Chapter 11

Time Delay Estimation

MANUEL T. SILVIA
Allied Signal Aerospace Company
Bendix Oceanics Division
Sylmar, California 91342

INTRODUCTION

The estimation of time delay (or time difference) has become an important problem in digital signal processing. For example, an ideal active radar or sonar, which employs a single omnidirectional sensor to transmit and receive signals, can measure the time difference between the time a signal was transmitted and the time a backscattered signal was received to estimate the range of a radar or sonar target. If these active systems employ an array of omnidirectional sensors, then, in addition to estimating range, the time delays between these sensors can also be used (i) to focus the transmitted energy in a specified direction and (ii) to estimate the direction of a radar or sonar target [see Fig. 11.1(a)]. On the other hand, an ideal passive sonar or radar generally employs an array of omnidirectional sensors (at least two) for the sole purpose of receiving acoustic or electromagnetic radiation from distant targets. When the radiation is received at the passive array, time delay estimation methods are used to estimate the travel time of an acoustic or electromagnetic wavefront between the sensors. Wavefront travel time gives the range and direction of a radiating target [see Fig. 11.1(b)]. In either case the estimation of these time delays is often corrupted by ambient and receiver-generated noise, and multipath and finite-length observation intervals [1].

In seismology an underground disturbance creates seismic waves. Seismic detectors, located on the earth's surface, record these waves at different times. By estimating the time delays associated with the propagation of these seismic waves to the various detectors, a seismologist can decide whether an underground disturbance was natural or created (e.g., disturbances that are localized deep within the earth are more likely to be natural, whereas shallow disturbances are more likely to be created [2]). In the speech and hearing area time delay estimation has been explored to measure the travel time of sound waves from the

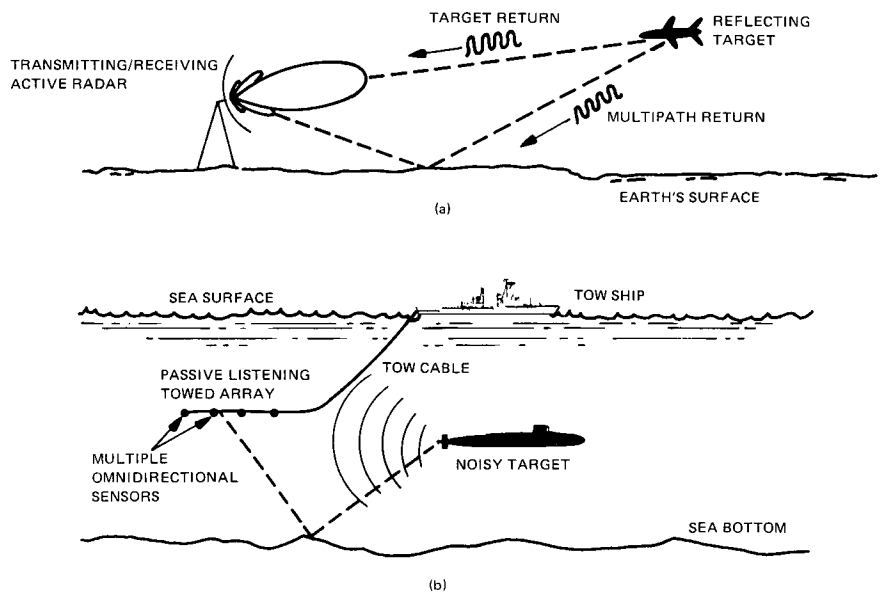


Fig. 11.1. (a) Active radar system. (b) Passive sonar system.

external ear to the eardrum and other parts of the internal ear. Analyzing these time delays helps researchers understand the hearing mechanisms of man and other species [3]. In biomedicine the electroencephalogram (EEG) represents the spontaneous electrical activity of the brain as measured from electrodes placed on a patient's scalp. When the patient is subjected to a sensory stimulus, electrical signals in the EEG are observed. By studying the time delays in these electrical signals, biomedical researchers can help doctors improve their neurological assessment of patients [4].

