general. In fact, it will become clear from later discussions that changing the signal waveform merely permits us to trade the size of the resolvable cell against the minimum strength of targets for which this degree of resolution can be achieved.

The Fourier constraint on the volume distribution must, of course, also apply in the delay domain. Interchanging the role of delay and Doppler in Eq. (12), we find the dual relation

$$\int_{-\infty}^{+\infty} |\chi(\tau, \nu)|^2 \, d\nu = \int_{-\infty}^{+\infty} |\chi(0, \nu)|^2 \, e^{j2\pi\nu\tau} \, d\nu \quad , \tag{13}$$

which has analogous consequences. The total volume distribution in the delay domain depends only on the ambiguity surface on the Doppler axis.

The minimum resolvable cell that can be provided (under circumstances to be considered later) depends on the shape of the central spike of the ambiguity surface. If this central spike is expanded into a double Taylor's series about the origin and terms of higher order than quadratic are dropped, a horizontal cut at 75 percent of peak height results in a contour of elliptical shape and widths along delay and Doppler axis that are given as the inverse effective signal bandwidth and duration, respectively. If a suitable scale is chosen for delay and Doppler axis, the situation is as shown in Fig. 2. For a given value of signal bandwidth and signal duration, the cross section of the central spike is smallest when the phase constant α is zero. This is indicated by the circle in Fig. 2. As the value of $|\alpha|$ increases, the circle is stretched into an ellipse, with the largest ellipse obtained when the signal is a linear FM pulse of Gaussian envelope. The area of the ellipse can be calculated from the series expansion as

$$A = \pi \frac{1}{2\beta_0} \frac{1}{2t_0} \frac{1}{\sqrt{1 - \left(\alpha^2/\beta_0^2 t_0^2\right)}} \quad , \tag{14}$$

which shows the effect of α in increasing the area over the value for $\alpha = 0$, the circle in Fig. 2.

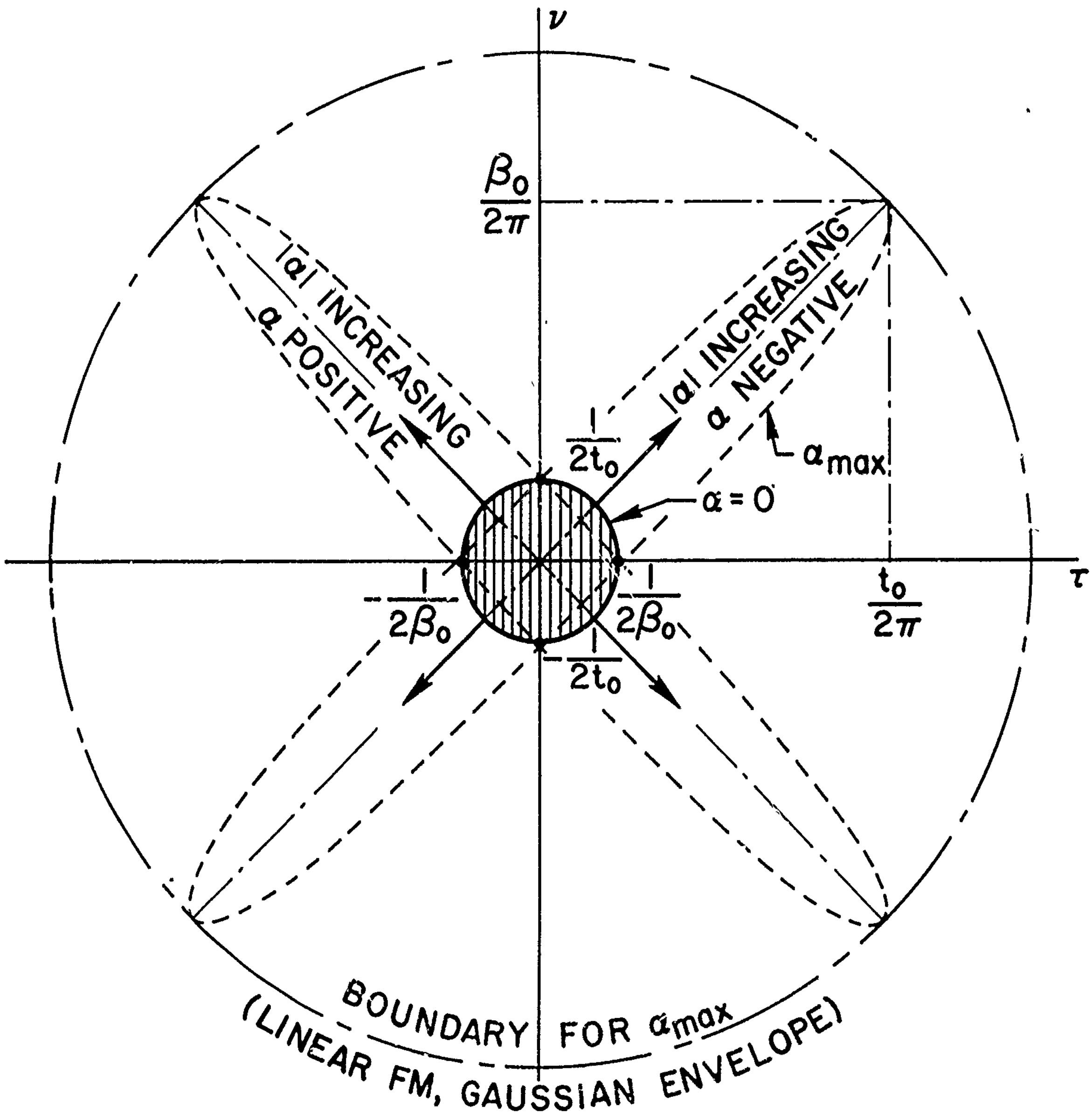

Fig. 2. Contour of the Central Spike of the Ambiguity Surface for Fixed Signal Bandwidth and Duration and Varying Phase Constant

III. SIGNAL DESIGN FOR RESOLUTION

The constraints on the ambiguity surface of a signal, or the radar uncertainty relation, will now be used for an exposition of the general resolution problem. This approach also leads to an understanding for the significance of existing and feasible principles of high-resolution radar.

A. RESOLUTION AS A TRADE BETWEEN TARGET SEPARABILITY AND BACKGROUND VISIBILITY

Starting with the simplest case, we assume that the radar is to resolve two point targets having the same radar cross section. The receiver response to the combined return can be obtained from the superposition of the two ambiguity functions (one translated with respect to the other by the amounts of the differential range delay and Doppler shift) by taking the value of the combined function at the Doppler frequency to which the filter is matched. When the targets are separated in range only, the width of the central response peak along the delay axis must be smaller than the differential range delay if the two targets are to be resolved in the presence of interfering noise. Analogously, when the targets are separated in Doppler only, the same must apply for the central response width along the Doppler axis. The extension to the case where the targets are separated in both range and range rate is also obvious.

