Mathematics of Radar

Bill Moran

1. Radar Fundamentals

1.1. *Introduction*

Radar is now used in many applications — meteorology, mapping, air traffic control, ship and aircraft navigation, altimeters on aircraft, police speeding control, etc. It is now being used in the form of ground penetrating radar for mineral exploration and delineation and for land-mine detection. Of course, its primary role is in defence. The theory of radar is well developed and has many interesting and difficult mathematical problems.

The aim of these notes is to provide a description of radar and its theory accessible to a mathematical audience with the hope of stimulating interest in the problems. Often mathematical treatments of radar ignore the problems of implementation. I intend to go as far into the engneering of a radar system as is necesssary to provide an understanding of these issues. In the first section we shall describe the operation of a radar system at a relatively detailed level. Later sections will cover some of the mathematics arising in radar design and use.

1.2. *A Simple Radar System*

The key references for this section are [3, 14, 16]. Typically radar systems comprise a transmitter and receiver, though recently passive radar systems relying on the ambient HF and VHF radiation have been built and made operational. We shall not consider such systems. We shall also assume that the radar transmitter and receiver are collocated (*monostatic*), again the usual situation. Nonetheless there is much current interest in *bistatic* and *multistatic* radars

The most important issues in designing a radar system are as follows.

1. What is its purpose? How it will be used significantly changes the method of processing the data. Some radars are used to detect targets, others to track them, others to produce images. Radars used to check the speeds of vehicles, for example, need to be able to measure the doppler accurately in a small range of velocities.
2. The next most serious issue is that of noise. Returns from relatively distant targets have very small amounts of energy which can be swamped by the noise generated

within the receiver. Methods of maximizing the post processing signal-to-noise ratio are crucially important in radar systems.

3. As well as noise, a radar system will receive returns from objects which are not important for the operational purpose of the radar. For example, shipborn radars, typically receive much of the return energy from the sea surface, particularly if it is rough. This can mask a more important return. Such artefacts are collectively called "clutter". Much work has been done on extraction of signals of importance from the clutter.
4. The power available to the transmitter is a key factor, particularly in reducing noise problems. Evidently this dictates much about the size of the transmitter and the type of electronics to be used.
5. The resolution needed in both doppler and range and the range of values of these parameters for the targets of importance are significant factors in the design of a radar.
6. Of increasing importance in a defence environment are issues around jamming and other electronic counter-measures to obscure or mask a target, or otherwise render a radar system ineffective. Radar system developers have to be ready to counter such interventions.

A simple radar system is sketched in Figure 1.
At a general level, the various parts of a radar system are:

1. The antenna — usually serving the dual role of transmitting and receiving signals. The shape of the antenna dictates the shape of the beam that is transmitted. Focussing allows radar systems to put power where it is most needed. Inevitably all beams have some sidelobes. Shaping of these is an interesting mathematical problem which we do not have time to cover here. Many modern radars use phased array antennas which comprise many small antennas each of which can be assigned a separate (complex) weighting. Such a system allows the radar beam to be steered electronically rather than mechanically as is the case for conventional parabolic dish antennas.
2. The duplexer — this switches the antenna between transmission and reception. Both cannot happen at the same time usually. Typicallya radar will transmit for no more than 10% of time. During most of the remainder of the time it will be in receive mode.
3. The radio frequency (RF) oscillator — provides the carrier signal for the radar. This will usually be at least 10 times the bandwidth of the (intermediate frequency coded) waveform it carries.
4. The RF mixers — one in each of the transmitter and receiver. These serve to put the waveform onto the carrier signal in the transmitter and to remove it in the receiver by mixing it with the carrier and then passing it to the low pass filters.
5. Filters and amplifiers — these amplify the various signals and filter out unwanted frequencies. Filtering is a significant operation for the removal of noise. Noise is typically "broadband" in that its spectrum is wide. By filtering out many regions of the spectrum such noise is reduced. Amplifiers increase the power in the signals (and

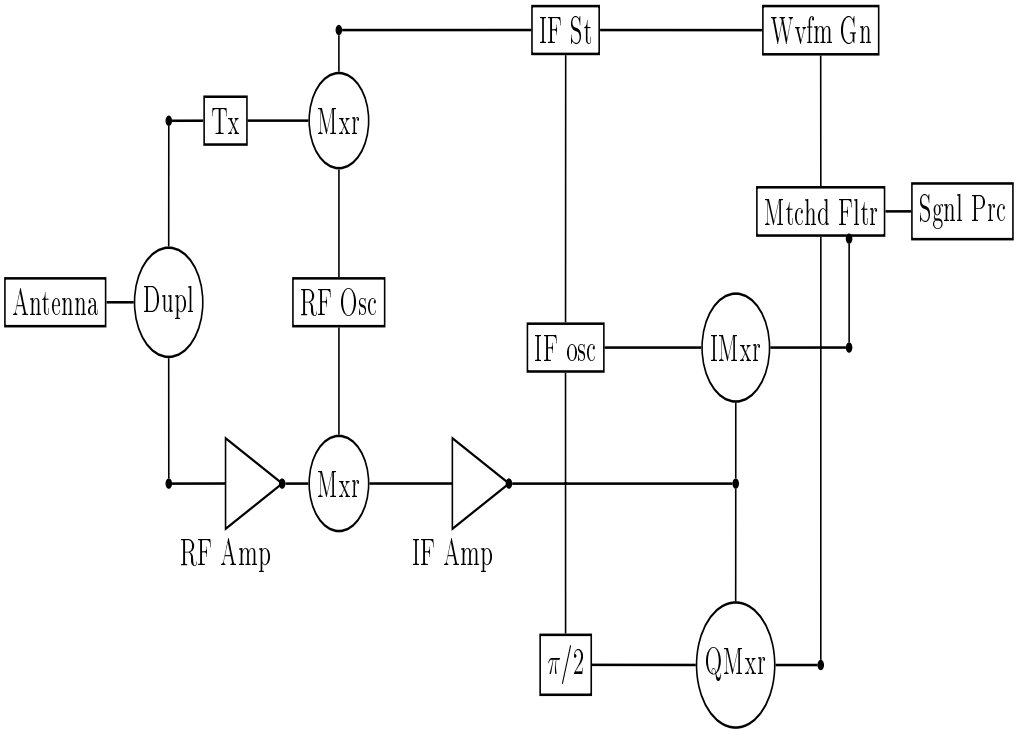

FIGURE 1. A simple radar system

the noise which is present) to make them less susceptible to noise in the later parts of the receiver. Filtering also serves to remove high frequency components after mixing.

6. Components corresponding to the ones in 3, 4, 5 for the intermediate frequency (IF) stages of the system; in particular the demodulation mixers and filters which produce

the I and Q channels. These are obtained by mixing with two IF sinusoids $\pi/2$ radians out of phase with each other.

7. Waveform generator — produces the desired waveform for modulation onto the IF carrier and then onto the RF carrier for transmission. The waveform is also used to " matched filter" the return. We shall spend some time discussing waveform design in section 3.
8. Signal processing block — this serves to detect, identify, estimate the range and/or doppler of targets, to remove clutter and noise not removed earlier in the processing, etc. As analogue to digital (A/D) converters become faster and more accurate, signal processing is increasingly accomplished by converting the signals to digital form and then using digital circuitry and even software. In fact technology already at the (low) IF stage and I/Q channel separation can be accomplished digitally.

1.3. *Doppler and Range*

We begin here to build the mathematical model of radar processing, by and large following Borden's treatment [3]. Usually a transmitted signal comprises a slowly varying *waveform* superimposed on a rapidly oscillating sinusoid. Thus

$$\mathrm{s}(t) = \mathrm{w}(t).\cos(2\pi(f_c t + \phi(t)) \tag{1.1}$$

where $\mathrm{w}(t)$ is the amplitude modulation waveform, $\phi(t)$ represents the frequency or phase modulation and f_c is the carrier frequency. It is important to make the rather obvious observation at this stage that all signals transmitted and received are real-valued. Of course we can represent the signal as the result of amplitude modulating two sinusoids with *opposite phase* (that is, differing by $\pi/2$), thus

$$\mathrm{s}(t) = \mathrm{w}(t).\cos(2\pi\phi(t))cos(2\pi f_c t) - \mathrm{w}(t)\sin(2\pi\phi(t))\sin(2\pi f_c t),$$

which in *complex form* is

$$\mathrm{s}(t) = \Re\big(\mathrm{w}(t).e^{2\pi\phi(t)}.e^{2\pi f_c t}\big), \tag{1.2}$$

provided the waveform $\mathrm{w}(t)$ is real. We write

$$\mathrm{s}_0(t) = \mathrm{w}(t)e^{2\pi\phi(t)} \tag{1.3}$$

for the complexification of the slowly varying part.

Much of the theory of radar processing takes place in the complex domain as we shall see later. To give an idea of the relative variability of the two components of (1.1), the carrier frequency will often be in the range 1–10GHz ($= 1 - 10 \times 10^9$ cycles per second). The variability of the waveform of course depends on its form. Often phase coded pulses are used which, in theory, switch instantaneously, but limitations in the practical implementation usually restrict the highest frequency components of the waveform to be no more than $1/10$ of the carrier frequency.

In the complex domain we can write the signal as

$$\mathrm{s}(t) = \Re\big(\mathrm{w}(t).e^{2\pi i\phi(t)}.e^{2\pi i f_c t}\big) = \Re\big(\mathrm{s}_0(t).e^{2\pi i f_c t}\big) \tag{1.4}$$

The transmitted signal hits a *target* whose distance from the (collocated) transmitter and receiver is R. Let us assume for the moment that the target is stationary relative to the radar system and that it comprises a single scatterer. Then the return signal will, to a good approximation (see Section [deleted -Ed.] for a more accurate view of radar scattering), be a delayed version of the original signal which depends on the range R of the target. Specifically the signal voltage at the antenna of the receiver is

$$\mathrm{s}_u(t) = A\mathrm{s}(t - \frac{2R}{c}) \tag{1.5}$$

where c is the speed of light and A represents the attentuation of the signal by the reflection. As I have already said, this is an approximation to the true situation, but a good one for most purposes. In fact, even if the target is totally unchanging the transmitted waveform really excites the electrons within the material of the target which then reradiate a modulated version of the signal. This effect is slight, but its use is being developed for the purposes of target identification. More importantly the phase of the waveform will change on reflection from the target. This can be accommodated within the complex form of the signal (that is, the complex signal whose real part is given in (1.2)) by allowing A to be complex. Otherwise an extra phase shift is imposed on the carrier.

When the waveform returns to the receiver some noise is added ("receiver noise") which arises from the thermal activity generated within the components of the receiver. It is important to realise here that the fall-off of the signal strength (ie power) as a function of distance of the target is as the fourth power of the range R, so that typically the received signal power is very small (for a distant target 10^{-18} watts is not unusual — there is a saying among radar engineers that the total signal energy received by all radars ever is not enough to turn a page) and in this context receiver noise can be significant. To a good approximation, receiver noise is Gaussian. That is we can write

$$\mathrm{s}_r(t) = \mathrm{s}_u(t) + N(t),$$

where $N(t)$ is a Gaussian process, for the signal after the initial stages of the receiver.

I have mentioned the issue of noise several times and its importance cannot be overemphasised. At any moment it will often be the most significant part of the signal in the system prior to the matched filter. Extraction of signal from noise is a major issue. In radar literature the noise process is assumed typically to be "white Gaussian noise". From a mathematical perspective there are problems with this assumption. Such a continuous time process cannot exist since it would have infinite power. To be more mathematically precise, we assume that the power spectral density of the noise is flat across the interval $[-B, B]$ where B represents the bandwidth of the receiver; that is, the maximum frequency it will handle, as would be the case if the signal had been passed through a (perfect) lowpass filter prior to

further processing. The treatment of noise in this situation involves interesting ideas in the field of mathematical statistics — specifically estimation theory and hypothesis testing, but we do not have time to address them here. Accordingly we refer the reader to [13, 18].