Hyperbolic location systems, commonly referred to as time difference of arrival systems, locate an active source or transmitter by processing signal arrival-time measurements at three or more passive stations [5]. The measurements at these passive stations are sent to another station, designated the master station, that does the time delay processing. The basic idea behind hyperbolic location systems is as follows. The arrival-time measurements at two stations are combined to produce a relative arrival time that, in the absence of noise and other interference, restricts the possible transmitter location to a hyperboloid with the two stations as foci. The transmitter location is then estimated from the intersection of three or more independently generated hyperboloids determined from at least four stations. If the transmitter and the passive stations lie in the same plane, then the transmitter location is estimated from the intersections of two or more hyperbolas determined from three or more stations. Figure 11.2 illustrates two hyperbolas, each of which has two branches, derived from the measurements at three stations. Notice that the two hyperbolas have two points

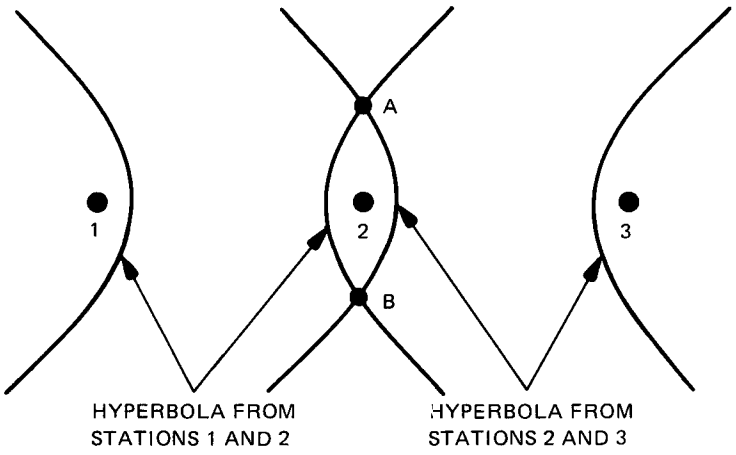


Fig. 11.2. Planar hyperbolic location system.

of intersection. We can resolve the resulting ambiguity in transmitter location by using a priori information about the location or a fourth station to generate an additional hyperbola.

If we now consider the three or more passive stations to be active sources and the one transmitter to be a passive station or receiver, then the single passive receiver can determine its own position by processing the signal arrival-time measurements from the three or more active sources. Similar to the previous hyperbolic location system, the position of the passive receiver will be determined by the intersection of three or more hyperbolas. This simple idea describes the fundamental principle behind a military and commercial navigation system called LORAN C [6]. Today, nearly every yachtsman has a commercial LORAN C receiver that estimates the signal arrival-time differences or time delays and converts these measurements into latitude and longitude coordinates, which are useful for navigational purposes (refer to Fig. 11.3). Thus, time delay estimation forms the basis of nearly all hyperbolic location and/or navigation systems.

As we have just seen, the time delay estimation problem spans the fields of radar, sonar, seismology, speech and hearing, biomedical research, and hyperbolic localization, just to name a few. The main purpose of this chapter is to provide a summary of the basic principles behind time delay estimation. In Section II we consider the time delay problem for active sensors (e.g., active radars or sonars). In Section III we discuss the time delay problem for passive sensors (e.g., passive sonars or radars). In Section IV we concern ourselves with the statistical theory of correlation and its relationship to the time delay estimation problem. In Section V we consider the implementation of some time delay estimation algorithms using the fast Fourier transform (FFT). Section VI, provides a table of various algorithms and some numerical results.