The effect of waveform design on the radar resolution performance is illustrated in Fig. 3. As indicated in Fig. 3a, narrowing of the ambiguity surface along the delay axis for a constant-carrier pulse, by shortening the pulse duration, is possible only at the expense of widening the response in the Doppler domain correspondingly. If the pulse duration is held constant and the desired increase in bandwidth is achieved by modulating the carrier frequency, linear FM twists the ambiguity surface as shown in Fig. 3b. Though resolution in range and range rate separately now is high, it is seen that targets lying in the direction of the ridge will not be resolved. On the

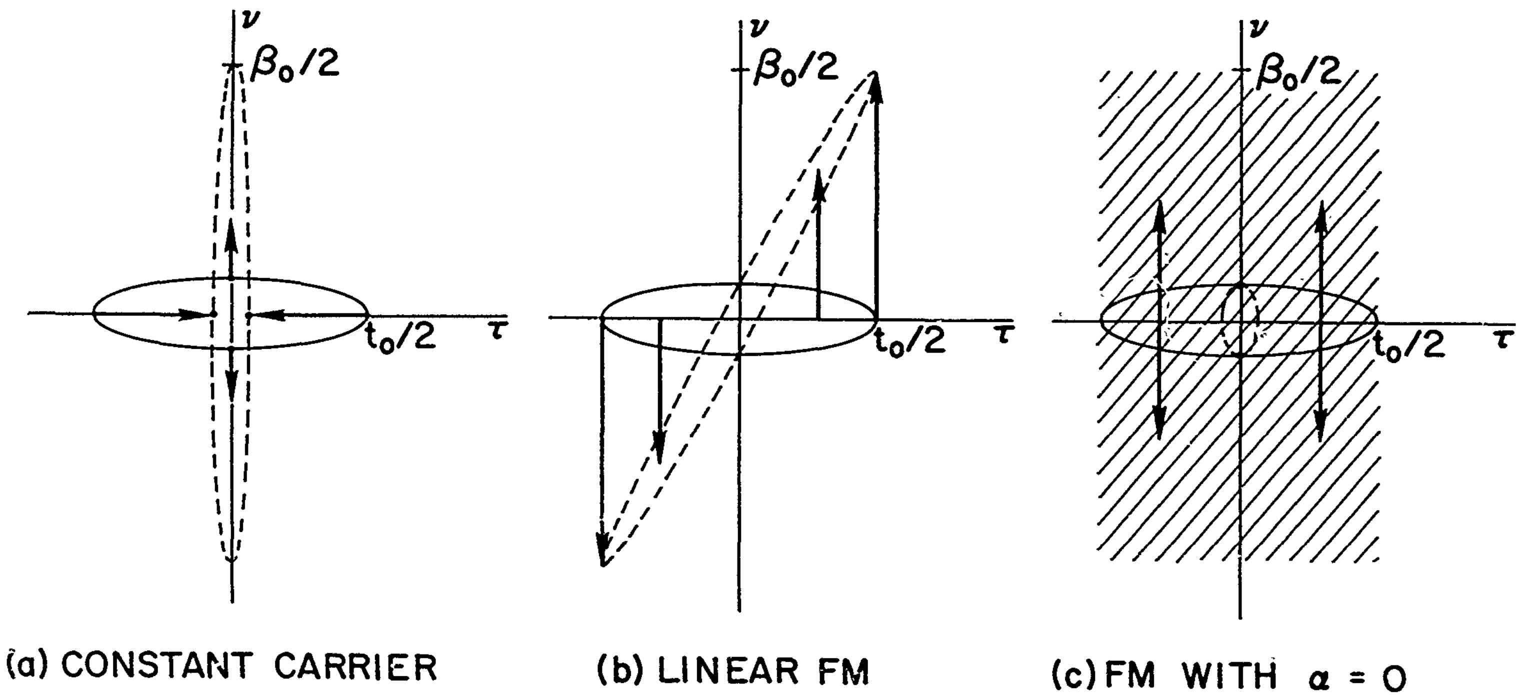

Fig. 3. Alternative Methods of Reducing the Width of the Central Spike in Range Delay

other hand, if the bandwidth increase is obtained by nonlinear FM or some other low-α modulation, in the limit with a type of modulation where the phase constant α is zero, the central peak is narrowed through smearing of the original surface in the manner of Fig. 3c.

For the case under consideration, two targets of the same cross section, a "pulse compression" signal with an ambiguity surface as shown in Fig. 3c is evidently optimum, as it will permit arbitrarily high resolution in range and range rate, provided the signal bandwidth and duration can be made sufficiently large. However, when the two targets have widely differing cross sections, the pedestal of the ambiguity surface of Fig. 3c will be objectionable. In fact, the height of the response pedestal of the strong target may be much larger than the peak of the response from the second target and thus may completely mask it. The radar then cannot "resolve" the two targets, no matter how widely separated they are in terms of the main response width. Furthermore, in the presence of a large number of targets, the clutter background introduced through the superposition of the individual pedestals may become so high that even relatively strong targets cannot be recognized.

The essence of the foregoing discussion is that target resolution cannot be specified in terms of the dimensions of the central spike of the ambiguity surface alone. Waveforms with low value of the phase constant α (and only these can truly reduce the area covered by the central spike) introduce what may be called "self-clutter" into the radar system, and the problem of target resolution becomes one of recognizing the desired target in this clutter background, in addition to separating the returns from closely spaced targets. This clutter is the effect of the radar uncertainty relation and is basic to the problem of combined resolution in range and range rate. In the following discussion, we shall treat target resolution as the task of recognizing a particular target in the combined interference from all other targets. Since target resolvability is of interest only if even the weakest target of interest can be reliably detected in the thermal noise of the system, we shall ignore this type of noise in the consideration of resolution.

B. CHOICE OF OPTIMUM RADAR WAVEFORMS

In arriving at an optimum waveform, we have to find first the optimum ambiguity surface for a particular application and then determine the corresponding signal waveform. In a strict sense, the problem of synthesizing a waveform in accordance with a prescribed ambiguity surface has not been solved yet; however, this is of little concern for the analysis of target resolution. Once the general properties of the ambiguity surface, as outlined above, and the significance of the three signal parameters β_0, t_0, and α are understood, it is relatively easy to visualize the interrelation between signal waveform and its associated ambiguity surface.

The returns from the illuminated targets will have delays and Doppler shifts that fall within certain boundaries in the τ, ν-plane. We shall refer to the area over which the returns are spread in this plane as the occupied target space. Since the radar uncertainty relation prevents the achievement of the ideal ambiguity surface, which is a single sharp spike, the problem of waveform design is evidently that of matching the ambiguity surface to the properties of the occupied target space. By this we mean that the volume under the ambiguity surface should be distributed over the τ, ν-plane in such a manner that it does not interfere with the measurement. To the degree to which this goal can be accomplished, waveform design will yield an improvement in the resolution performance of radar. (Since the discussion will be given in terms of delay and Doppler rather than range and range rate, it is well to point out that the target Doppler spread, and hence the size of the occupied target space, is proportional to the carrier frequency.)

As seen from Fig. 3, the problem of waveform design is simple when resolution in only one coordinate is needed. When all targets have the same range rate, we choose a constant-carrier pulse of short duration and obtain an ambiguity surface as indicated by the dashed ellipse of Fig. 3a. At some point, the signal may become so short that the transmitter cannot supply adequate energy within so short a duration. Then frequency modulation can be used to increase the signal bandwidth without decreasing the duration, and the most practical type of FM to choose may be linear FM as shown in Fig. 3b.

The advantage of linear FM, or Chirp, signals comes from the fact that the signal stays matched to the filter (or is properly "compressed") even if it has the wrong Doppler shift. Furthermore, a Chirp signal is relatively easy to generate and process. The practical problem is to design the signal such that the residual volume in the vicinity of the delay axis is small or, in more practical terms, that the remaining range "side lobes" after compression are very much smaller than the central peak. Theoretically, however, any other type of FM would serve just as well, and there are no limitations on achievable resolution.

The situation is similar when all targets have the same range and only range rate resolution is needed. Since high range rate resolution requires long signals, the limitation is not the transmitter power but the receiver isolation. For long-range operation, the signal duration cannot exceed the width of the occupied target space in the delay domain, so that coherent pulse trains must be used when still higher range rate resolution is required than corresponds to the limiting value of the single-pulse duration. The properties of pulse train type signals will be considered later.