Now we insert the possibility that the target is moving. This has the effect of modifying the transmitted waveform other than just by delay. It imposes the doppler effect on it. If this is done correctly (ie relativistically) it results in a " time dilation" of the return signal, so that, if the target has a radial velocity v, the return signal $s_u(t)$ becomes

$$\mathrm{s}_u(t) = A\mathrm{s}(\alpha t - \frac{2R}{c}),$$

where

$$\alpha = \frac{(1 - \frac{v}{c})}{(1 + \frac{v}{c})}.$$

When v is much smaller than c this is approximated by $\alpha = (1 - 2v/c)$. A further approximation is possible if, as is usually the case, the signal is " narrow band"; that is, if its (Fourier) spectrum is essentially in a range $(f_c - \delta, f_c + \delta)$ where δ is small compared to f_c. For most radar applications this is a reasonable assumption since the signal modulating the carrier will have relatively low bandwidth. In this case, the return signal is approximated by shifting the frequency of the return from a stationary target at the same range by $f_d = (2v/c)f_c$. This is best written in terms of the complex signal

$$\mathrm{s}_u(t) = \Re\big(\mathrm{s}_0(t - \frac{2R}{c}).e^{2\pi i f_c(1-2v/c)(t-\frac{2R}{c})}\big)$$

This equation is the standard one used in most radar calculations.

1.4. *I and Q channels*

As we have already said, processing of radar signals is done largely in the complex domain. Of course the signal transmitted from and returned to the radar antenna is real. However in a very natural way as we have seen (1.2) the transmitted signal is treated as the real part of a complex signal which is simpler to understand. When the return signal is received it is turned into a complex signal — that is, it is made into two real signals. This is usually done in two steps. First the signal is " demodulated" to one whose carrier frequency (the "intermediate frequency" or IF) is much lower and so in practical terms easier to work with. Demodulation is done by mixing (that is, multiplying) the return signal with a signal (pure tone) whose frequency differs from the carrier frequency by the IF. In other words we form

$$\mathrm{s}_m(t) = \mathrm{s}_r(t)cos2\pi(f_c + f_{IF})t$$

where f_{IF} is the IF. The resulting signal is then passed through a low-pass filter which removes all frequencies greater than the carrier frequency and leaves intact signals whose spectrum lies well below the carrier frequency. Typically for an S-band radar (ie $f_c \sim 3 \times 10^9$

GHz) the IF will be around 100MHz or less and may be accomplished by two stages of demodulation. Using the product formula for sine and cosine we see that the resulting signal is approximately

$$\begin{aligned} s_{IF}(t) &= w(t).\cos 2\pi\phi(\tau)\cos 2\pi((f_{IF}-f_d)\tau)) \\ &\quad -w(\tau)\sin 2\pi\phi(\tau)\sin 2\pi(f_{IF}-f_d)\tau), \end{aligned}$$

where $\tau = t - \frac{2R}{c}$. It needs to be realised that when I say "approximately" the degree of approximation in this game is very large. Because of the need to detect very small signals the filters used will often attenuate the power in unwanted frequencies by 10^{-8}.

At this point the conventional radar system will mix the signal with both $\cos 2\pi f_{IF}t$ and $\sin 2\pi f_{IF}t$ and again low pass filter to form the I and Q channels respectively. The radar engineer really would like to obtain an analytic signal so the correct thing to do is to form the Hilbert transform of the real signal $s_r(t)$. This is impractical, and so it is approximated by this method. The approximation is very good for narrow band signals. The complex signal which results from these transformations is

$$s_c(t) = s_0(t - \frac{2R}{c})e^{-2\pi i f_d t}. \tag{1.6}$$

1.5. *Ambiguity*

Let us return to the case when there is no doppler component; that is, when the target is stationary. The return signal is as in (1.5) together with some noise. What happens next in the radar receiver depends on the information we want to extract from the signal. However at this point there are just two types of information we might want to obtain:

1. detection of the target — that is to decide on the basis of a statistical hypothesis test whether we indeed have a signal or just noise;
2. estimation of the range R of the target.

It turns out that the optimal detector (in the sense of the Neymann-Pearson Lemma) and the maximum likelihood range estimator both require the same operation — to take the correlation of the received signal with a copy of the transmitted signal. This also has the effect of maximizing the post-processing signal-to-noise ratio. In fact it makes sense to do this in the complex domain. Thus we form the function

$$Q(x) = \int s_c(t)\overline{s_0(t-x)}\, dt$$

where the bar over the signal indicates (for when we treat complex signals) complex conjugation. The maximum absolute value of this signal tells us the best estimate of the range in Gaussian noise for this choice of transmitted signal. Note that we will usually have some control over the shape of the latter and we shall spend some time later on issues concerned with its design.

When there is doppler present the radar behaves in a similar way. The return signal is correlated (" matched filtered") with the transmitted signal, so that the resulting signal looks like this:

$$\chi(x,f) = \int_{\mathrm{R}} \mathrm{s}_c(t)\overline{\mathrm{s}_0(t-x)e^{2\pi i f t}}\,dt \tag{1.7}$$

where now the return signal $\mathrm{s}_c(t)$ is as in (1.6). The effect of the matched filter is to maximise the signal power relative to the noise power in the post-processing signal.

We call this the *radar ambiguity function* of the signal s_0 and write it as $\chi_{\mathrm{S}_0}(x,f)$. It is the output we would obtain from our receiver for a signal delay $x = 2R/c$ and velocity $v = f.c/2f_c$. An example for a rectangular pulse is given in Figure 2. We shall deal with

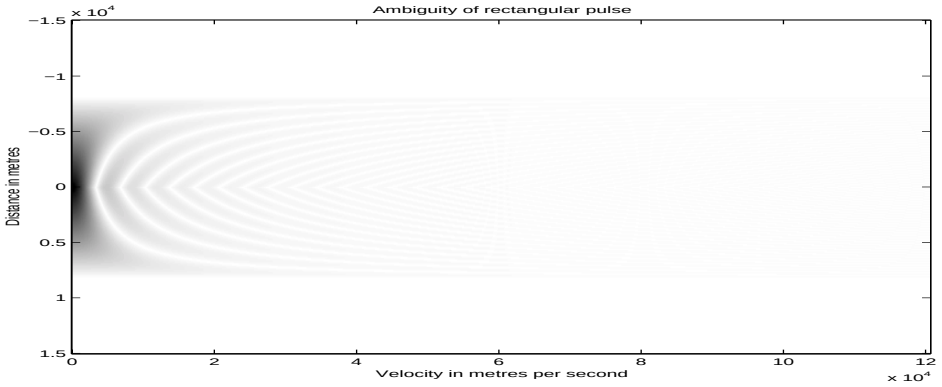

FIGURE 2. Ambiguity of rectangular waveform

the ambiguity function in Section 2. We shall return to the choice of waveforms in Section 3.

1.6. *Extended targets*

When the transmitted signal is reflected from an extended scatterer or from multiple scatterers which are not time varying, the resulting signal is obtained as a convolution of the so-called *reflectivity kernel* of the target(s) with the transmitted signal. Thus if $k(t)$ is this kernel, where now t is a measure of range (which is the same as time in this context), the reflected signal is

$$\mathrm{s}_u(t) = \int_{\mathrm{R}} \mathrm{s}(t-\tau)k(\tau)\,d\tau.$$

Normally velocity does not vary from one part of the target to another so that this subjected to a constant phase shift, though it is straightforward to derive the formula for varying doppler.

2. Ambiguity Functions

This section discusses the mathematical structures lying behind the ambiguity function (particularly the narrow band case) as we have described it in Section 1.5. We shall do so in terms of the unitary representation theory of the Heisenberg group and we shall deal with the theoretical ideas needed first.

2.1. *Representations of the Heisenberg Group*

The narrow band ambiguity function can be described in several different ways at varying levels of abstraction. In an attempt to use some of the power of the abstract setting without demanding too much of the reader, I have attempted to steer a middle course between papers of Schempp [19–22] and the "naive" viewpoint of Wilcox [4]. In many respects the point of view adoped here is that of Auslander and Tolimieri [12]; this couches the theory in terms of the infinite dimensional irreducible representations of the Heisenberg group. As these are probably the most sophisticated ideas discussed here, we present them in detail.

2.1.1. *The Heisenberg Group.* The three dimensional Heisenberg group is the group of $\mathbf{3} \times \mathbf{3}$ upper triangular matrices with $\mathbf{1}$'s on the diagonal, thus:

$$u = \begin{matrix} \mathbf{1} & x & z \\ \mathbf{0} & \mathbf{1} & f \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{matrix} . \tag{2.1}$$

We denote the collection of all such matrices with real entries by $\mathcal{H}$ and note that they form a group under matrix multiplication. Observe that the inverse of u above is

$$u^{-1} = \begin{matrix} \mathbf{1} & -x & -z + xf \\ \mathbf{0} & \mathbf{1} & -f \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{matrix} .$$

This group has the topological structure of R^3, so that we can talk about continuity of functions on $\mathcal{H}$. Lebesgue measure m on R^3 is invariant in the sense that

$$\int_{\mathrm{R}^3} f(\mathrm{ux})\ dm(\mathrm{x}) = \int_{\mathrm{R}^3} f(\mathrm{x})\ dm(\mathrm{x}), \tag{2.2}$$

where the variable x represents an element of $\mathcal{H}$ and u is an arbitrary element of $\mathcal{H}$. Here f is any suitable function for which the integrals make sense (such as continuous complex-valued functions with compact support). In fact, for this group, the measure is both left invariant and right invariant, so that equation (2.2) also holds when u multiplies x on the right in the argument of f.

2.1.2. *Unitary Representations.* A *unitary representation of* $\mathcal{H}$ is a map $\mathrm{x} \to U_{\mathrm{x}}$ from $\mathcal{H}$ into the group of unitary operators $\mathcal{U}(\mathfrak{H})$ of some (separable) Hilbert space $\mathfrak{H}$, which is a homomorphism in the sense that

$$U_{\mathrm{x}_1}.U_{\mathrm{x}_2} = U_{\mathrm{x}_1.\mathrm{x}_2},$$

where the . on the left side indicates multiplication of operators and on the right multiplication of group elements in $\mathcal{H}$. We shall also assume that this map is continuous into the weak-operator topology on $\mathcal{U}(\mathfrak{H})$; that is,

$$\mathrm{x} \mapsto \langle \xi, U_{\mathrm{x}}\eta \rangle$$

is continuous for every ξ and η in the Hilbert space.

There are some simple such representations of $\mathcal{H}$ for which $\mathfrak{H}$ is one dimensional (and so its unitary group can be identified with the group of complex numbers of absolute value 1 under multiplication). An example is the *trivial representation* 1 where

$$1_{\mathrm{x}} = 1$$

for all $\mathrm{x} \in H$. A slightly less trivial example is of the form

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix} \mapsto e^{2\pi i(\alpha x + \beta f)} \tag{2.3}$$

where α and β are any real numbers. This gives a two-dimensional family, parametrized by α and β, of one-dimensional representations, but these are not important for our purposes.

2.1.3. *Irreducibility and Equivalence.* We need to make two more general definitions from representation theory. First we need to say when two representations, say $\mathrm{x} \to U^{(1)}_{\mathrm{x}}$ on $\mathfrak{H}_1$, and $\mathrm{x} \to U^{(2)}_{\mathrm{x}}$ on $\mathfrak{H}_2$, are *equivalent*. By this we mean that they are the same up to a change of Hilbert basis, or equivalently, that there is a unitary operator V from $\mathfrak{H}_1$ onto $\mathfrak{H}_2$ such that

$$VU^{(1)}_{\mathrm{x}} = U^{(2)}_{\mathrm{x}}V$$

for all $\mathrm{x} \in \mathcal{H}$.