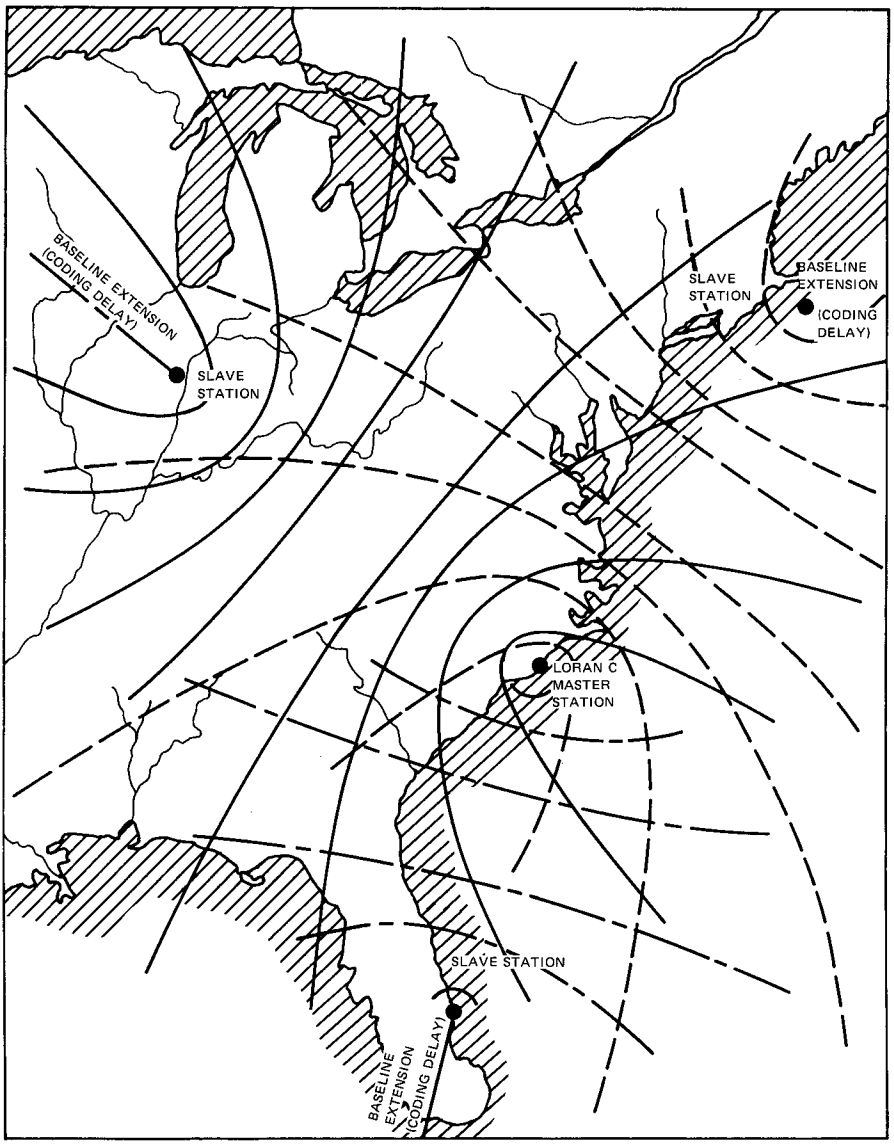


Fig. 11.3. LORAN C chart showing hyperbolas used for time delay navigation.

TIME DELAY ESTIMATION FOR ACTIVE SENSORS II

The Time Delay Estimation Problem for a Single Omnidirectional Active Sensor A

In general, an active radar or sonar system contains an array of omnidirectional sensors, a transmitter, and a receiver. The sole purpose of the transmitter is to excite the sensors with electrical signals. In radar the sensors convert these electrical signals into electromagnetic energy and radiate electromagnetic waves. In sonar the sensors convert these electrical signals into acoustic energy and radiate acoustic waves. In both cases the radiated waves eventually strike a target. However, only some of the transmitted energy returns to the radar or sonar sensors. Since the sensors are assumed to be reciprocal devices, they convert the returned waves back to electrical signals. The sole purpose of the receiver is to filter or prepare these electrical signals for further signal processing.

If the active system employs only one omnidirectional reciprocal sensor, then it radiates the same amount of energy in all directions and it receives energy in the same way for every direction. See Fig. 11.4 for a graphical illustration.

In the theory of analog and digital filters the most trival of all linear time-invariant filters is the one whose frequency response treats all the frequencies of an input signal in the same way. That is, its frequency response has a constant gain of unity and a constant phase shift of zero. This filter is sometimes referred to as the trival all-pass filter. Its frequency response is shown in Fig. 11.5. Thus, we can think of the single omnidirectional reciprocal sensor as a trival all-pass *spatial filter*. That is, this sensor has a spatial response that treats all the angles ϕ associated with an input signal in the same way. Analogously, its radiation and receive spatial response has a constant gain of unity and a constant phase shift of zero. The radiation and receive spatial response of a single omnidirectional reciprocal sensor is shown in Fig. 11.6. Although the spatial gain function $G(\phi)$ can be plotted in either Cartesian or polar coordinates [Fig. 11.6(a), (c), respectively], a commonly used graphical representation of the spatial gain function is the *beam pattern*, as defined in Fig. 11.6(d).