Of more interest is waveform design when targets have to be resolved in both range and range rate. In a common practical situation, we know no more about the properties of the target environment than that targets may have a certain spread in range delay and Doppler. In other words, the only prior information available and usable is that the targets fall within certain boundaries in target space. The optimum ambiguity surface for this case consists of a narrow central spike, with the remainder of the volume spread out into a uniformly low pedestal. This ambiguity surface is usually referred to as a "thumbtack" surface, and its characteristics may be deduced from the Fourier constraints of Eqs. (12) and (13): for a central spike width in the order of $1/\beta_0$ and $1/t_0$ along delay and Doppler axes, respectively, essentially all of the unity volume must be contained within $|\tau| \leq t_0$ and $|\nu| \leq \beta_0$. Furthermore, for the high time-bandwidth product necessary to achieve a thumbtack type surface the volume in the spike is evidently negligible, and the height of the pedestal must be $1/\beta_0 t_0$ to satisfy the volume constraint. These over-all dimensions of the thumbtack surface are indicated in Fig. 4.

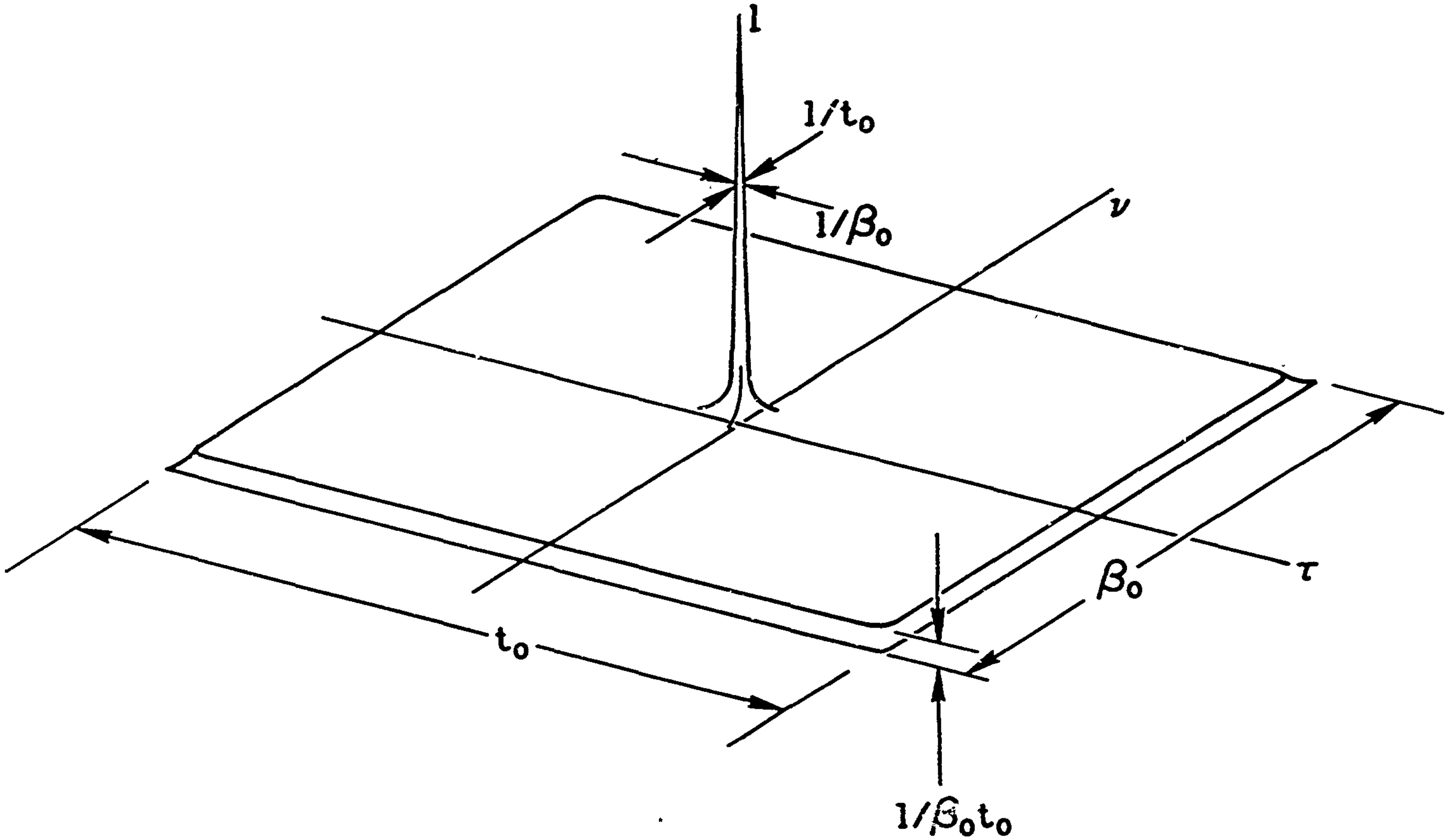

Fig. 4. Thumbtack Ambiguity Surface and its Constraints

As will be shown later, the pedestal of the thumbtack surface can lead to a serious degradation of resolution performance even though it is relatively low for high time-bandwidth product signals. In those cases where the occupied target space has a small size, it may be preferable to use an ambiguity surface where the volume from around the central spike is shifted to other places in the τ, ν-plane. No interfering self-clutter will then be generated as long as the cleared area is at least twice as wide as the occupied target space in both delay and Doppler domain. This raises the question as to how large an area can be cleared.

Along either axis, the ambiguity surface is determined by the autocorrelation function of the signal in the respective domain. Since only signals having pulse train structure can have an autocorrelation function that is zero over an extended interval, it is clear that the desired ambiguity surface can be achieved only with pulse trains in the time domain or "pulse trains" in the frequency domain. Choosing the former for purposes of illustration, uniform repetition of an arbitrary signal at a repetition period T will deposit "volume" at periodic intervals on the delay axis, as indicated by the marks in Fig. 5. Because of Eq. (12), most of the volume in the τ, ν-plane must then be concentrated in strips parallel to the delay axis and repeated at intervals $1/T$. If, in this process, volume is deposited on the Doppler axis, it must fall within the horizontal strips, and through Eq. (13) then confines the volume to the vertical strips in Fig. 5. The result is the well-known ambiguity surface of the uniform pulse train, with spikes of unity height at the intersecting points of the grid. As shown by the rectangle in Fig. 5, the maximum size of the occupied target space which the pulse train can accommodate without introducing serious self-clutter is unity.

It can be shown that the size of the clear area cannot be increased through either intra-pulse or inter-pulse modulation of the pulse train. As is seen from Fig. 3c, modulation with small phase constant α achieves narrowing of the response peak through smearing of the undesired volume into a pedestal, thus introducing self-clutter. High-α modulation translates ridges