A representation of $\mathcal{H}$ ($\mathrm{x} \mapsto U_{\mathrm{x}}$ on $\mathfrak{H}$, say) is *irreducible* if it cannot be broken up into sub-representations, that is, if there are not two subspaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$ each invariant under all of the operators U_{x} ($\mathrm{x} \in H$), such that $\mathfrak{H} = \mathfrak{H}_1 \oplus \mathfrak{H}_2$ in the sense of Hilbert spaces. The restriction of U to these subspaces produces representations $U^{(1)}$ and $U^{(2)}$ on $\mathfrak{H}_1$ and $\mathfrak{H}_2$ respectively, such that, for all $\mathrm{x} \in H$,

$$U_{\mathrm{x}}(\xi_1 + \xi_2) = U^{(1)}_{\mathrm{x}}\xi_1 + U^{(2)}_{\mathrm{x}}\xi_2,$$

where $\xi_j \in \mathfrak{H}_j$ ($j = 1, 2$). While it may not be obvious, this is equivalent to the non-existence of a non-trivial subspace of $\mathfrak{H}$ which is left invariant by each of the operators U_{x}, $(\mathrm{x} \in H)$. This is a consequence of the unitary property of the operators. Equally it is equivalent by Schur's Lemma to the non-existence of a non-trivial bounded operator on $\mathfrak{H}$

which commmutes with all of the U_{x}. Evidently all of the one-dimensional representations described in (2.3) are irreducible.

2.1.4. *The Schrödinger Representation.* Now we define a representation of $\mathcal{H}$ on the Hilbert space $L^2(\mathrm{R})$ of complex square integrable functions on R (or equivalently complex-valued finite energy signals on R) by letting the matrix (2.3) act on a signal s as a point target with doppler shift f and delay x would. Specifically we write

$$U^{\gamma}_{\mathrm{x}}(s)(t) = e^{2\pi i\gamma(ft-z)}s(t-x). \tag{2.4}$$

Note that the variable z acts as a scalar multiplier (namely the scalar $e^{2\pi i\gamma z}$) which makes the homomorphism property work out correctly. Here γ is a non-zero real number. Now the Stone-von Neumann theorem says that, up to equivalence of representations, the representations of the form U^{γ} are the only irreducible representations which are not one dimensional, so that (2.3) and (2.4) are descriptions of all of the irreducible unitary representations of the Heisenberg group up to equivalence. The particular forms of the infinite dimensional irreducible representations of the Heisenberg group given in (2.4) are called the *Schrödinger representation.*

In fact the representation in (2.4) is characterized via the Stone–von Neumann theorem as the unique (up to equivalence) irreducible representation whose restriction to the centre of $\mathcal{H}$ (that is, to the matrices as in (2.1) for which $f = x = 0$) is just $z \to e^{2\pi i\gamma z}I$ where I is the identity operator on $L^2(\mathrm{R})$.

2.2. *Ambiguity Functions*

It is now easy to define the ambiguity function corresponding to a signal s in terms of the infinite dimensional irreducibles of the Heisenberg group. Writing $U = U^1$, we define

$$\chi_{\mathrm{s}}(x,f) = \langle \mathrm{s}, U_{\mathrm{x}}\mathrm{s}\rangle \tag{2.5}$$

where

$$\mathrm{x} = \begin{pmatrix} 1 & f & 0 \\ 0 & 1 & x \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.6}$$

This produces the function we have defined in (1.7). This view of ambiguity functions can be found in the paper [12] of Auslander and Tolimieri. The choice of element of the group may seem arbitrary, but remember that the only effect of changing the top right hand corner of this matrix is to multiply the ambiguity function by a scalar of absolute value 1. In fact, it is really the absolute value of the ambiguity function which is important, since it represents the response of the system to a target of a given range and frequency, and this is unchanged by a change in the top right hand corner. In effect we are making a choice of cross section of the quotient of $\mathcal{H}$ by its centre Z.

Once we have made this definition, it makes sense to allow the *cross ambiguity function*

$$\chi_{s,s'}(x,f) = \langle s, U_{\mathrm{x}} s' \rangle,$$

where x is as in (2.6).

2.2.1. *Properties.* The (cross-)ambiguity function has a number of remarkable properties, one of the most important of which is Moyal's identity:

THEOREM 1. *Let s and s' be in $L^2(R)$. Then*

$$||\chi_{s,s'}||_{L^2(R^2)} = ||s||||s'||.$$

This is a simple consequence of the following theorem:

THEOREM 2. *Let s_1, s_2, s_3, s_4 be in $L^2(R)$. Then*

$$\langle \chi_{s_1,s_2} \cdot \chi_{s_3,s_4} \rangle_{L^2(R^2)} = \langle s_1, s_3 \rangle \langle s_2, s_4 \rangle.$$

One of the important consequences of Moyal's identity is that the unitary representation $\mathrm{x} \mapsto U_{\mathrm{x}}$ is a *square integrable* representation modulo the centre of $\mathcal{H}$. Such representations have been widely studied.

Another remarkable property of this quadratic form on $L^2(\mathrm{R})$ is its variation under the Fourier transform operator.

THEOREM 3. *Let $s \in L^2(R)$. Then, for any $(x,f) \in R^2$,*

$$\chi_{\mathcal{F}(S)}(-f,x) = e^{2\pi i f x} \chi_S(x,f). \tag{2.7}$$

In fact this arises from an interesting intertwining property of the Schrödinger representation. The map

$$(x,f,z) \mapsto (f,-x,-z)$$

is an anti-automorphism of the Heisenberg group which we call κ and so $\mathrm{x} \mapsto \kappa(\mathrm{x})^{-1}$ is an automorphism. Call this ν. Then

$$\nu(x,f,z) = (-f,x,z-fx).$$

Moreover

$$\mathcal{F}(U_{\mathrm{x}}(\mathrm{s})) = U_{\nu(\mathrm{x})}(\mathcal{F}(\mathrm{s})),$$

so that the Fourier transform operator intertwines the two representations U and D where

$$D_{\mathrm{x}} = U_{\nu(\mathrm{x})}. \tag{2.8}$$

Now, for $\mathrm{s} \in L^2(\mathrm{R})$,

$$e^{2\pi i z} \chi_{\mathrm{S}}(x,f) = \langle \mathrm{s}, U_{\mathrm{x}} \mathrm{s} \rangle = \langle \mathcal{F}(\mathrm{s}), \mathcal{F}(U_{\mathrm{x}} \mathrm{s}) \rangle,$$

by the Plancherel Theorem, and in view of (2.8) this equals

$$\langle \mathcal{F}(\mathrm{s}), D_{\mathrm{x}} \mathcal{F}(\mathrm{s}) \rangle = \langle \mathcal{F}(\mathrm{s}), U_{\nu(\mathrm{x})} \mathcal{F}(\mathrm{s}) \rangle = e^{2\pi i f x} \chi_{\mathcal{F}(\mathrm{S})}(-f,x),$$

which proves equation (2.7). A corresponding equality holds for the bilinear form.

2.3. *Weil-Brezin Formula (Zak Transform)*

There is another way of realising the unitary representations we have described and used above. It will simplify formulae if we deal only with the case $\gamma = 1$, though of course the theory extends in an obvious way to all other non-zero values of γ. Instead of using the Hilbert space $L^2(\mathrm{R})$ we consider a Hilbert space $\mathfrak{K}$ of functions F on the Heisenberg group itself. Let Γ be the subgroup of $\mathcal{H}$ consisting of matrices of the form

$$\begin{pmatrix} 1 & n & z \\ 0 & 1 & m \\ 0 & 0 & 1 \end{pmatrix} \tag{2.9}$$

where $n, m \in \mathbf{Z}$ and $z \in \mathrm{R}$. Consider the space $\mathfrak{K}$ of functions F on $\mathcal{H}$ which satisfy

1. $F(\gamma \mathrm{x}) = e^{2\pi i z} F(\mathrm{x})$ for all $\gamma \in \Gamma$ as in (2.9);
2. $\int_{\mathcal{H}/\Gamma} |F(\mathrm{x})|^2 \, d\mathrm{x} < \infty$,

with the inner product

$$\langle F, G \rangle_{\mathfrak{K}} = \int_{\mathcal{H}/\Gamma} F(\mathrm{x}) \overline{G(\mathrm{x})} \, d\mathrm{x}. \tag{2.10}$$

Next we define a representation V of $\mathcal{H}$ just by the simple formula

$$V_{\mathrm{x}}(F)(\mathrm{y}) = F(\mathrm{yx}). \tag{2.11}$$

We need to make a few remarks about these formulae. First note that 1. makes $|F|$ constant on the right cosets of Γ, so that the integral in 2., which is over the right coset space $\Gamma_0 \backslash H$, makes sense. This applies equally to the integral in (2.10). Secondly, the space of functions $\mathfrak{K}$ may be realised as the space of L^2 sections of a line bundle on the torus $\mathbf{T}^2$. Thirdly, those familiar with induced representations will recognize that this is just the representation τ, induced from the subgroup Γ to $\mathcal{H}$, where τ is given by

$$\begin{pmatrix} 1 & n & z \\ 0 & 1 & m \\ 0 & 0 & 1 \end{pmatrix} \longrightarrow e^{2\pi i z},$$

that is, $V = \mathrm{ind}_{\Gamma}^{\mathcal{H}} \tau$, in the terminology of Mackey. We note that Γ is a normal subgroup of $\mathcal{H}$ and so we may apply standard Mackey theory. The stabilizer of this representation is just Γ itself and so the induced representation is irreducible. Now the Stone–von Neumann Theorem tells us that the only irreducible representation whose restriction to the centre of $\mathcal{H}$, the subgroup of matrices with x and y both zero, is just $e^{2\pi i z}$ times the identity operator, is the one we have called U^γ earlier with $\gamma = 1$. It follows that the representation V and $U = U^1$ are equivalent, so there is a unitary operator W intertwining U and V.

It is not difficult to calculate that the intertwining operator is $W : L^2(\mathrm{R}) \mapsto \mathfrak{K}$ given by

$$W(s)(f,x,z) = e^{2\pi iz} \sum_{k\in\mathrm{Z}} s(x+k)e^{2\pi ikf}. \tag{2.12}$$

This formula is called the *Weil-Brezin formula* or (usually without the term $e^{2\pi iz}$) the *Zak transform*. In essence it is a map between functions on R and functions on the unit square (the cross sections of the line bundle) and maps $L^2(\mathrm{R})$ onto $L^2(I^2)$ where I represents the unit interval. In fact the map is unitary. Since it intertwines two irreducible representations it has to be unitary up to a scalar multiple.

This gives us another way of representing ambiguity functions, since for $s \in L^2(\mathrm{R})$,

$$e^{2\pi iz}\chi_s(x,f) = \langle s, U_{\mathrm{x}}\rangle = \langle W(s), V_{\mathrm{x}}W(s)\rangle. \tag{2.13}$$

This alternative view is a useful device in realising theoretical features of χ.

2.4. *Resolution and Hermite functions*

This section is taken largely from the work of Wilcox [4]. He uses a slightly modified form of the ambiguity function. Indeed in many papers the ambiguity of signal s is defined in a *symmetrised* form:

$$A_{\mathrm{S}}(x,f) = \int_{\mathrm{R}} \mathrm{s}(t+\frac{x}{2})\overline{\mathrm{s}(t-\frac{x}{2})}e^{2\pi ift}\,dt. \tag{2.14}$$

This formalism corresponds to a slightly different description of the Heisenberg group as the group on R^3 whose multiplication is given by

$$(x_1,f_1,z_1)*(x_2,f_2,z_2) = (x_1+x_2, f_1+f_2, z_1+z_2+B(x_1,f_1;x_2,f_2)) \tag{2.15}$$

where

$$B(x_1,f_1;x_2,f_2) = \frac{1}{2}(x_1f_2 - f_1x_2). \tag{2.16}$$

It is important to be aware that this does not change the significant properties of the ambiguity function. In fact

$$A_{\mathrm{S}}(x,f) = e^{\pi ifx}\chi_{\mathrm{S}}(x,f).$$

The Hermite functions play a central role in the theory of this object. For our purposes we define them by follows:

$$W_n(t) = \frac{2^{1/4}}{\sqrt{n!}}H_n(2\sqrt{\pi}t)e^{-\pi t^2} \qquad (n = 0,1,2,3,\ldots)$$

where H_n is the nth Hermite polynomial:

$$H_n(t) = (-1)^n e^{x^2/2}\Big(\frac{d}{dx}\Big)^n e^{-\frac{x^2}{2}} \qquad (n = 0,1,2,3,\ldots).$$

The Hermite functions form an orthogonal basis of $L^2(\mathrm{R})$. Their ambiguities are easily calculated in terms of known functions:

$$A_{W_n}(x,y) = e^{-\frac{\pi}{2}(x^2+y^2)} L_n\left(\frac{1}{2}(\pi(x^2+y^2)\right)$$

where L_n is the nth Laguerre function

$$L_n(x) = \frac{1}{n!} e^x \left(\frac{d}{dx}\right)^n (x^n e^{-x}) \text{ for } x > 0.$$

Evidently such functions are radially symmetric. In fact this characterizes Hermite waveforms, as Wilcox observes in [4].