As we have just seen, a single omnidirectional active sensor cannot focus its transmitted energy in a specified direction. Since we have assumed that the sensor is a reciprocal device, it cannot tell which direction the returned or backscattered energy came from. Thus, any radar or sonar that uses a single omnidirectional reciprocal sensor to transmit and receive cannot estimate the direction of a reflecting target. However, by estimating the time delay or time difference between the time a signal was transmitted and the time a signal was received, a "single-sensor" active radar or sonar can estimate the reflecting target's range. Hence, the time delay estimation problem for a single omnidirectional active sensor is equivalent to the range estimation problem. Let us elaborate.

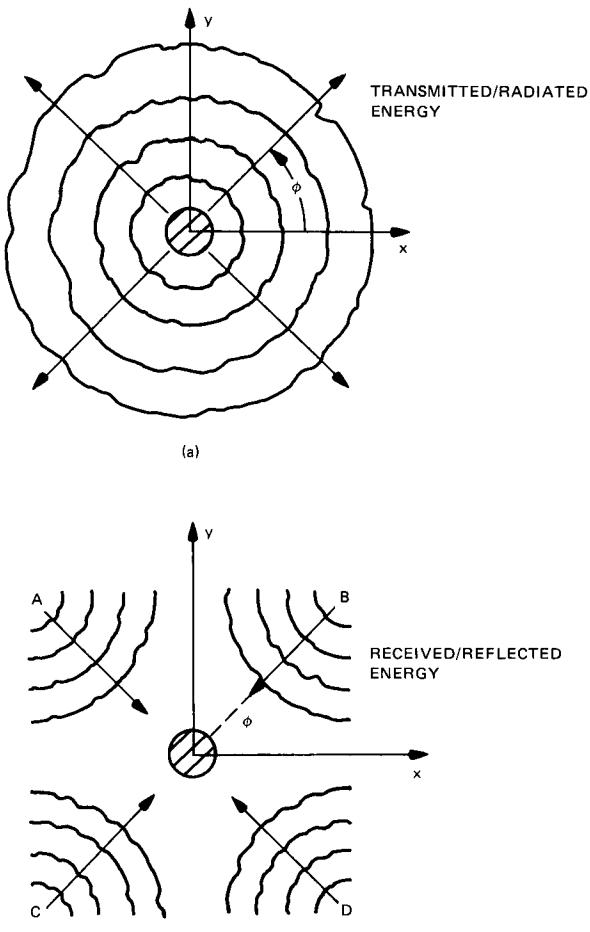


Fig. 11.4. Graphical illustration. (a) A single omnidirectional active sensor radiates the same amount of energy in all directions ϕ. (b) A single omnidirectional active sensor, which is assumed to be a reciprocal device, receives reflected energy in the same way for every direction ϕ. That is, this sensor would receive the same reflected energy regardless of where the target is located (e.g., at points A, B, C, or D). Thus, it has no spatial discrimination.

Suppose the transmitter creates an electrical signal, say $s_T(t)$, which excites a single omnidirectional sensor. The resulting electromagnetic or acoustic waves radiate outward [Fig. 11.4(a)] toward a reflecting target. When the waves strike the target, a complicated backscattering process generates a reflected or backscattered wave that returns to the sensor. The sensor then converts this returned wave to an electrical signal, say $s_R(t)$. If the ambient noise of the medium and the noise generated in the receiver can be combined into an additive process,

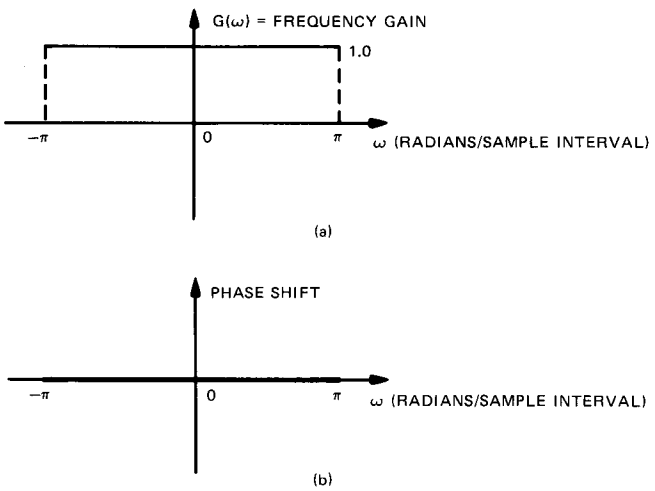


Fig. 11.5. Frequency response of a trivial all-pass digital filter: (a) frequency gain function; (b) frequency phase shift function.