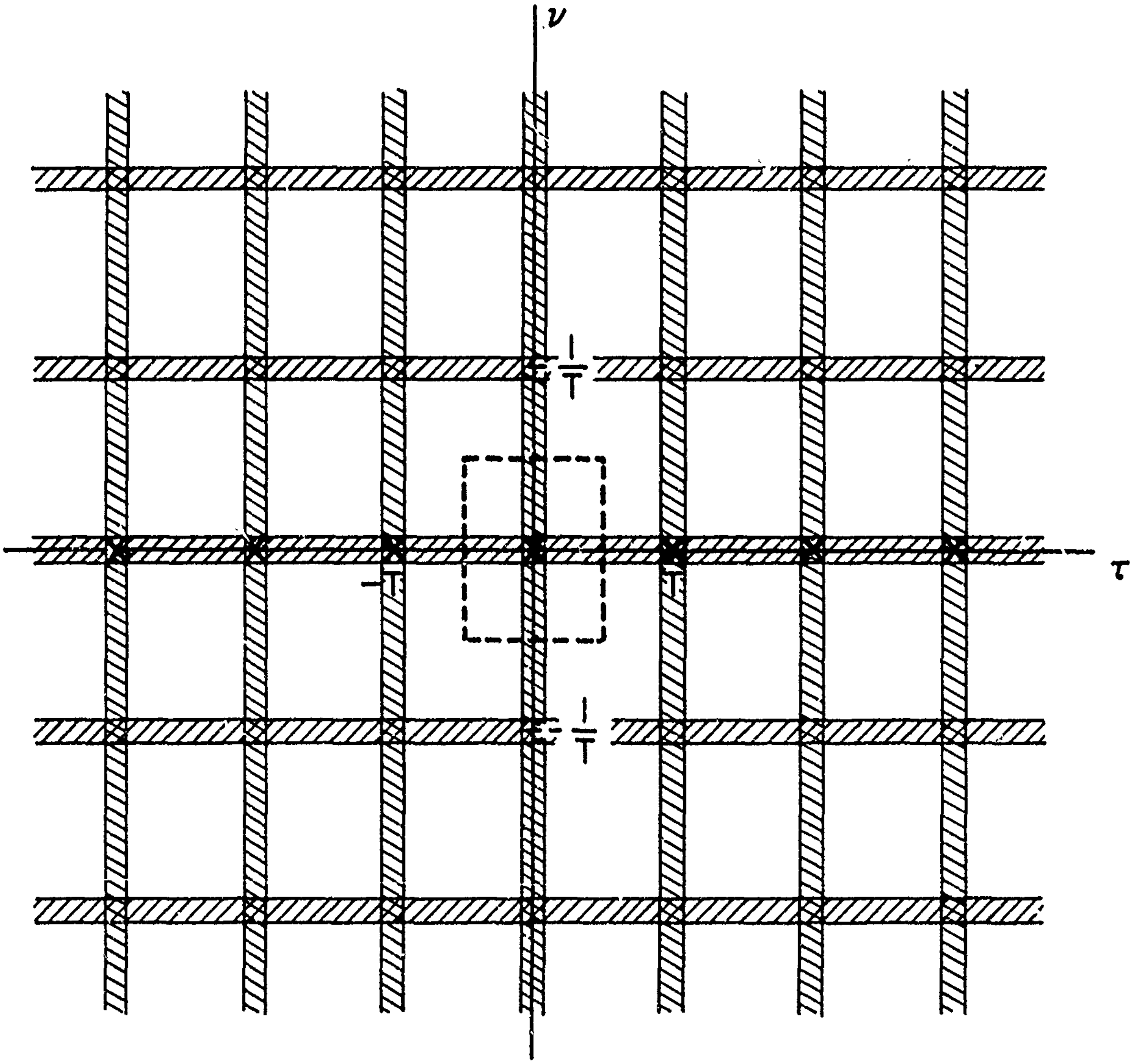

Fig. 5. The "Clear Area" of Signals with the Structure of Pulse Trains

or spikes in the manner of Fig. 3b and merely distorts the shape of the clear area without altering its size. This behavior of pulse trains is studied in more detail in Ref. 6.

It should be mentioned that the "clear area" cannot be strictly free from all volume. As a consequence of the definition of time and frequency as a pair of Fourier transforms, no signal can have extended zero intervals in both time and frequency domain. Thus there will always be some volume left between the individual spikes of the ambiguity surface; for example, a pulse train in time will have some volume spread out within the vertical strips of Fig. 5. In some cases, the interference from such secondary response lobes can be very objectionable, yet these effects are second order compared to those from the full pedestal.

C. APPROXIMATION OF THE THUMBTACK SURFACE THROUGH PRACTICAL WAVEFORMS

In the preceding section, we have found that when resolution in only one coordinate, range or range rate, is needed, the signal design is dictated only by practical considerations. When the occupied target space is smaller than unity, the proper choice for the signal is the pulse train, where this term includes a signal consisting of the periodic repetition of frequency bands. In the general case where the illuminated target space is large compared to unity, however, the radar should be using a signal having a thumbtack ambiguity surface. We shall now add some remarks about waveforms which approximate this surface.

It is clear from Fig. 2 that the signal must have a large time-bandwidth product and a phase constant α which approaches zero. As it is not practical in high-power radar to use signals of other than roughly rectangular envelope, the desired increase in the signal bandwidth can be achieved only by FM or PM of the carrier. From the definition of the effective phase constant, Eq. (7), we then find that α is zero whenever the mean of the FM function $\phi'(t)$ is zero. One way to achieve this is to use symmetrical FM, such as the combination of the frequency up-sweep and down-sweep of the simple Chirp signal, or "V-type" Chirp. The corresponding ambiguity surface is approximately indicated by the contours for fixed Doppler shift in Fig. 6a. Since each half of the signal is a Chirp signal and contains a strong linear FM component, the ambiguity surface still shows the pronounced ridges reminiscent of a Chirp signal. For a further suppression of these ridges, we should use symmetrical FM without linear segments, for example, quadratic FM. The corresponding ambiguity surface is sketched in Fig. 6b.

With signals having smooth modulation functions, it is unavoidable that much of the volume is concentrated in the central portion of the ambiguity surface, giving rise to near sidelobes of appreciable magnitude. A more even distribution of the volume can be achieved with modulation functions of random character, such as noise modulation or phase reversal coding in accordance with some pseudo-random binary code. The ambiguity surface then takes the form of Fig. 6c, with a fairly uniform distribution of the volume but occasional subsidiary peaks of appreciable magnitude. The practical difficulty consists in selecting the modulation code such that the interfering spikes stay below a given level.

When the desired range rate resolution is so high that the signal duration exceeds the target spread in range delay, a coherent pulse train must be used. However, the pulse train must employ some kind of pulse-to-pulse coding to smear the ambiguous spikes of the uncoded pulse train into a low pedestal (Ref. 6). Jittering of the repetition interval, changing the waveform from pulse to pulse, or pseudo-random frequency hopping are some codes which produce the desired smearing effect. Codes with high phase constant α, such as linear frequency hopping from pulse to pulse, must of course be ruled out since they result in a translation of the volume without any leveling of the subsidiary spikes.