THEOREM 4. *An ambiguity function $A_s(x,y)$ is radially symmetric, that is,*

$$A_s(x,y) = f(x^2+y^2)$$

if and only if $s = cW_n$ for some n.

We mention here too the many papers of Schempp on this subject and on the ambiguity function in general (see, for example, [19–22]).

The sharpness of the peak of the ambiguity function (at $(0,0)$, of necessity) is a measure of the waveform's capacity to resolve nearby (in both doppler and range) targets. One way of measuring resolution capabilities is as follows. Fix a minimum response resolution ϵ and consider the level curve

$$|A_s(x,y)|^2 = 1 - 4\pi^2\epsilon^2.$$

This is approximately an ellipse, which Wilcox calls the *resolution ellipse*. Write its equation as

$$\beta^2 x^2 + 2\gamma xy + \tau^2 y^2 = \epsilon^2.$$

In fact

$$\beta = \int_{\mathrm{R}} f^2 |\mathcal{F}(s)(f)|^2 \, df \qquad \tau = \int_{\mathrm{R}} t^2 |s(t)|^2 \, dt$$

are respectively measures of the *bandwidth* and the time duration of the signal. Two close point targets will produce a response which comprises two copies of the ambiguity function with slightly different centres. These two targets will be most difficult to resolve if their resolution ellipses share a common major axis. The two targets are said to be *resolvable* if the centres of their resolution ellipses are one major semi-axis apart. This number — the *smallest resolvable separation* — is given by

$$R(\epsilon)^2 = \frac{2\epsilon^2}{\tau^2 + \beta^2 - \sqrt{(\tau^2-\beta^2)^2 + 4\gamma^2}}.$$

Wilcox calls $R(\epsilon)/\epsilon$ the *resolution factor* ρ, but perhaps for mathematical precision the resolution factor should be the derivative of $R(\epsilon)$ at 0. Wilcox has shown that the best

resolution factor obtainable subject to constraints $\beta \leqslant \beta_0$ and $\tau \leqslant \tau_0$ on the bandwidth and time duration is

$$\rho = \frac{1}{\min(\beta_0, \tau_0)}.$$

The Hermite functions arise in this context as follows. Let Ω_n be the subspace of $L^2(\mathrm{R})$ consisting of all functions of the form $P(t)e^{-\pi t^2}$ where P is a polynomial of degree at most n. The following theorem of Hardy characterizes Ω_n.

THEOREM 5. *Let $s \in L^2(R)$ satisfy*

$$s(t) = O(t^n e^{-\pi t^2}), \qquad |t| \to \infty,$$
$$\mathcal{F}s(f) = O(f^n e^{-\pi f^2}), \qquad |f| \to \infty.$$

Then $s \in \Omega_n$

Now we have the following theorem of Wilcox.

THEOREM 6. *The Hermite waveform W_n has the smallest resolution factor of all waveforms in Ω_n.*

2.5. *Characterization*

We discuss here the problem of how to describe the class $\mathcal{A}$ of ambiguity functions χ_s as s ranges over all functions in $L^2(\mathrm{R})$. One easy characterization arises from the observation that the Fourier transform of the ambiguity function with respect to the doppler variable is

$$\mathcal{F}_2^{-1}(\chi_S)(x, \tau) = s(\tau)\overline{s(\tau - x)},$$

where we use the notation $\mathcal{F}_2$ to mean the Fourier transform with respect to the second variable. Making the change of variable

$$\tau - x = u, \qquad \tau = \tau,$$

we obtain the function

$$H(u, \tau) = s(\tau)\overline{s(u)}.$$

We write $H = \mathcal{R}(\chi_S)$ to denote this transformation; that is,

$$\mathcal{R}(F)(s, t) = \mathcal{F}_2^{-1}(F)(t - s, t). \tag{2.17}$$

To describe ambiguity functions then it is enough to describe these functions intrinsically and this is easily accomplished. The following is a theorem of Wilcox [4], but see also [12].

THEOREM 7. *Let $H \in L^2(R^2)$ satisfy*

1. $\overline{H(u, \tau)} = H(\tau, u)$
2. $H(x, x) \geqslant 0$
3. $H(v, v)H(x, \tau) = H(x, v)H(v, \tau)$

Then there is a signal $s \in L^2(R)$ *such that* $\mathcal{R}(\chi_S) = H$*; that is,*

$$\chi_S(x, f) = \mathcal{F}_2(K)(x, f)$$

where $K(u, \tau) = H(\tau - u, \tau)$.

A somewhat similar theorem, also due to Wilcox, characterizes ambiguity functions in terms of their expansions in tensor product bases of the Hilbert space

$$L^2(\mathrm{R}^2) = L^2(\mathrm{R}) \otimes L^2(\mathrm{R}).$$

A *tensor* orthonormal basis of $L^2(\mathrm{R}^2)$ is a basis $\psi_{m,n}(x, y) = \phi_m(x)\overline{\phi_n(y)}$ where (ϕ_n) is an orthonormal basis of $L^2(\mathrm{R})$. Any member F of $L^2(\mathrm{R}^2)$ has an expansion

$$F(x, y) = \sum_{m,n} c_{m,n}\phi_m(x)\overline{\phi_n(y)} \tag{2.18}$$

which converges in $L^2(\mathrm{R}^2)$. The following simple theorem characterizes ambiguity functions in terms of their expansion in such a basis.

THEOREM 8. *A function* $F \in L^2(R^2)$ *is an ambiguity function if, in equation (2.18),*

$$c_{m,n} = a_n\overline{a_m}$$

for some sequence (a_n).

This has the following immediate corollary

COROLLARY 1. *A function* $F \in L^2(R^2)$ *is an ambiguity function if, in equation (2.18),*

$$c_{k,k}c_{m,n} = c_{m,k}c_{k,n} \qquad \textit{and} \qquad c_{k,k} \geqslant 0 \textit{ for all } k, m, n.$$

Another characterization of ambiguity functions arises from their description in terms of the representation of the Heisenberg group. To this end, we recall that a complex-valued function p on $\mathcal{H}$ is *positive definite* if, for any choice of N complex numbers c_n and elements of the group x_n,

$$\sum_{m,n=1}^{N} c_m\overline{c_n}p(\mathrm{x}_m\mathrm{x}_n^{-1}) \geqslant 0.$$

positive definite functions arise as functions of the form

$$p(\mathrm{x}) = \langle \xi, R(\mathrm{x})\xi \rangle,$$

where R is a representation of the group and ξ and is an element of the Hilbert space on which it acts. Among the positive definite functions, the ones which arise from irreducible representations are characterized by the following property: we call a positive definite function p *extremal* if, for all other positive definite functions q for which $p - \alpha q$ is also positive definite for some $\alpha > 0$ q is a multiple of p.

Since we know the irreducible representations of the Heisenberg group, we also know all extremal positive definite functions. An ambiguity function is now a function $\chi(x, f)$ in $L^2(\mathrm{R})$ for which the function

$$p(x, f, z) = e^{2\pi i z}\chi(x, f)$$

is an extremal positive definite function. While this characterization is not as easily checked as the earlier ones it is of theoretical importance.

2.6. *The Abiguity Problem*

The simple abiguity problem is that of determining a signal s from its ambiguity χ_{S}. To make it mathematically meaningful we restrict the signal to belong to $L^2(\mathrm{R})$, so that for such signals we are interested in the question: is $s \mapsto \chi_{\mathrm{S}}$ one to one. The answer to this question is yes.

It is relatively simple to show the following theorem (see, for example, [12]),

THEOREM 9. *Let s_1 and s_2 be two $L^2(R)$ functions for which $\chi_{s_1} = \chi_{s_2}$ then $s_1 = s_2$.*

However this is not really an answer to the radar engineer's problem for several reasons, one of which is that the ambiguity function is not what is important. It is the absolute value of the ambiguity function which is really of interest. In this context, the problem arises of determining all possible waveforms s whose ambiguity functions have the same absolute value. Much of this section is adapted from work of Philippe Jaming indexJaming, Philippe and his collaborators (see [7, 10, 15]). Following Bueckner ([9]) we state here the following:

DEFINITION 1 (General Abiguity Problem). *Given $s_1, s_2 \in L^2(R)$ what is the set of all pairs s_1', s_2' such that*

$$|\chi_{s_1,s_2}(x, f)| = |\chi_{s_1',s_2'}(x, f)| \tag{2.19}$$

for all $(x, f) \in R^2$?

This problem is unsolved in general, but the following simple observation shows that equation (2.19) does not force $(s_1, s_2) = (s_1', s_2')$. Let

$$\mathrm{s}_j' = U_{\mathrm{x}'}\mathrm{s}_j \text{ where } \mathrm{x}' \in \mathcal{H}. \tag{2.20}$$

Then , with $\mathrm{x} = (x, f, 0) \in \mathcal{H}$,

$$\begin{aligned}|\chi_{\mathrm{s}_1',\mathrm{s}_2'}(x, f)| = |\langle \mathrm{s}_1', U_{\mathrm{x}}\mathrm{s}_2'\rangle| &= |\langle U_{\mathrm{x}'}\mathrm{s}_1, U_{\mathrm{x}}U_{\mathrm{x}'}\mathrm{s}_2\rangle| \\ &= |\langle \mathrm{s}_1, U_{\mathrm{x}'}^* U_{\mathrm{x}}U_{\mathrm{x}'}\mathrm{s}_2\rangle| = |\langle \mathrm{s}_1, U_{\mathrm{x}}U_{\mathrm{z}}\mathrm{s}_2\rangle|\end{aligned}$$

where $\mathrm{z} = \mathrm{x}^{-1}\mathrm{x}'^{-1}\mathrm{x}\mathrm{x}'$. Now z is a commutator and so (since the Heisenberg group is two step nilpotent) belongs to the centre of $\mathcal{H}$. It follows from the irreducibility of U that U_{z} is just a scalar times the identity operator, say $U_{\mathrm{z}} = e^{2\pi i z}I$, and hence that

$$|\langle \mathrm{s}_1, U_{\mathrm{x}}U_{\mathrm{z}}\mathrm{s}_2\rangle| = |\langle \mathrm{s}_1, U_{\mathrm{z}}U_{\mathrm{x}}\mathrm{s}_2\rangle| = |\langle e^{2\pi i z}\mathrm{s}_1, U_{\mathrm{x}}\mathrm{s}_2\rangle| = |\langle \mathrm{s}_1, U_{\mathrm{x}}\mathrm{s}_2\rangle| = |\chi_{\mathrm{s}_1,\mathrm{s}_2}(x, f)|.$$

We will say that s_1', s_2' are *Heisenbery related* to s_1 and s_2 if a relationship holds as in (2.20). Two such related pairs have the same absolute values for their ambiguity functions. The problem is now restated as: Does (2.19) force s_1', s_2' to be Heisenberg related to s_1, s_2?.

Unfortunately this is not the case. For instance if

$$s_1(t) = s_2(t) = \left(\frac{\sin t}{t}\right)^n \sin(2\pi n t)$$
$$s_1'(t) = s_2'(t) = \left(\frac{\sin t}{t}\right)^n \cos(2\pi n t)$$

then

$$|\chi_{s_1,s_1}(x,f)| = |\chi_{S_1',S_1'}(x,f)|,$$

for all $(x,f) \in R^2$, though s_1 is not Heisenbery related to s_1'. The proof of this is relatively easy if one works in the Fourier domain and uses the identity (2.7). The Fourier transforms of these functions comprise two non-zero identical "pieces" around $\pm 2n\pi$; the difference between the two being in the signs of these pieces. The ambiguity therefore is non-zero only when translated of these pieces overlap. It is now clear that the ambiguities can differ only is sign.