then the electrical signal, as seen by the receiver, is

$$\underline{z}(t) = s_R(t) + \underline{v}(t)$$

where $\underline{v}(t)$ is the electrical signal that represents the combined additive noise process. In general, we note that $s_R(t)$ depends on the physical properties and shape of the reflecting target, $s_T(t)$, and the spatial response of the sensor. For example, it could happen that the direction in which the active sensor decides to focus its transmitted energy may not coincide with the target's direction, which would cause a reduction in the amplitude of $s_R(t)$ (refer to Fig. 11.7). However, since we are considering a single omnidirectional reciprocal sensor, the amplitude of $s_R(t)$ will not be affected by the spatial response of the sensor.

We are now ready to state the time delay estimation problem for the case of a single omnidirectional active sensor. That is, given $\underline{z}(t)$ for $t_p \leq t \leq t_p + t_0$, where t_0 is the length of the observation interval, the problem is to estimate the time delay between the time $s_T(t)$ was transmitted, say $t = 0$, and the time $s_R(t)$ was received, say $t = \tau$. See Fig. 11.8.

In the absence of additive noise [e.g., $\underline{v}(t) = 0$], estimating the time delay τ is not difficult, because the onset or leading edge of $s_R(t)$ is easily detected when no interference or noise is present, provided the amplitude of $s_R(t)$ is not too small. For the noise-free case this onset detection problem seems to be independent of the shape and amplitude of $s_R(t)$. However, when noise is present, then detecting the leading edge of $s_R(t)$ becomes more difficult. By examining Fig. 11.8(b), we can intuitively argue that the level of difficulty increases as the signal-to-noise ratio (SNR) decreases. That is, the onset of $s_R(t)$ becomes more difficult to detect when the noise amplitudes approach the signal amplitudes. Moreover, this onset

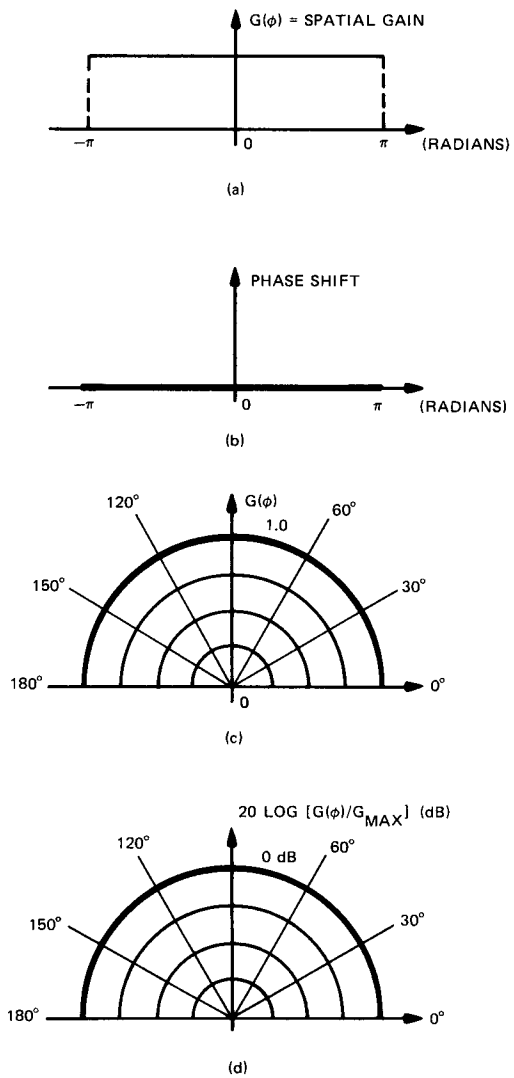


Fig. 11.6. Spatial response. (a) spatial gain function $G(\phi)$ for a single omnidirectional reciprocal sensor. Graph is in Cartesian coordinates; (b) spatial phase-shift function for a single omnidirectional reciprocal sensor; (c) spatial gain function plotted in polar coordinates. Graph is for $0 \leq \phi \leq \pi$; (d) the function 20 Log $[G(\phi)/G_{\max}]$, plotted in polar coordinates, is commonly called the beam pattern. Graph is for $0 \leq \phi \leq \pi$.