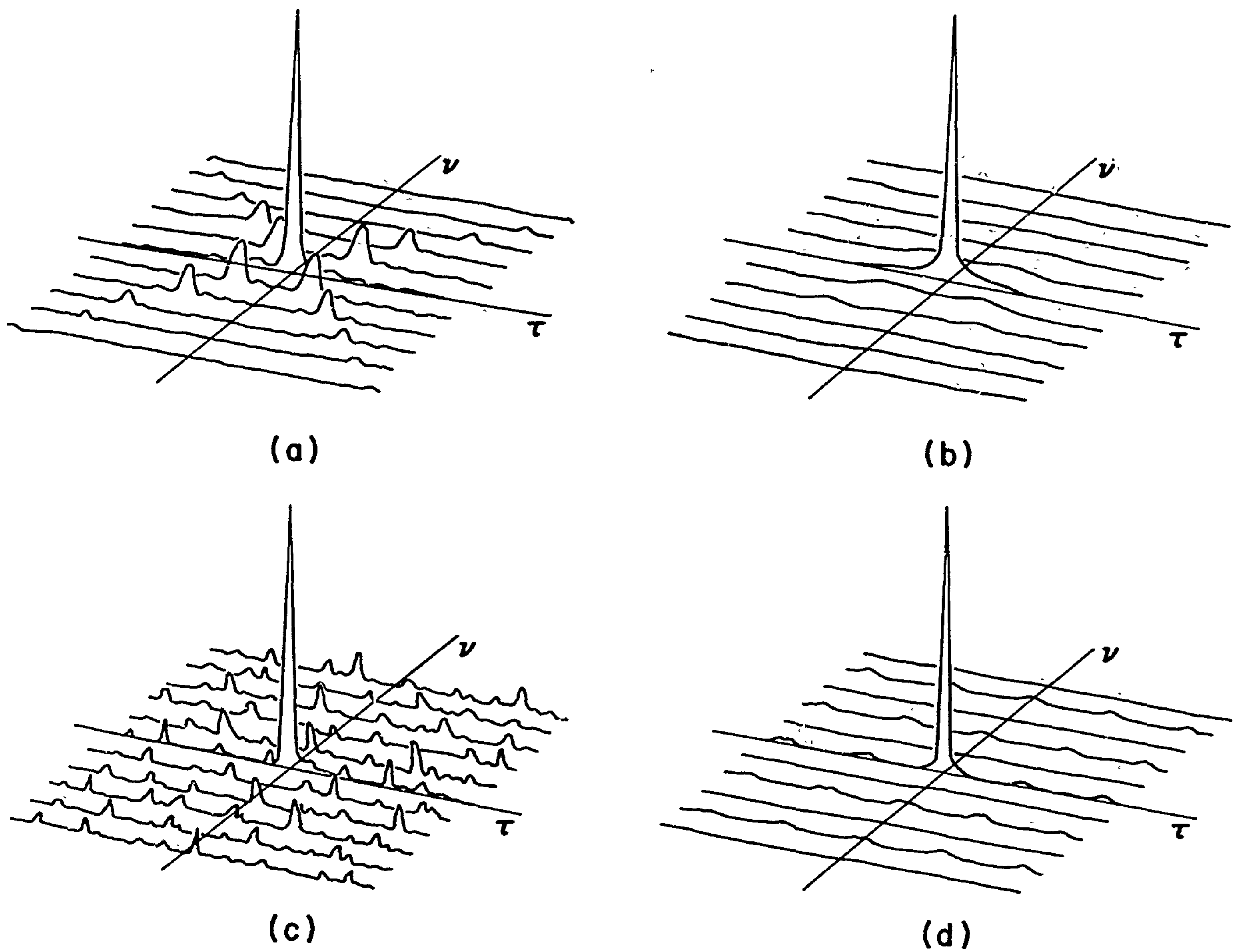

Fig. 6. Representative Examples for Practical Approximations of the Thumbtack Surface. (a) V-type Chirp (b) Quadratic FM (c) Phase-Reversal Coding (d) Pulse Train with Pulse-to-Pulse Coding

IV. FUNDAMENTAL LIMITATIONS ON RADAR RESOLUTION

It was found earlier that in the attempt to decrease the width of the central spike of the ambiguity surface and improve proximal target resolution, a background of self-clutter is generated which in turn may prevent the recognition of targets altogether. We shall now demonstrate that this self-clutter imposes a fundamental limit on the resolution performance of radar.

Assume that the targets illuminated by the radar beam are constrained to an interval $\Delta\tau$ in range delay and an interval $\Delta\nu$ in Doppler, and the size of this (occupied) target space, $\Delta\tau\Delta\nu$, is large compared to unity. The optimum ambiguity surface under these conditions is the thumbtack surface. When only a few targets are present, the amount of target masking obtained with a given signal waveform can be readily estimated from the ambiguity surface of the particular waveform. When many targets are present, and this is where the limitations on resolution become apparent, the superposition of the many pedestals from returns having different delays and Doppler shifts will have an averaging effect on the combined pedestal so that its detail structure is of little interest. Hence we can estimate radar resolution performance on the basis of an ideal thumbtack surface and, for the present purpose, ignore the problem of approximating the ideal thumbtack surface by a practical waveform.

The relation between the size of the occupied target space and the spread of the pedestal is depicted in Fig. 7. As the first case, we assume that the target space contains N discrete point scatterers of a total cross section σ_t and an average cross section $\sigma_0 = \sigma_t/N$. Furthermore, N will be assumed large enough to allow treating the superposition of the pedestals as a power superposition, that is, superposing the pedestals of the ambiguity surfaces rather than ambiguity functions. (This actually corresponds to the calculation of the variance of the combined clutter which, in the case of many contributions, will very nearly have a Gaussian distribution, so that target recognition in the clutter becomes the well-known problem of signal detection in Gaussian noise.)

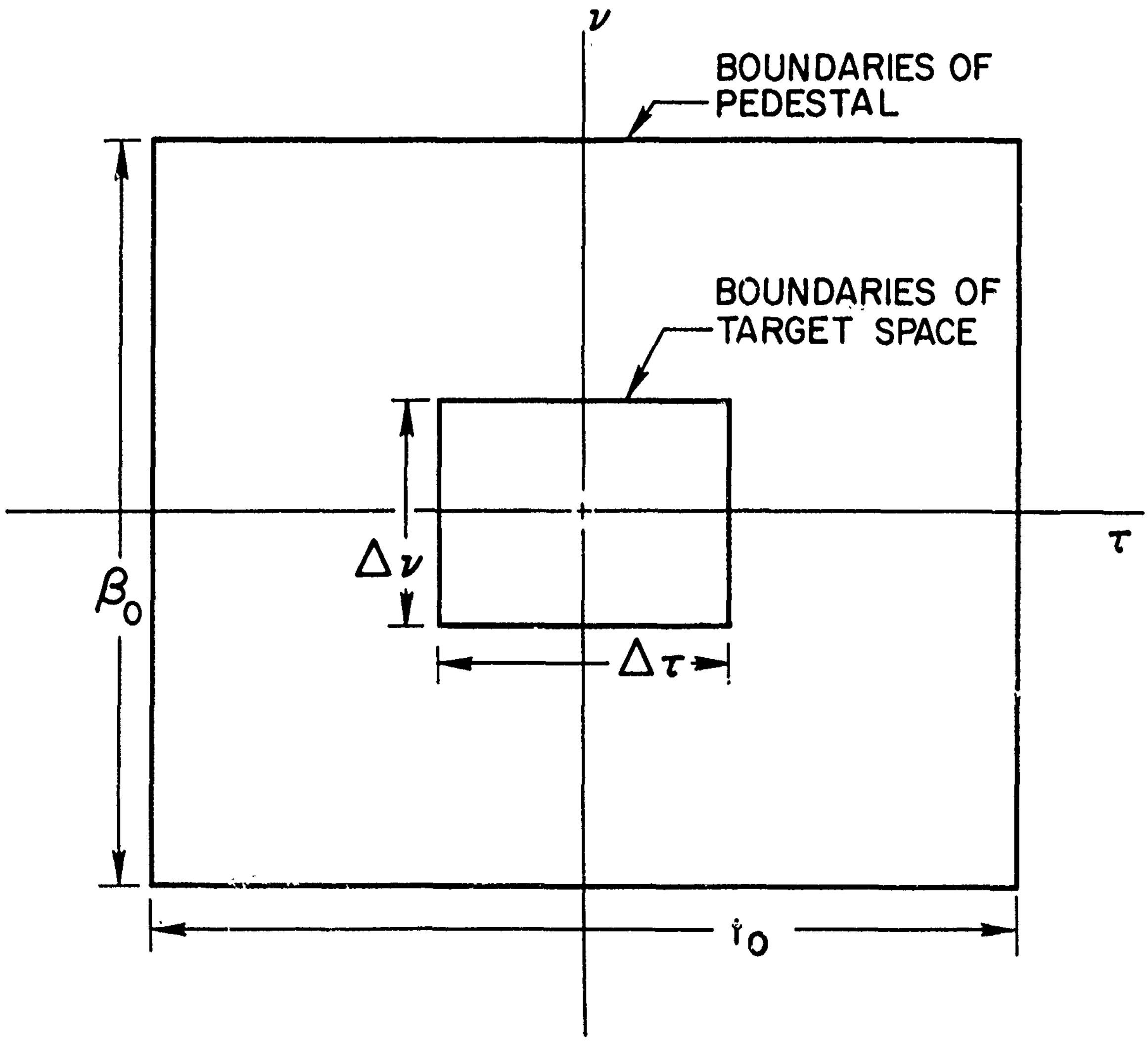

Fig. 7. Occupied Target Space and Spread of the Pedestal of the Ambiguity Surface