For special classes of functions it is the case that all s_1', s_2' are obtained by such transformations. For instance, Bueckner [9] and De Buda [5] have proved that, if s_1 and s_2 are both of the form $P(t)\exp(-\frac{t^2}{2})$ with P a polynomial, then s_1' and s_2' are Heisenbery related to s_1 and s_2.

Jaming indexJaming, Philippe [10] has obtained some results on the abiguity problem for compactly supported functions. His results are stated in terms of the symmetrised form, though obviously simple modifications convert to the χ-form of the ambiguity function. Let $u \in L^2(R)$ be a compactly supported function and suppose v satisfies $|A_v| = |A_u|$. Then it is relatively easy to see that v is also compactly supported. Moreover if the support of u is contained in an interval of length $2a$ then the support of v is also contained in an interval of length $2a$. In fact Jaming shows that the interval of support must have the same length for both u and v.

We may now assume that supports of both u and v are contained in $[-a,a]$ and no smaller interval, in particular, u and v are compactly supported. The Paley-Wiener theorem ensures that $A_u(x,y)$ and $A_v(x,y)$ are both entire functions of exponential type in the y variable. Now (cf : [10]) the following holds.

$$A_u(x,y)\overline{A_v(x,\bar{y})} = A_v(x,y)\overline{A_u(x,\bar{y})} \quad \text{for all } x \in R, y \in C. \tag{2.21}$$

On the other hand, by the Hadamard factorisation theorem, an entire function $f(z)$ of exponential type is entirely determined by its zeros, up to a factor $\lambda e^{\mu z}$ with $\lambda, \mu \in C$. Unfortunately (2.21) only tells us that, for fixed x, if z is a zero of $A_u(x,.)$ then either z or $\bar{z}$ is a zero of $A_v(x,.)$.

Several cases occur, for instance, A_u may only have real zeros (e.g. if $u = A_{[a,b]}$), then A_u and A_v have the same zeroes.

There are some functions u for which every ambiguity partner v is such that either A_u and A_v have the same zeroes, or A_u and A_{Zv} have the same zeroes, where $Zu(t) = u(-t)$. The final alternative is that $A(u)$ and $A(v)$ may have some common non-real zeroes and some conjugate zeroes.

In what follows, after replacing u by Zu or by some function Heisenbery related to u, we shall assume that $A(u)$ and $A(v)$ have the same zeroes. In other words we now consider the following restricted radar abiguity problem :

PROBLEM 1 (Restricted Radar Abiguity Problem). *Given a compactly supported $u \in L^2(R)$, what is the set of ambiguity partners v of u, such that for every $x \in R$, $A(u)(x,.)$ and $A(v)(x,.)$ have the same zeroes in the complex plane ?*

Jaming calls such ambiguity partners *restricted ambiguity partners* and shows that there exist compactly supported functions u which have ambiguity partners that are not restricted ambiguity partners either of u or of Zu. He also shows the following result.

THEOREM 10. *Let $u \in L^2(R)$ be a compactly supported function and let v be a restricted ambiguity partner of u. If Ω is the open set of all x such that $A(u)(x,.)$ is not identically 0, there exists a locally constant function ϕ on Ω such that, for every t_0, t_1, t_2 belonging to the support of u,*

$$\phi(t_2 - t_1) + \phi(t_1 - t_0) \equiv \phi(t_2 - t_0) \quad (2\pi)$$

and

$$v(x) = ce^{i\phi(x-a-x_0)} e^{i\omega x} u(x-a)$$

for some $a \in R, \omega \in R$, $c \in \mathbf{T}$ and some x_0 belonging to the support of u. Conversely, every function v of that form is a (restricted) ambiguity partner of u.

This theorem essentially states that if u is "simple" (in particular, the support is an interval) then the solutions of the abiguity problem are "simple", whereas for complicated u (for example, when the support has big gaps) the solutions are also complicated. In [7], the discrete abiguity problem is considered and various interesting results are obtained.

2.7. *Approximation*

An important problem for the radar engineer is to invent waveforms with specific ambiguity properties. The abiguity problem of the previous section is one part of that problem — it examines the possibility of uniquely specifying a waveform in terms of its ambiguity (or its absolute value), and have we have seen without further constraints this is not successful. Nonetheless the radar engineer needs despite the non-uniqueness to be able to find waveforms with appropriate ambiguity fucntions. One way to do this is first to invent a function $F \in L^2(\mathrm{R}^2)$ with the appropriate properties and then find the ambiguity function which

most closely approximates it. In other words, the problem is to find the signal $s \in L^2(\mathrm{R})$ which satisfies

$$s = \operatorname{argmin}||A_s - F||_{L^2(\mathrm{R}^2)}.$$

Using the unitary operator $\mathcal{R}$ described in equation (2.17), we obtain

$$||A_s - F||^2 = ||s(t)\overline{s(\tau)} - \mathcal{R}(F)||^2 = 2(1 - \langle \mathcal{R}(F), s \otimes \overline{s} \rangle)$$

if both F and s are normalized. The problem now is one of maximizing

$$\langle \mathcal{R}(F), s \otimes \overline{s} \rangle)$$

which corresponds to finding the eigenvector whose eigenvalue has the largest absolute value of the integral operator with kernel $\mathcal{H}$, that is,

$$\mathcal{H}(u)(t) = \int_{\mathrm{R}} H(t, \tau) u(\tau)\, d\tau.$$

This operator is compact and so its eigenvalues $\mu_1, \mu_2, \ldots$ decrease in absolute value and tend to zero. We choose for our signal s the eigenvector corresponding to μ_1.

2.8. *The Wide-Band Ambiguity Function*

The classical ambiguity function of Woodward is a result of making the narrow band approximation. There are situations where this is not valid. For example in sonar work where the speed of sound through water is not sufficiently high compared to velocities of targets that the approximation is valid. Also in radar the approximation breaks down in two ways:

1. the velocity is a substantial proportion of the velocity of light. This is an unlikely scenario in practice.
2. the radar is "wide-band"; that is, it uses a broad spectrum. This is an area of increasing interest.

The *wide band cross ambiguity* function is defined by

$$W_{\mathrm{s}_1,\mathrm{s}_2}(x, \alpha) = \alpha^{1/2} \int_{\mathrm{R}} \mathrm{s}_1(t)\overline{\mathrm{s}_2(\alpha(t + x))}\, dt,$$

where $\alpha > 0$ represents the scaling due to doppler and x corresponds to range as before; the bar over the second term in the integral corresponds to complex conjugation. The (auto)-ambiguity function of a signal s is $W_s(x, \alpha)$.

This formula too corresponds to a representation of a group. This time the group in question is the so called $ax + b$ group G of 2×2 matrices

$$\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$$

where $a > 0$ and $b \in \mathrm{R}$. The representation is again on $L^2(\mathrm{R})$ and is given by

$$U_{a,b}(\mathrm{s})(t) = a^{1/2}\mathrm{s}(at + b)$$

for $f \in L^2(\mathrm{R})$. Thewide-band ambiguity function is then just

$$W_{s_1,s_2}(x,\alpha) = \langle s_1, U_{\alpha,x}(s_2)\rangle_{L^2(\mathrm{R})}.$$

Several properties of the narrow band ambiguity function carry over to the wide band case. However it is necessary to restrict to a smaller class of functions to make the theory closely approximate that of the narrow band ambiguity. We write $H^2(\mathrm{R})$ for the Hardy class of functions whose Fourier transforms vanish on the negative half-line. Note that this subspace is invariant under the action of $U_{a,b}$ for all $a > 0$ and $b \in \mathrm{R}$. The original representation therefore cannot have been irreducible, though the restriction to $H^2(\mathrm{R})$ is. We need to restrict further to the those functions s which satisfy

$$\int_{\mathrm{R}^+} |\mathcal{F}(s)(f)|^2 \, \frac{df}{f} < \infty.$$

This class is designated $\mathfrak{H}$ and we write

$$\langle s_1, s_2\rangle = \int_{\mathrm{R}^+} \mathcal{F}(s_1)(f)\overline{\mathcal{F}(s_2)(f)} \, \frac{df}{f}.$$

It is now possible to obtain a Moyal type identity. For u in $H^2(\mathrm{R})$ and $v \in \mathfrak{H}$, $W_{u,v}$ is in $\mathfrak{K} = L^2(\mathrm{R}^+ \times \mathrm{R}; a^{-1}\, da\, db)$, and

$$\langle W_{u,v}, W_{u',v'}\rangle_{\mathfrak{K}} = \langle v, v'\rangle_{L^2(\mathrm{R})}\langle u, u'\rangle_{\mathfrak{H}}.$$

There are more aspects of the narrow band theory which carry over to the wide band theory with appropriate modifications. For example the approximation results of Section 2.7 have a counterpart in the wide band theory. The $ax + b$ group is an example (the simplest) of a non-unimodular group — its left and right invariant measures are different. It is this fact which produces the need for the two separate Hilbert spaces $H^2(\mathrm{R})$ and $\mathfrak{H}$. For more information on this issue, the reader is referred to [6]. For more on the general theory of the wide-band ambiguity, we refer the reader to [1, 2, 11, 17, 23]. Jaming [10] has considered the ambiguity problem for wide-band ambiguity.

3. Waveform Design and Processing

For many purposes the ideal radar waveform would produce an ambiguity function which was the so-called “thumbtack” — that is, zero everywhere except at the origin. This would have ideal range and doppler discrimination. However Moyal’s identity tells us that no (finite energy) signal gives rise to that waveform since the result has to be in $L^2(\mathrm{R}^2)$ and have norm equal to the square of the norm of the signal. A very short pulse would seem to have many advantages in this sense. It has a good thumbtack-like ambiguity at least in the range direction. However it has many disadvantages. If it is not to have extremely high power (which of course means that the electronics have to be able to deliver this amount of power to the antenna) then the total electromagnetic energy hitting the target is very small and the resulting energy returned to the receiver orders of magnitude smaller. The received energy falls off as

the fourth power of the distance of the target so total *energy on target* is a significant factor in radar detection, particularly of distant targets. From this perspective longer waveforms are better. Another disadvantage of a short pulse is that it is a very wideband signal, and so not so good in the doppler direction. Moreover, to produce the shape of siuch a pulse and effectively deal with it on reception requires electronics capable of handling such signals. Maintaining linear responses over such a frequency spread is difficult. Radar engineers then aim to have long waveforms whose auto- correlation (matched filter) produces something approximating a thumb-tack. This process is called *pulse compression*. A purely random signal of infinite length (and therefore energy) — that is one comprising a bi-infinite sequence of Gaussian random variables of zero mean and constant variance has an expected ambiguity which is a thumbtack. Figure 3 gives a the ambiguity of a finite waveform obtained using a random number generator. As a result pseudo-random codes (which are finite and deterministic approximations to random signals) are often used as waveforms in radar. Many papers have been written examining the properties of different waveforms. We give a small sample of these here.

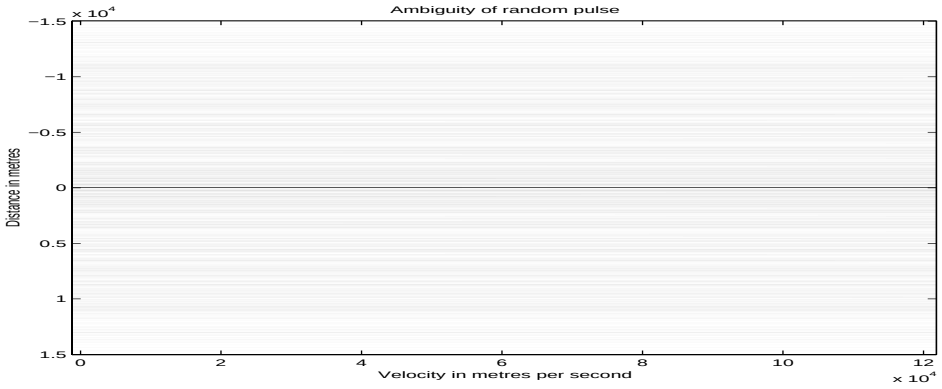

FIGURE 3. Ambiguity of random waveform

Before we do, however, we note that for some purposes a thumbtack is not the best waveform. For example there are circumstances in which a waveform which is extremely doppler tolerant (that is its ambiguity is a ridge along the doppler axis) and in others we may require range tolerance.