detection problem could be made less difficult if, for example, $s_R(t)$ had a large amplitude and a fast rise time or a sharp leading edge. This implies that our success in detecting $s_R(t)$'s leading edge, in the presence of noise, depends on the shape and amplitude of $s_R(t)$. In any event, we can intuitively argue [see Fig. 11.8(b)] that the accuracy of our time delay estimate will be greater if the

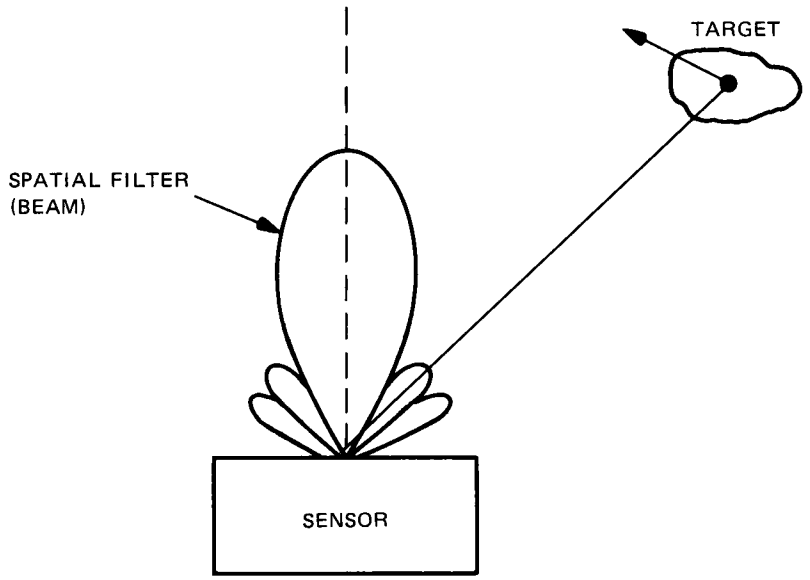


Fig. 11.7. Amplitude of reflected signal (echo) is affected by spatial response of sensor.

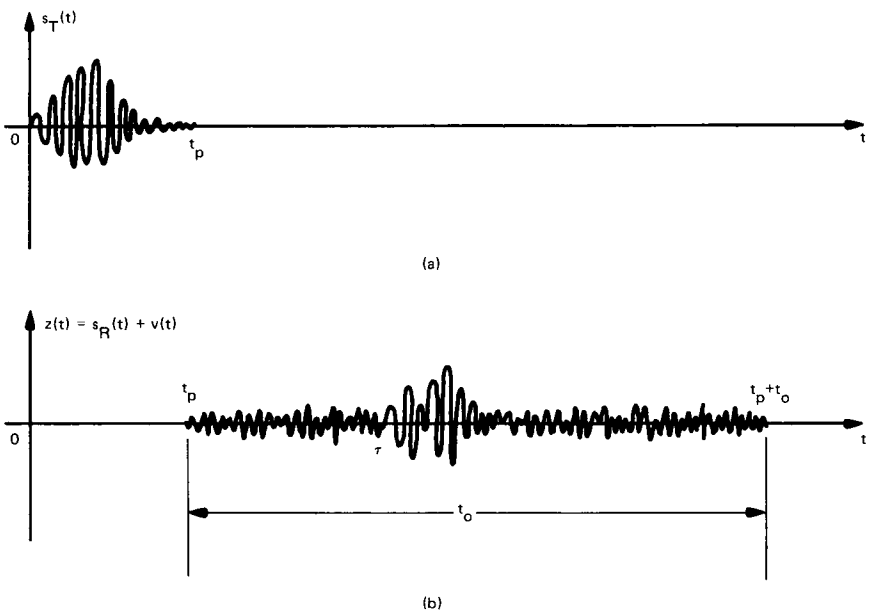


Fig. 11.8. Time delay estimation problem for a single omnidirectional active sensor: (a) transmitted signal; (b) backscattered signal as seen by the receiver.