Since the peak of the normalized thumbtack surface has a height of unity and a pedestal of height $1/\beta_0 t_0$, proper amplitude scaling of each return signal requires that the heights of both peak and pedestal be multiplied by the radar cross section of the corresponding target. Superposing the scaled ambiguity surfaces from all N targets, the height of the combined pedestal becomes

$$H_p = \sum_1^N \frac{\sigma_\nu}{\beta_0 t_0} = \frac{\sigma_t}{\beta_0 t_0} = N \frac{\sigma_0}{\beta_0 t_0} \quad . \tag{15}$$

For a target of cross section σ_ν, the signal-to-clutter ratio thus is

$$\frac{S_\nu}{C} = \frac{\sigma_\nu}{H_p} = \beta_0 t_0 \frac{1}{N} \frac{\sigma_\nu}{\sigma_0} \quad , \tag{16}$$

and for a target of average cross section,

$$\frac{S_0}{C} = \frac{\beta_0 t_0}{N} \quad . \tag{17}$$

The value of the signal-to-clutter ratio needed for proper target identification in the background will depend on the requirements of the particular application. In most instances, a minimum signal-to-clutter ratio in the order of 6 db might be typical. However, in order to avoid obscuring the essence of the discussion by numerical factors, we shall here assume a minimum required signal-to-clutter ratio of unity. Equation (17) then shows that the maximum number of targets which can be accommodated if a target of average cross section is to be resolved is given by the time-bandwidth product of the signal,

$$N_{max} = \beta_0 t_0 \quad . \tag{18}$$

Provided the area of the pedestal is larger than the occupied target space, as in Fig. 7, the target capacity of the radar can be increased by increasing the time-bandwidth product of the signal. It is seen from Eq. (16) that if a target of lower than average cross section is to be resolved, the target capacity of the radar decreases accordingly. Specifically, a target m db below average can be resolved only if the number N of targets is m db below N_{max} as given by Eq. (18). As an example, if 100 targets are present and targets with a cross section 20 db below average are to be resolved, the transmitted signal must have a time-bandwidth product of 10,000, a rather expensive requirement for a practical radar.

In some applications, the relation between size of the occupied target space and area of the pedestal may be the reverse from that of Fig. 7, and it is of interest to consider resolution when $\Delta\tau\Delta\nu$ is large compared to $\beta_0 t_0$. Now it is preferable to introduce the average spacing of the targets in delay and Doppler, ξ_τ and ξ_ν, rather than the total number of targets in target space. The number of contributors to the combined pedestal is given as the number of targets within the area of the single pedestal, or

$$N = \frac{\beta_0 t_0}{\xi_\tau \xi_\nu} \quad , \tag{19}$$

and the signal-to-clutter ratio for a target of cross section σ_ν is

$$\frac{S_\nu}{C} = \xi_\tau \xi_\nu \frac{\sigma_\nu}{\sigma_0} \quad . \tag{20}$$

As shown by the preceding result, target resolvability in the clutter now depends only on the target strength and the average target density, but not on the time-bandwidth product of the signal. Though increasing the time-bandwidth may improve proximal target resolution for targets of sufficient strength, it does not improve resolvability in the self-clutter. For a target of average cross section, setting the signal-to-clutter ratio in Eq. (20) equal to unity yields

$$\xi_\tau \xi_\nu = 1 \quad . \tag{21}$$

This means that a target of average cross section can be resolved only if the target density is no higher than one target per unit area in target space. In other words, the size of the resolvable cell cannot be made smaller than unity, which is the resolvable cell that can be provided by a simple constant-carrier pulse. If the target density is higher than one target per unit area, only targets proportionately stronger than the average target can be resolved, with a strict correspondence between the increase in target density and increase in the strength of resolvable targets.

The situation is similar for continuously distributed targets. We analyze this case by assuming that the target cross section in each cell of dimensions $1/\beta_0 \times 1/t_0$, the area of the central spike of the ambiguity surface, is concentrated into a single point target. The task of the radar then is to measure the cross section of each such target and obtain a "map" of the cross section distribution in target space. Of course, the self-clutter will interfere with the measurement, and the cell size will have to be large enough to ensure that the cross section per cell, or the equivalent point target, is so large that it can be recognized in the clutter. We again have the trade-off between cell size and adequate target strength.

When the occupied target space is large compared to unity, the thumbtack ambiguity surface is again optimum. The relation between target space and pedestal is as shown in Fig. 7, and the number of scatterers contributing to the clutter background equals the number of cells of size $1/\beta_0 t_0$ within the occupied target space, or

$$N = \frac{\Delta\tau\Delta\nu}{1/\beta_0 t_0} \quad . \tag{22}$$

The signal-to-clutter ratio for a cell containing the target cross section σ_ν is, from Eq. (15),

$$\frac{S_\nu}{C} = \frac{1}{\Delta\tau\Delta\nu}\frac{\sigma_\nu}{\sigma_0} = \frac{1}{\Delta\tau\Delta\nu}\frac{\sigma_\nu^*}{\sigma_0^*} \quad , \tag{23}$$

where the asterisk is used to designate cross section densities in target space. It is seen that an average cross section density can be properly mapped only if the total occupied target space is unity in size or smaller. At those portions in target space where the cross section level is m db below average, correct mapping (or target "resolution") is possible only when the occupied target space has a size m db below unity. Conversely, for target spaces larger than unity, mapping of the distribution of cross section is feasible only where the cross section density is proportionately higher than average. For a given target space, changing the time-bandwidth product of the signal thus merely permits trading the size of the resolution cell against the cross section level for which a resolvable cell of size $1/\beta_0 t_0$ can be truly provided in the self-clutter.

Considering the other case of interest, a target space large compared to the area of the pedestal, the number of cells contributing to the background of self-clutter is

$$N = \frac{\beta_0 t_0}{1/\beta_0 t_0} = (\beta_0 t_0)^2 \quad , \tag{24}$$

and the signal-to-clutter ratio for a cell with cross section σ_ν is

$$\frac{S_\nu}{C} = \frac{1}{\beta_0 t_0}\frac{\sigma_\nu}{\sigma_0} = \frac{1}{\beta_0 t_0}\frac{\sigma_\nu^*}{\sigma_0^*} \quad . \tag{25}$$

A resolution commensurate with the area of the central spike of the ambiguity surface, or a cell of dimensions $1/\beta_0 \times 1/t_0$, can evidently only be achieved in those portions of target space where the cross section density is higher than average by a factor $\beta_0 t_0$ or more. Where the cross section density is of average strength, the truly resolvable cell is unity in size, and when the cross section density is still lower, the size of the resolvable cell is proportionately larger than unity.

The preceding results have revealed very fundamental limitations on the combined range and range rate resolution achievable with radar when the occupied target space is large compared to unity. When it is equal to unity or smaller, these limitations can be avoided by using pulse train type signals to clear the area around the central spike of the ambiguity surface and achieve an arbitrarily high resolution, except for the effects of the residual volume in the "clear" area. As the occupied target space exceeds unity by a larger and larger factor, the benefits from being able to clear an area of unity size will become of less value. Furthermore, we shall have to smear the ambiguous spikes of the ambiguity surface of the uniform pulse train by pulse-to-pulse coding, and in this process approximate the thumbtack surface. Hence, this type of ambiguity surface must be used whenever the size of the occupied target space significantly exceeds unity, and we then have the limitations on resolution as discussed above. Of course, in those simple cases where resolution in only one coordinate is needed, range or range rate, the limitations on achievable resolution are only of practical nature, such as the difficulties of realizing a large time-bandwidth product or reducing the spurious responses to the desired low level.

V. PRACTICAL IMPLICATIONS OF THE CONSTRAINTS ON RESOLUTION

It is instructive to illustrate the practical significance of the preceding results by considering two typical applications for high-resolution radar. As the first example, we choose the synthetic aperture ground mapping radar, a system whose task it is to provide a map of the continuous cross section distribution on the ground. As a second example, typical for resolution of high-density discrete targets, we consider a radar operating against dense-target clouds such as aircraft or missiles accompanied by chaff and decoys.