3.1. *Conventional waveforms*

We shall first talk about waveforms for a convventional radar.

3.1.1. *Rectangular and Gaussian pulses.* The simplest waveform used in radar is just a simple pulse, either rectangular or Gaussian. Such pulses are shown in Figure 4 and their ambiguity functions in Figure 5. In fact, of course, we plot the absolute value of the ambiguity

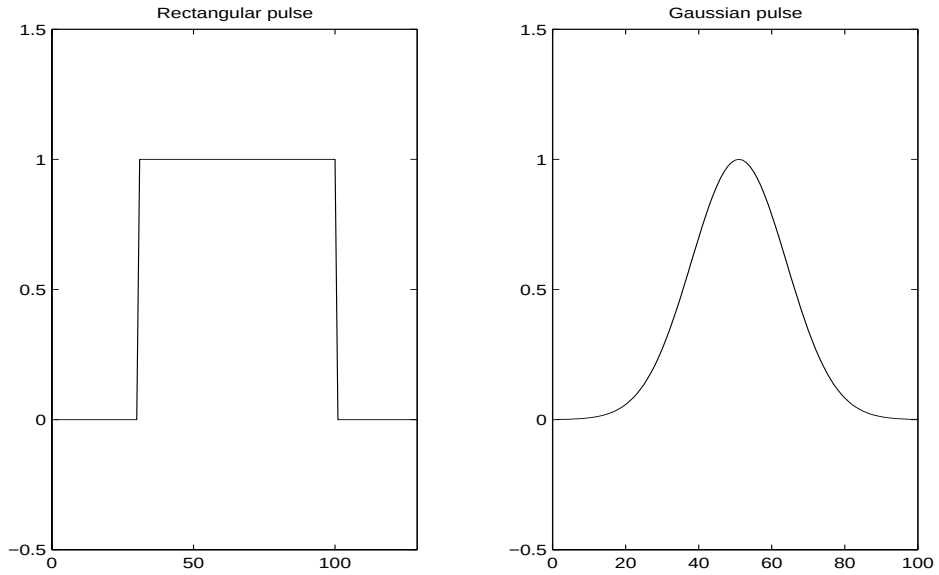

FIGURE 4. Rectangular and Gaussian waveforms

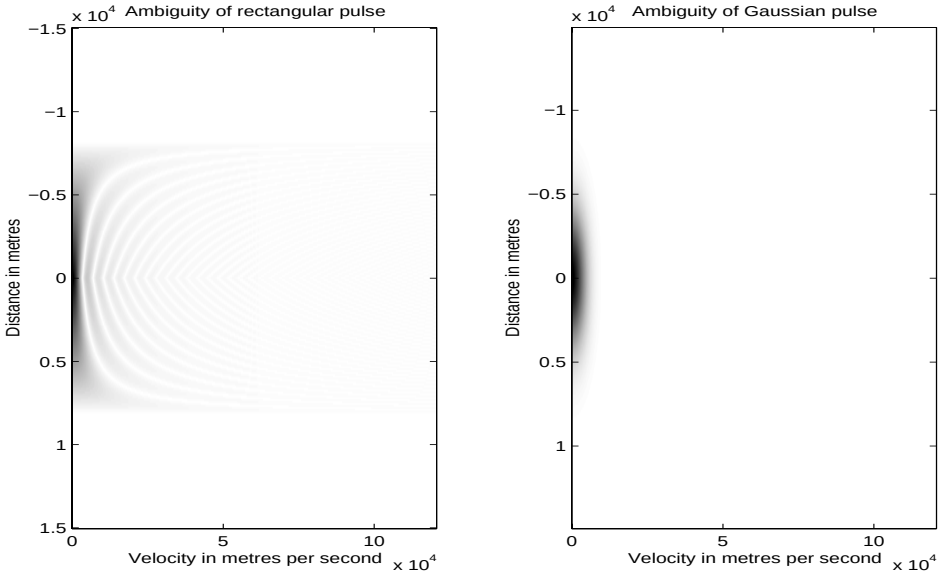

FIGURE 5. Ambiguities of rectangular and Gaussian waveforms

since in general it is a complex function. Notice the “sinc” behaviour on the doppler direction of the rectangular pulse. It has a relatively sharp peak in the doppler direction but then has

"sidelobes". The effect of these sidelobes is to mix targets of differing dopplers into the same doppler bin. This is an issue for shipborne radars, in particular, where the reflection from the sea close to the ship might obscure a distant (and therefore fainter) airplane whose doppler should separate it.

A glance at the doppler axes in Figure 5 shows that I have plotted doppler shifts corresponding to quite unrealistic velocities. I should point out that the values assume a 100 microsecond pulse (a long one) at 10 GHz. A more realistic ambiguity plot is shown in Figure 6 — it gives the ambiguities of the two waveforms for velocities up to mach 3 — about 1km metres per second. This too is an unrealistic representation since the grey scales do not adequately convey the significance of the pixels in the image. Again from an engineering viewpoint it is better to plot the logarithm of the values of the ambiguity — to be precise we need a dB plot where the value in decibels is equal to $\mathbf{20}\log_{10}|\chi|$. These are shown in Figure 7. As can be seen, the main problems with these waveforms are

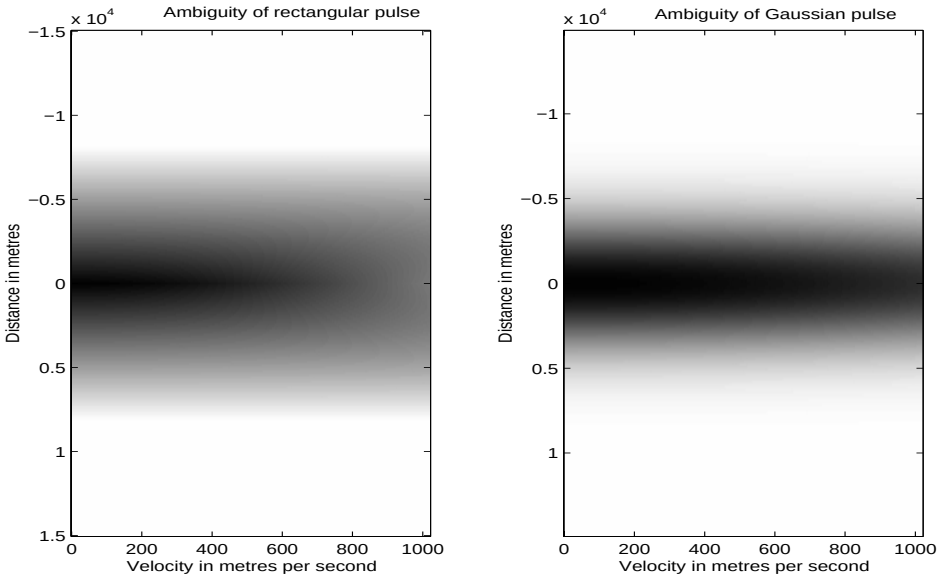

FIGURE 6. Ambiguities of rectangular and Gaussian waveforms out to mach 3

- significant doppler tolerance;
- very low range resolution.

To reiterate, a long pulse is used here and better range resolution could be obtained with a shorter pulse. However, compared to many other waveforms both of these have significant weaknesses.

3.1.2. *Chirp waveforms.* Perhaps the next most common waveform is called a "chirp" or "linear chirp". It is a signal with a linear increase or decrease in frequency over time. Using

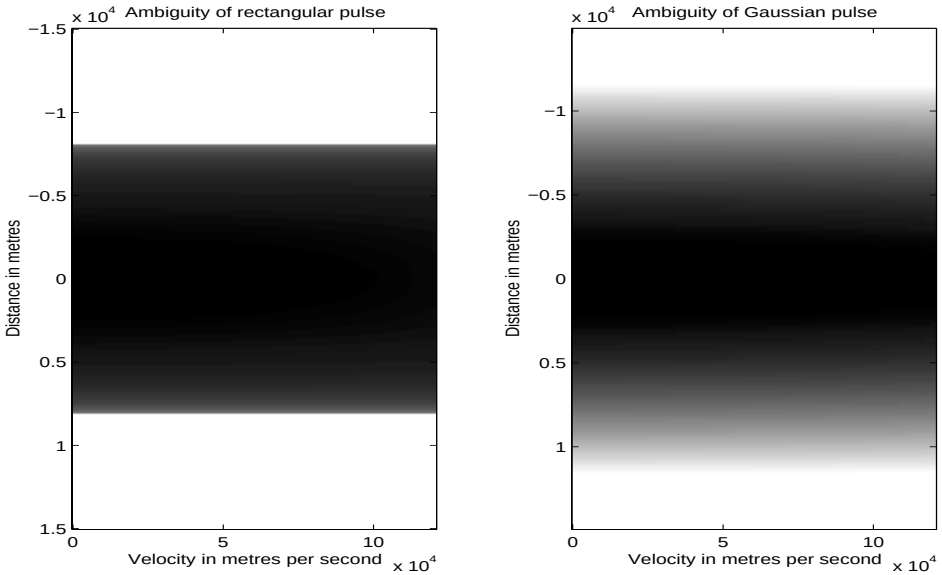

FIGURE 7. dB plots of ambiguities of rectangular and Gaussian waveforms

the fact that we have effectively real and complex (or I and Q channels) it is possible to transmit the signal $e^{-\alpha i t^2}$ (rather than just its real part, say). The real part is shown in Figure 8, along with its auto- correlation and the ambiguity of the complex signal is shown in Figure 8. It will be seen that the ambiguity of the chirp has a ridge running across the diagram which has quite a sharp peak, however, it does have a slight fall from left to right across the image. This gives rise to an ambiguity between range and doppler. The sharpness of the peak is a very useful feature of the chirp It is used in many applications and in particular in synthetic aperture radar. As we have said, radar engineers describe waveforms with auto-correlations looking like that of the chirp as *pulse compression waveforms*. They have the property that they behave to some extent (at least for low doppler values) like a very sharp pulse while retaining high energy.

3.1.3. *Barker sequences.* The fourth favourite waveform is based on Barker codes. Such a code is a sequence of ± 1s with the property that the auto correlation takes just the values $0, \pm 1$ and the length of the code. For example the code $[1 - 1111]$ has auto correlation

$$[101050101].$$

Barker codes have almost perfect auto correlation properties and quite doppler intolerant ambiguity properties. The ambiguity is shown in Figure 10. Unfortunately they are only known to exist in lengths 3, 5, 7, 11 and 13. Much mathematical effort has gone into trying to show that these are the only lengths of Barker codes and much computing effort into finding longer ones without success in both cases.