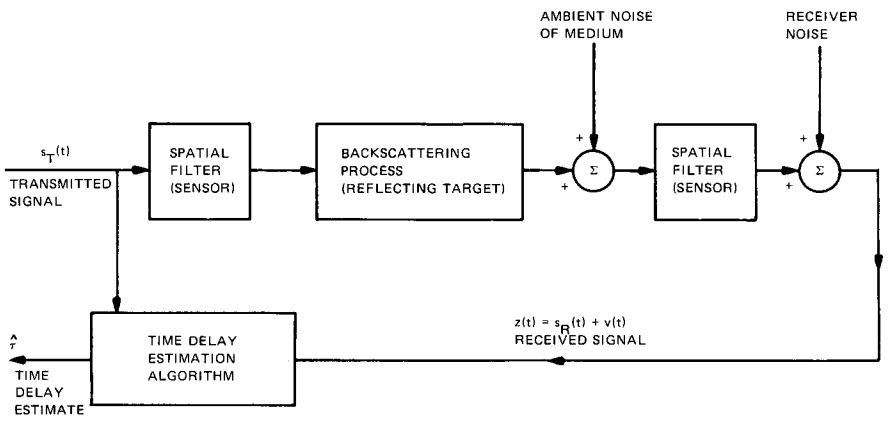


Fig. 11.9. Block diagram of the time delay estimation problem for a single omnidirectional active sensor.

rise time of $s_R(t)$ is small and the SNR is high. In the following sections we will quantify these intuitive notions.

We close this section with Fig. 11.9, a block diagram of the time delay estimation problem for a single omnidirectional active sensor.

B The Time Delay Estimation Problem for an Array of Omnidirectional Active Sensors

In the previous section we considered the most trivial of all active systems—that is, the one with a single omnidirectional reciprocal sensor. This system had no spatial discrimination capability. In this section we consider a more practical active system, one that employs an array of omnidirectional reciprocal sensors. As we will soon see, active systems that use more than one omnidirectional reciprocal sensor have the ability to perform spatial discrimination or spatial filtering.

1 *Two-Sensor Array*

Let us begin by considering the simplest of all arrays—the linear array containing only two omnidirectional reciprocal sensors separated by a distance l. Figure 11.10 gives a pictorial description of this array.

To explain how the two-sensor array can provide spatial discrimination, let us consider the following experiment. We assume that the radar or sonar transmitter is capable of generating the waveform or pulse

$$s_T(t) = \begin{cases} a(t)\cos[2\pi F_0 t + \theta(t)], & 0 \le t \le t_p \\ 0 \quad \text{otherwise} \end{cases} \tag{11.1}$$

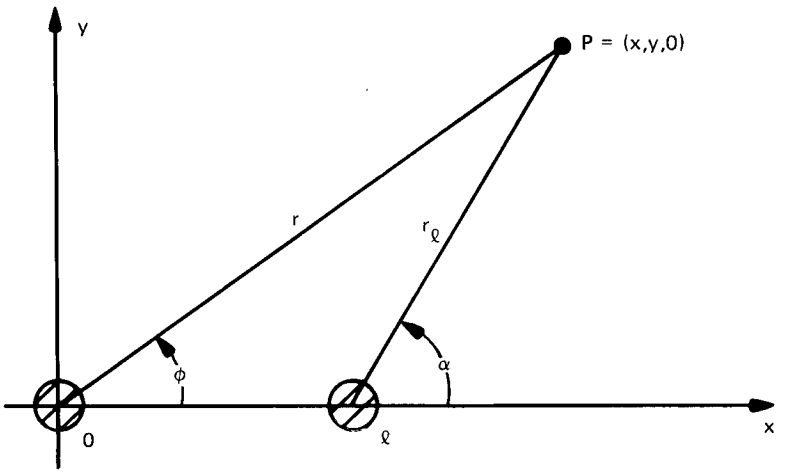


Fig. 10. A simple array with two omnidirectional reciprocal sensors a distance l apart. The z-axis is pointing out of the page.

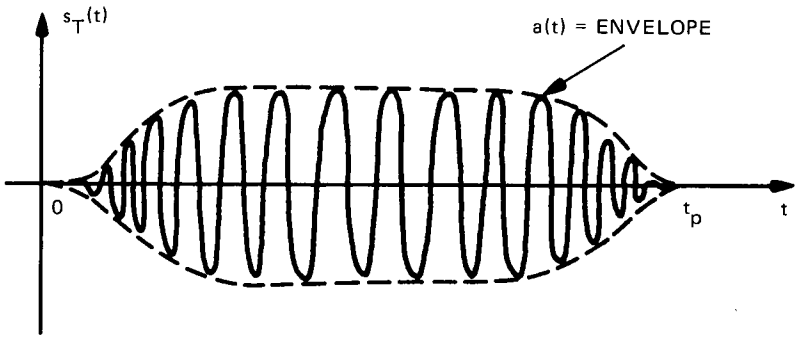


Fig. 11.11. Description of a radar or sonar pulse whose amplitude and phase are both slowly varying with respect to the sine-wave fluctuations.