The geometry of synthetic aperture radar is shown in Fig. 8. Briefly summarizing its operation (Ref. 7), the radar flies along a straight path with velocity v and periodically illuminates a ground swath of width M. The return signals are stored and then processed such as if they had been received by a long array antenna with a correspondingly narrow beam in azimuth. As an alternative and, for our purposes, more interesting interpretation, the radar uses coherent processing of a very long signal to achieve high resolution in Doppler which, because of the known relative motion of radar and target, can be converted into high resolution in azimuth. Range resolution is obtained as in any conventional high-resolution radar.

Referring to previous discussions, in a ground mapping radar the targets are continuously distributed over large intervals in range and Doppler. The occupied target space in the sense defined above is that portion of the total target space which is actually illuminated by the radar beam. Having found that proper mapping is achievable only when the occupied target space does not exceed unity, we can immediately conclude that the antenna aperture must be large enough to limit the size of the illuminated target space to unity, in which case the interfering clutter background is avoided. Then there will be no theoretical limit on achievable resolution, but the constraint on the size of the occupied target space now imposes a limit on the size of the map that can be obtained.

Preceding page blank

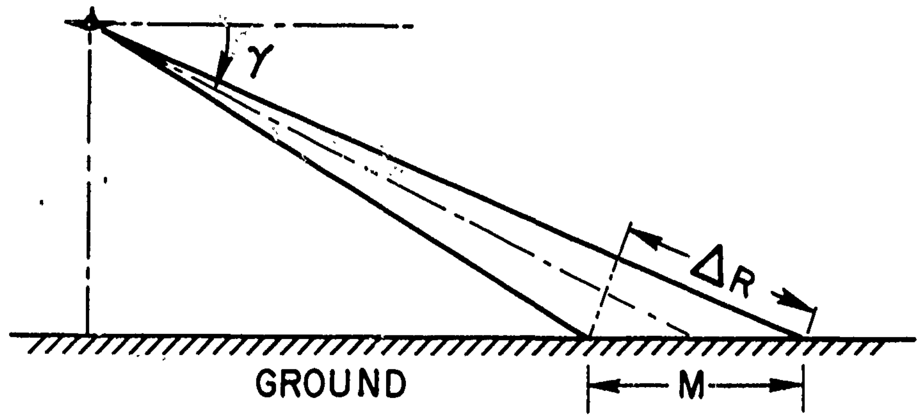

Fig. 8. Geometry for Synthetic Aperture Radar Operation

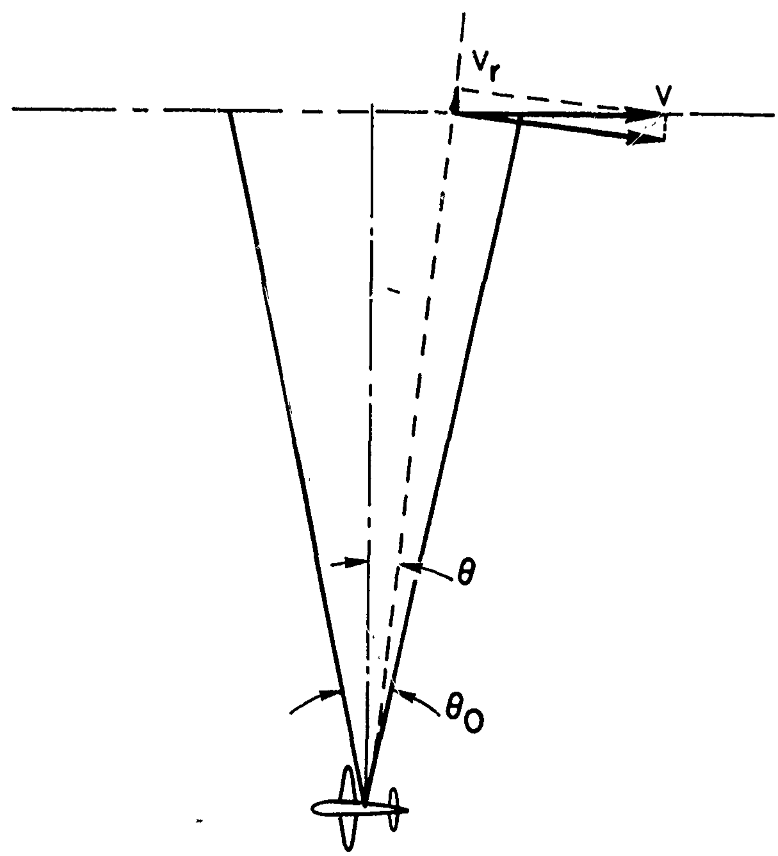

Fig. 9. Derivation of the Target Doppler Spread

Converting the limit on the occupied target space into a limit on map size, the spread of the targets in range delay is found from Fig. 8 as

$$\Delta\tau = \frac{2\Delta R}{c} = \frac{2M \cos \gamma}{c} \quad . \tag{26}$$

The target spread in Doppler can be calculated from the relative target motion as indicated in Fig. 9. For a target at angular position θ, the Doppler shift is

$$\nu = 2\frac{v_r}{\lambda} = 2\frac{v}{\lambda} \sin \theta \approx 2\frac{v}{\lambda}\theta \quad , \tag{27}$$

where the approximation is good because the antenna beam is rather narrow in practice. (Though target acceleration is not negligible in synthetic aperture applications, it is predictable from target range and hence can be taken out in data processing.) The total Doppler spread of the targets within the beam at any one time is determined by the azimuth width θ_0 of the beam as (Eq. (27))

$$\Delta\nu = 2\frac{v}{\lambda}\theta_0 \quad , \tag{28}$$

and the unity constraint on the occupied target space becomes

$$\Delta\tau\Delta\nu = 2\frac{v}{\lambda}\theta_0 \frac{2M \cos \gamma}{c} = 1 \quad . \tag{29}$$

From Eq. (27), the relation between resolution in Doppler and resolution in azimuth angle is

$$\delta\nu = \frac{2v}{\lambda}\delta\theta \quad . \tag{30}$$

We now replace the angular resolution $\delta\theta$ by the lateral resolution in azimuth at target range R, $a = R\delta\theta$. Furthermore, Doppler resolution is in the order of the inverse signal duration, and this duration is given by the time it takes the target to traverse the antenna beam; hence

$$\delta\nu = \frac{1}{T} = \frac{2v}{\lambda}\frac{a}{R} \quad , \tag{31}$$

and with $T = \theta_0 R/v$, Eq. (31) and the constraint of Eq. (29) lead to

$$M = \frac{1}{2\cos\gamma} a \frac{c}{v} \quad . \tag{32}$$

This result states that the self-clutter in high-resolution radar limits the achievable swath width. For a given vehicle velocity v, the swath width decreases proportionately to the width of the azimuth resolution element, a. As an ancillary result, for a given resolution in azimuth, the achievable swath width is inversely proportional to the velocity of the vehicle carrying the radar, which means that low-velocity vehicles can map wider swaths than high-velocity vehicles. This is an expression of the fact that the size of the occupied target space in range delay and Doppler depends on the vehicle velocity through the relation between Doppler frequency and velocity as given in Eq. (27).

For the second example, a radar operating against a target "cloud", the resolution performance again depends on the size of the target space illuminated by the beam. When the occupied target space is unity in size or smaller, coherent pulse trains can be used to produce an ambiguity surface with a clear area around the central spike, and resolution is limited only by the residual volume in the clear area. In practice, this leads to a limitation on the size of weak targets that can be seen in the presence of strong targets and becomes a matter of successful system design. On the other hand, when the occupied target space is large compared to unity (and this can be made deliberately so in weapon system design), the problem is much more fundamental. This case will be considered below.