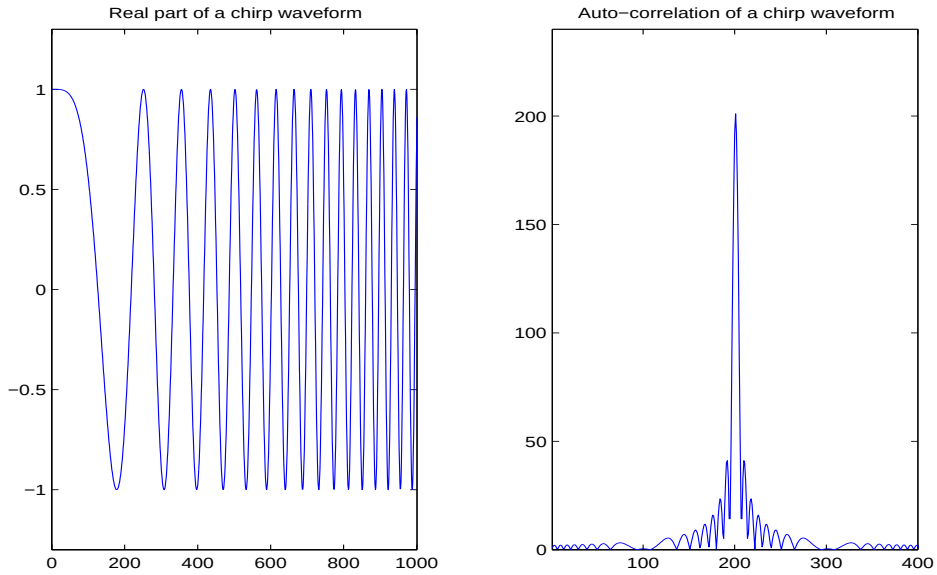

FIGURE 8. Chirp waveform and auto- correlation

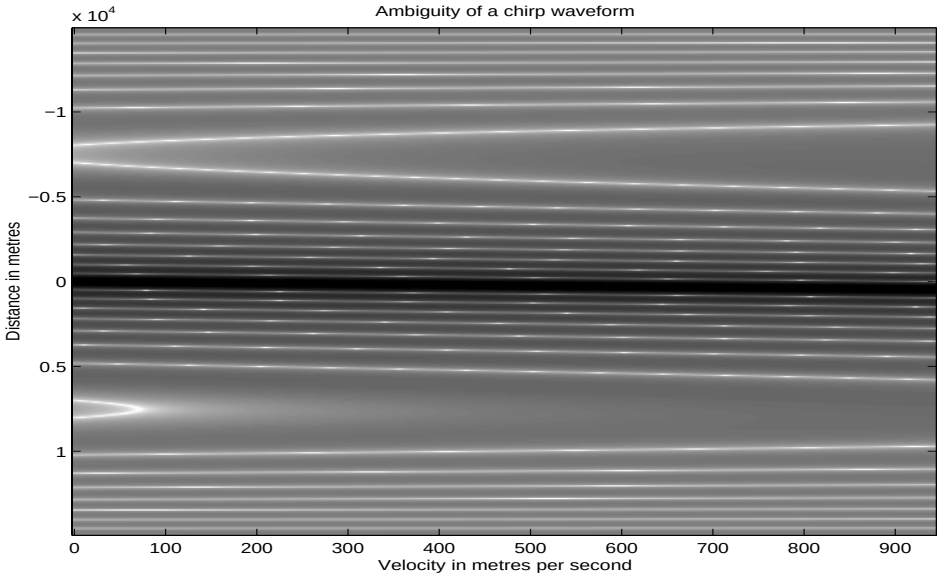

FIGURE 9. Ambiguity of Chirp waveform

One way in which radar engineers have sought to overcome the lack of longer Barker sequences is to use so called “Barker on Barker” sequences. These are formed by choosing

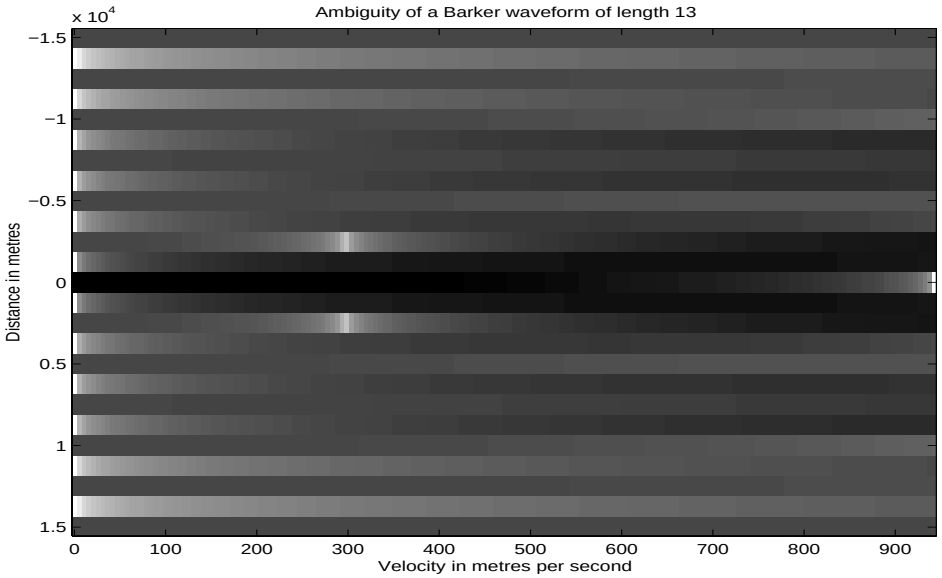

FIGURE 10. Ambiguity of Barker waveform of length 13

two Barker codes (which may be the same) and substituting a multiple of one for each entry in the other using that entry to determine the mulltiplicative factor. Thus a Barker of length 3 is $[\mathbf{1}\ -\mathbf{1}\ -\mathbf{1}]$ and using it together with the Barker of length 5 given above we obtain the following code of length 15:

$$[\ \ \mathbf{1}\ -\mathbf{1}\ -\mathbf{1}\ -\mathbf{1}\quad \mathbf{1}\quad \mathbf{1}\quad \mathbf{1}\ -\mathbf{1}\ -\mathbf{1}\quad \mathbf{1}\ -\mathbf{1}\ -\mathbf{1}\quad \mathbf{1}\ -\mathbf{1}\ -\mathbf{1}].$$

Repeated substitution of this kind allows us to obtain arbitrarily long codes. Unfortunately their auto correlation and ambiguity properties are considerably less than ideal. Figure 11 illustrates this for a Barker 7 on a Barker 13.

3.1.4. *Costas Arrays.* Costas arrays are used in the design of stepped frequency radars; that is, radars in which the waveform comprises a sequence of pure tones at differing frequencies. Thus a chosen collection of $\boldsymbol{N}$ frequencies $\mathbf{f} = (\boldsymbol{f_1}, \boldsymbol{f_2}, \ldots, \boldsymbol{f_N})^{\boldsymbol{T}}$ are transmitted in consecutive time intervals of equal duration in this order. The frequencies are chosen from equally spaced ones $\boldsymbol{\phi} = (\boldsymbol{\phi_1}, \boldsymbol{\phi_2}, \ldots, \boldsymbol{\phi_N})^{\boldsymbol{T}}$, so that

$$\boldsymbol{\phi_j} = \boldsymbol{\phi_1} + (\boldsymbol{j} - \mathbf{1})(\boldsymbol{\phi_2} - \boldsymbol{\phi_1}). \tag{3.1}$$

The choice of order can be described bya permutation matrix $\boldsymbol{P}$ so that $= \mathbf{P}\boldsymbol{\phi}$. In order to obtain good ambiguity performance, the matrix $\boldsymbol{P}$ is chosen so that for any shift of the matrix horizontally or vertically the non-zero terms overlap in at most one place. Such a matrix is called a *Costas array*. This can be re-expressed in terms of the set $\boldsymbol{S}$:

$$\boldsymbol{S} = \{(\boldsymbol{i}, \boldsymbol{j}) | \boldsymbol{P_{ij}} = \mathbf{1}\}. \tag{3.2}$$

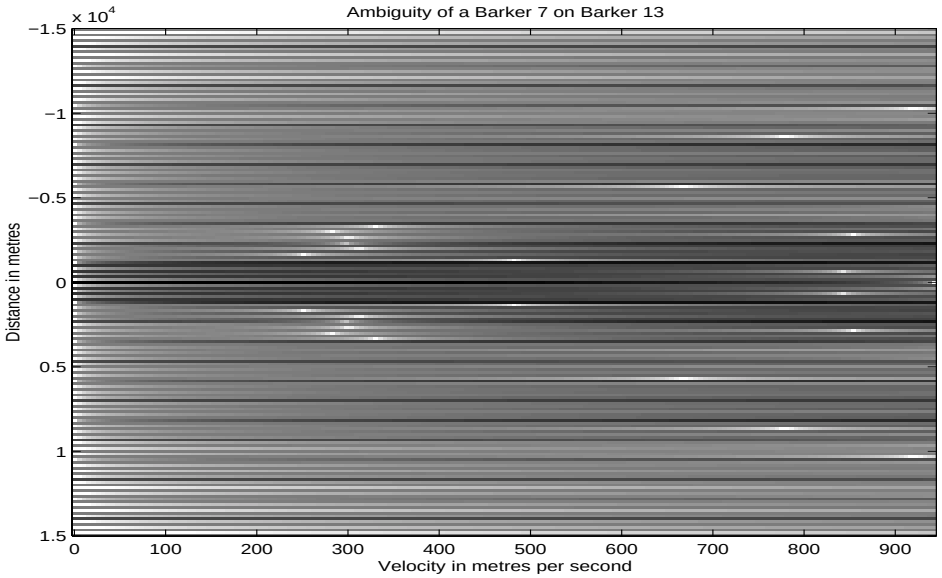

FIGURE 11. Ambiguity of a "Barker on Barker" waveform

Now $\boldsymbol{P}$ is a Costas array if, for $\mathrm{r}, \mathrm{s}, \mathrm{r}', \mathrm{s}'$ distinct elements of $\boldsymbol{S}$, no equation of the form

$$\mathrm{s} - \mathrm{r} = \mathrm{s}' - \mathrm{r}' \tag{3.3}$$

can hold.

There are several ways to construct Costas arrays (see Golomb [8]). Here is one due to Golomb. Let $\boldsymbol{\alpha}$ and β be two distinct primitive elements in the Galois field $\boldsymbol{GF}(\boldsymbol{p}^{\boldsymbol{n}})$, and let $\boldsymbol{P}$ be the matrix whose entries are $\mathbf{0}$ except at the elements of the set

$$\boldsymbol{A} = \{(i, j) : \boldsymbol{\alpha}^{i} + \boldsymbol{\beta}^{j} = \mathbf{1} \qquad (\mathbf{1} \leqslant \boldsymbol{i} \leqslant \boldsymbol{q} - \mathbf{2})\}. \tag{3.4}$$

Then $\boldsymbol{P}$ is a permutation matrix and indeed a Costas array. Figure 3.1.4 is an example of a Costas array of size $\mathbf{31}$ and Figure 3.1.4 is its ambiguity.

3.2. *Complementary Waveforms*

Following the view that the ideal waveform is a thumb-tack, we should require of such a waveform that at least its auto- correlation be a spike at the origin and zero elsewhere. As we have seen, the chirp approximates this quite well. No single waveform can achieve this perfect auto correlation, though Barker waveforms approach the ideal. It is possible however to find a pair of waveforms which have the property that if transmitted separately, each correlated against a copy of itself and the results added then we do obtain a perfect spike. Such waveforms are called *complementary pairs*. We require of a pair of waveforms $\boldsymbol{w}_1(\boldsymbol{t})$

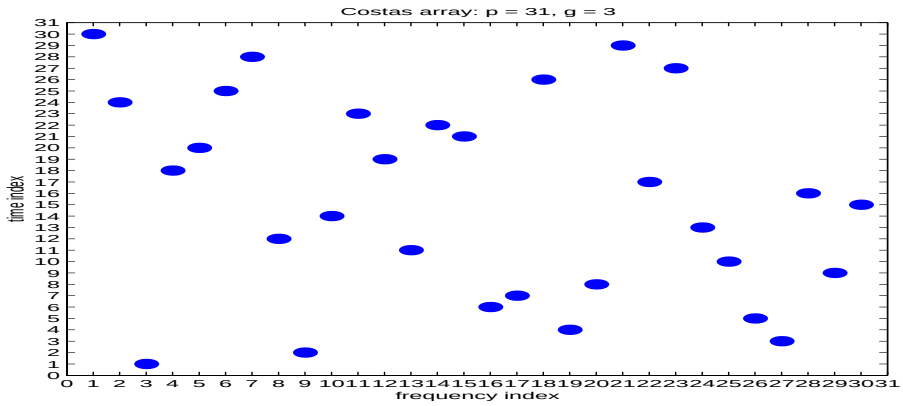

FIGURE 12. Costas array

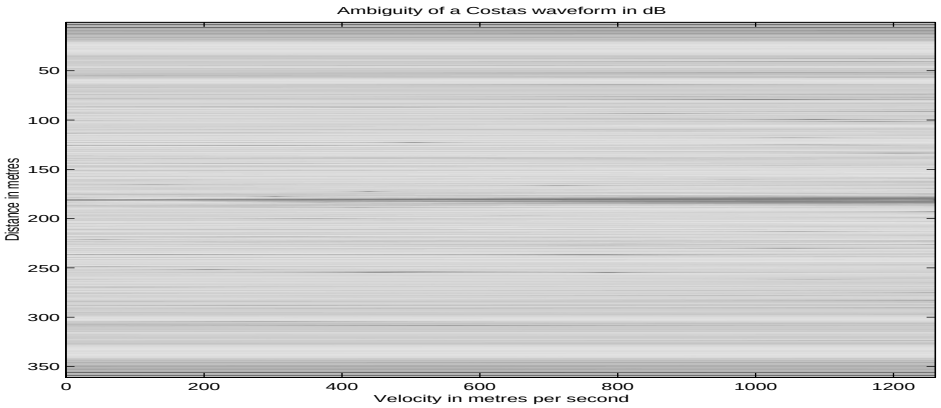

FIGURE 13. Costas array

and $w_2(t)$ that

$$\text{corr}(w_1, w_1)(t) + \text{corr}(w_2, w_2)(t) = \Delta_T(t),$$

where Δ_T is a short (triangular) spike at 0 of width T. This is equally expressed in terms of convolutions:

$$w_1 * \widetilde{w_1}(t) + w_2 * \widetilde{w_2}(t) = \Delta_T(t). \tag{3.5}$$

where $\widetilde{w}(t) = \overline{w(-t)}$. As we have said, it is possible to find such pairs.

In fact it is possible to find a pair of discrete codes p_1 and p_2 which are finite sequences of ± 1s satisfing

$$\text{corr}(p_1, p_1)(k) + \text{corr}(p_2, p_2)(k) = 0 \text{ except when } k = 0, \tag{3.6}$$

and from these easily construct waveforms satisfying (3.5). The classical construction of these is due to Golay and independently, Shapiro. It is an inductive construction starting with the two codes

$$p_1^{(1)} = [1, \quad 1]$$
$$p_2^{(1)} = [1, -1].$$

Then longer codes are constructed by the formula:

$$p_1^{(k)} = [p_1^{(k-1)}, \quad p_2^{(k-1)}]$$
$$p_2^{(k)} = [p_1^{(k-1)}, -p_2^{(k-1)}].$$

Our intention is to transmit the two waveforms in separate channels and on reception keep them separate and matched filter each against the corresponding transmitted form. The appropriate ambiguity function for such a pair of codes is the sum of the ambiguities of each code separately. By Moyal's identity, since p_1 and p_2 are orthogonal, so are χ_{p_1} and χ_{p_2}. After normalisation so that the sum of the energies in p_1 and p_2 equals 1, we have the following:

$$||\chi_{p_1} + \chi_{p_2}||^2 = ||\chi_{p_1}||^2 + ||\chi_{p_2}||^2 = ||p_1||^4 + ||p_2||^4 = \frac{1}{2}. \tag{3.7}$$

This shows that the ambiguities of complementary waveforms already have some chance of being more thumbtack-like than single channel waveforms, since the corresponding answer for a single channel waveform is 1, and the height at the centre $((0,0))$ is the same in each case. The ambiguity of a complementary pair is illustrated in Figure 14. In fact at dopplers of interest in typical applications (up to mach 3, say) complementary pairs are quite doppler tolerant. What is remarkable is that at all dopplers the range ambiguity is zero outside half the range of the waveform.

It may seem that complementary sequences offer ambiguity properties superior to single waveforms. However there are serious issues involved in the implementation of these waveforms. How does one maintain a significant separation of the two complementary sequences? If they are separated in time, then the target must remain coherent over the time span of the transmission for the advantages to be worthwhile. Separation in frequency, while feasible, also carries with it disadvantages. The responses of targets are frequency dependent as is attentuation through the atmosphere. Moreover, for extended targets, the effects of doppler change with different frequencies. These result in degradation from the ideal situation described here. The complexity of the electronics is also increased. One final observation: ultimately there is really only one signal transmitted, whether it be spread over a long time period (as time separation would require) or have a high bandwidth (in the case of frequency separation). The use of complementary waveforms is merely a device to manipulate the ambiguity function of a single waveform to push the lobes into more desired parts of the range-doppler plane.

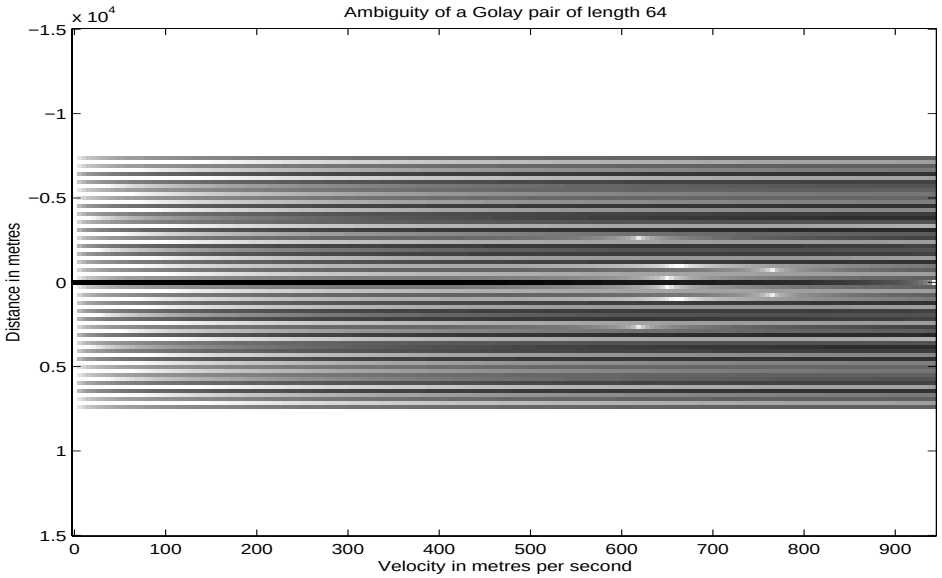

FIGURE 14. Ambiguity of a Golay pair of length 64

It is possible to go beyond just two waveforms. We recall the construction of the PONS matrix. This works by analogy with the Golay pairs but produces 2^n waveforms of length 2^n which are complementary in pairs, so that one could use any number of complementary sequence pairs separated by frequency or time or a mixture of both. The recursive method of construction is to take two "parent" codes s_1 and s_2 of length l, say, and use them to construct four "children" of double the length:

$$\begin{aligned} s_1 \cdot s_2 \\ s_1 \cdot -s_2 \\ s_2 \cdot s_1 \\ -s_2 \cdot s_1, \end{aligned}$$

where $\cdot$ indicates concatenation. It is relatively straightforward to see that, if the two parent codes are orthogonal, so are the four children. Classical PONS is now formed by starting with the two codes

$$\begin{matrix} 1 & 1 \\ 1 & -1 \end{matrix}$$

and repeatedly applying the above construction.

Of course the issues raised in the preceding paragraph are correspondingly exacerbated, but for N waveforms the power in the ambiguity is now (by a repeat of calculation (3.7))

just $1/N$. As Figure 15 illustrates, when we use half of the PONS matrix of length 64 the improvement in ambiguity is remarkable. We have perfect range sidelobes at all dopplers.

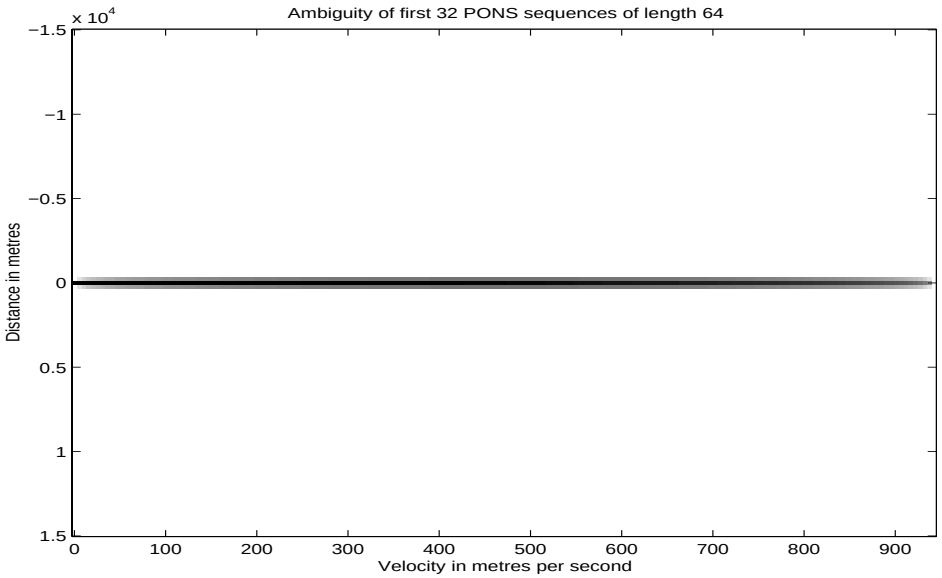

FIGURE 15. Ambiguity of the first 32 rows of the PONS matrix of length 64

References

[1] R.A. Altes. *Methods of wide-band signal design for radar and sonar systems*. PhD thesis, Univ. of Rcohester, Rochester, N.Y., 1970.

[2] R.A. Altes. Wide-band, proportional bandwidth wigner-ville analysis. *IEEE Trans. Acoustics, Speech, Signal Processing*, 38(6):1005–1012, June 1990.

[3] B. Borden. *Radar Imaging of Airborne Targets*. Institute of Physics Publishing, Bristol and Philadelphia, 1999.

[4] Wilcox C.H. The synthesis problem for radar ambiguity functions. Technical Report 157, MRC Tech. Summary Report, 1960. republished in Radar and Sonar part I vol 32 I.M.A. vol in Math. and its Appl,1991, 229–260.

[5] de Buda R. Signals that can be calculated from their ambiguity function. *IEEE Trans. Information Theory*, IT16:195–202, 1970.

[6] H. Dym and H.P. McKean. *Fourier Series and Integrals*. Academic Press, New York and London, 1972.

[7] G. Garrigós, Ph. Jaming, and J.-B. Poly. Zéros de fonctions holomorphes et contre-exemples en théorie des radars. To appear.

[8] S. Golomb. Costas arrays.

[9] Bueckner H.F. Signals having the same ambiguity functions. Technical Report 67-C-456, General Electric, Research and Development Center, Schnectady, N.Y., 1967.

[10] Ph. Jaming. Phase retrieval techniques for radar ambiguity functions. to appear.

[11] Auslander L. and Gretner I. Wide-band ambiguity functions and the $ax + b$ group. In *Signal processing part I: Signal Processing Theory*, volume 22 of *I.M.A. vol in Math and its Appl*, pages 1–12, 1990.
[12] Auslander L. and Tolimieri R. radar ambiguity functions and group theory. *SIAM J. Math Anal*, 16:577–6, 1985.
[13] Scharf L. *Statistical Signal Processing*.
[14] F.E. Nathanson. *Radar design principles*. McGraw-Hill, New York, 1969.
[15] Jaming Ph. *Trois problèmes d'analyse harmonique*. PhD thesis, Université d'Orléans, 1998.
[16] M.I. Skolnik. *Introduction to Radar Systems*. McGraw-Hill, New York, 2nd edition, 1981.
[17] J. Speiser. Wide-band ambiguity functions. *IEEE trans. Inform. Theory*, 13:122–3, 1967.
[18] Poor V. *Detection and Estimation Theory*. Springer-Verlag.
[19] Schempp W. *R. Math. Rep.Acad. Sci.Canada*, 4, 1984.
[20] Schempp W. *R. Math. Rep.Acad. Sci.Canada*, 5, 1984.
[21] Schempp W. *R. Math. Rep.Acad. Sci.Canada*, 6, 1984.
[22] Schempp W. *Math. Appl.*, pages 217–260, 1984.
[23] L.B. White. The wide-band ambiguity function and altes' q-distribution: constrained synthesis and time-scale filtering. *IEEE Trans. Inform. Theory*, 38(2):886–892, March 1992.