For large target spaces, the optimum ambiguity surface is the thumbtack surface, and for high resolution requirements the sketch of Fig. 7 applies. Using Eq. (16), the signal-to-clutter ratio for a target of cross section σ_ν can be written

$$\frac{S_\nu}{C} = \beta_0 t_0 \frac{\sigma_\nu}{\sigma_t} \quad , \tag{33}$$

where σ_t is the total target cross section within the target space illuminated by the radar beam. In order for a target to be recognizable in the clutter background, its cross section must evidently be at least as large as the fraction $1/\beta_0 t_0$ of the total cross section σ_t. Although weak target resolution can, in theory, be improved by choosing higher time-bandwidth products, there exist definite practical limitations on how high a time-bandwidth product an operational radar can employ. Furthermore, after a certain point, the decrease in the average pedestal height of the ambiguity surface for larger time-bandwidth product becomes meaningless, as it will be progressively more difficult to keep the subsidiary peaks above the average pedestal level low enough to prevent weak target masking by such individual peaks rather than the average clutter level.

The difficulties of dense-target resolution are even more serious than suggested above. The lowering of the clutter level by increasing the time-bandwidth product of the signal is, of course, paid for by a corresponding spreading of the pedestal, as indicated in Fig. 10. This shows that the signal-to-clutter ratio of Eq. (33) exists not only for targets within the densely occupied target space, but over an area which can be very much larger. In other words, if the signal-to-clutter ratio for a target within the target cloud is insufficient for proper target resolution, it will remain insufficient even if the target moves outside the cloud. Practically, the target of interest may have appreciably different range and range rate than the bulk of the targets and still be effectively hidden. The penalty paid for

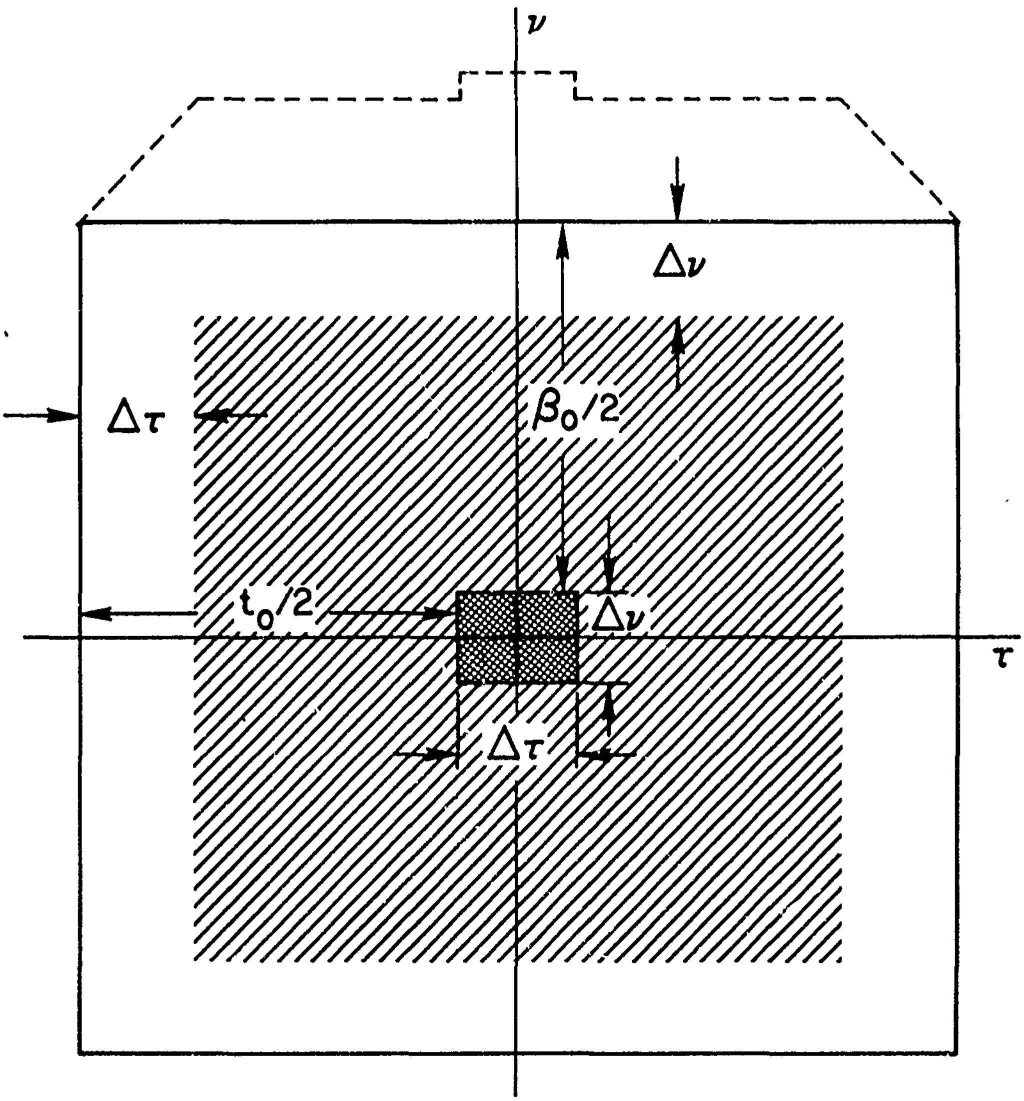

Fig. 10. Spreading of the Clutter in Target Space for "High-Resolution" Signals

decreasing the size of the resolution cell is a corresponding spreading of the masking effects of the target cloud beyond its boundaries in range and range rate. Again, these effects trace back to the radar uncertainty relation, and their severity can be easily appreciated by substituting practical numbers in the relations derived above.

VI. CONCLUDING REMARKS

Modern radar offers the possibility of significantly reducing the size of the resolvable cell in target space from the value of unity achieved with constant-carrier, single-pulse signals. However, aside from the practical difficulties of implementing such a radar, increasing the time-bandwidth product of the transmitted signal introduces self-clutter which itself limits radar resolution performance. It was shown that, as a consequence of the radar uncertainty relation, arbitrarily high resolution can even in theory be achieved only when the total occupied target space does not exceed unity in size, in which case pulse train type signals must be used to generate an ambiguity surface with a clear area around the central spike. For large occupied target spaces, and given characteristics of the target environment, signal design merely permits trading the size of the resolvable cell against the target strength for which this resolvable cell can actually be provided. Two practical examples were given in order to illustrate that these limitations are not merely of theoretical interest but do lead to very practical constraints on the resolution performance of even the most sophisticated radar.

REFERENCES

1. P. M. Woodward, "Probability and Information Theory, with Applications to Radar", Pergamon Press, New York (1953).

2. W. M. Siebert, "A Radar Detection Philosophy", IRE Trans. on Information Theory, Vol. IT-2, p. 204 (Sept. 1956).

3. E. C. Westerfield, R. H. Prager, and J. L. Stewart, "Processing Gains against Reverberation (Clutter) Using Matched Filters", IRE Trans. on Information Theory, Vol. IT-6, p. 342 (June 1960).

4. E. N. Fowle, E. J. Kelly, and J. A. Sheehan, "Radar System Performance in a Dense-Target Environment", 1961 IRE Internat. Conv. Rec., Part 4, p. 136.

5. D. Gabor, "Theory of Communication", J. Inst. Electr. Engrs. Vol. 93, Part III, p. 429 (1946).

6. A. W. Rihaczek, "Radar Resolution Properties of Pulse Trains", Proc. IEEE, Vol. 52 (1964), p. 153.

7. L. J. Cutrona, W. E. Vivian, E. N. Leith, and G. O. Hall, "A High-Resolution Radar Combat-Surveillance System", IRE Trans. on Military Electronics, Vol. MIL-5, p. 127 (April 1961).