where $a(t)$ is the envelope, F_0 is the carrier frequency in hertz, $\theta(t)$ is the phase modulation, and t_p is the pulse duration. Figure 11.11 is a graphical description of Eq. (11.1). At $t = 0$ the omnidirectional sensor located at $x = 0$ is excited by $s_T(t)$. After a time delay δ the omnidirectional sensor located at $x = l$ is also excited by $s_T(t)$. If the two sensors radiate waves into a homogeneous, isotropic, and lossless medium, then the resulting waveform at point P in Fig. 11.10 has the approximate form [7]

$$s_P(t) \simeq \frac{1}{r} s_T\left(t - \frac{r}{c}\right) + \frac{1}{r} s_T\left(t - \frac{r}{c} + \Delta - \delta\right) \tag{11.2}$$

where c is the wave propagation speed in the medium of interest and

$$\Delta = \frac{r}{c} - \frac{r_l}{c} \simeq \frac{l}{c} \cdot \cos\phi \tag{11.3}$$

represents the *travel time difference*. Equations (11.2) and (11.3) assume that $l \ll r$, the lines r and r_l are approximately parallel (e.g., $\alpha \simeq \phi$), and point P is "very far" from the array, so $s_P(t)$ is essentially a plane wave.

Notice that Δ depends on the sensor spacing l and the array angle ϕ. For fixed l, Δ is controlled by ϕ. Now if the $x = 0$ sensor is always excited at $t = 0$, then at what later time should we excite the $x = l$ sensor? That is, how should we select the excitation or *intersensor time delay* δ? Let us proceed to answer this question.

Suppose that the radar or sonar wishes to transmit most of its energy in a specific direction, say $\phi = \phi_0$. It turns out that $s_P(t)$ [refer to Eq. (11.2)] experiences "constructive interference" along a specific direction ϕ_0 when the time delay δ satisfies the condition

$$\delta = \Delta_0 = \frac{l}{c}\cos\phi_0 \tag{11.4}$$

In other words, the radiated energy will be a maximum along the line $\phi = \phi_0$ when δ satisfies Eq. (11.4). To see this let us consider the angular distribution of radiated energy under Eq. (11.4). Mathematically, the radiated energy at a "distant" point P for all angles ϕ is

$$\begin{aligned} E_P &\equiv \int |s_P(t)|^2\, dt = \int |S_P(F)|^2\, dF \\ &\equiv \int |S_T(F)|^2 |A(F,\phi)|^2\, dF \end{aligned} \tag{11.5}$$

which follows from the Fourier transform of Eq. (11.2). Here, the quantity

$$A(F,\phi) \equiv \sum_{n=0}^{1} a(n) e^{j2\pi nF\Delta} \tag{11.6}$$

describes a *spatial filter*, where

$$a(0) \equiv 1, \qquad a(1) \equiv e^{-j2\pi F\delta} \tag{11.7}$$

Δ is given by Eq. (11.3), and δ is given by Eq. (11.4). Since most radars and sonars transmit narrowband waveforms, most of the energy associated with $s_T(t)$ is concentrated around the carrier frequency F_0. Thus, we can say

$$E_P \propto |A(F_0,\phi)|^2 \tag{11.8}$$

From Eqs. (11.3), (11.4), (11.6), and (11.7) we conclude that the radiated energy E_P is a maximum along the line $\phi = \phi_0$ when $\delta = \Delta_0$. When $\delta = \Delta_0$, the line $\phi = \phi_0$ is generally referred to as the *main response axis* (MRA) of the array.

The *beam pattern*, defined by

$$B(\phi) \equiv 20\log\left[\frac{G(\phi)}{G_{\max}}\right] \tag{11.9}$$

where

$$G(\phi) \equiv |A(F_0, \phi)| = \text{spatial gain}$$

describes the angular distribution of radiated energy for a given spacing l, analog carrier frequency F_0, and intersensor time delay δ. $B(\phi)$ is an important quantity in the design of radar and sonar arrays.

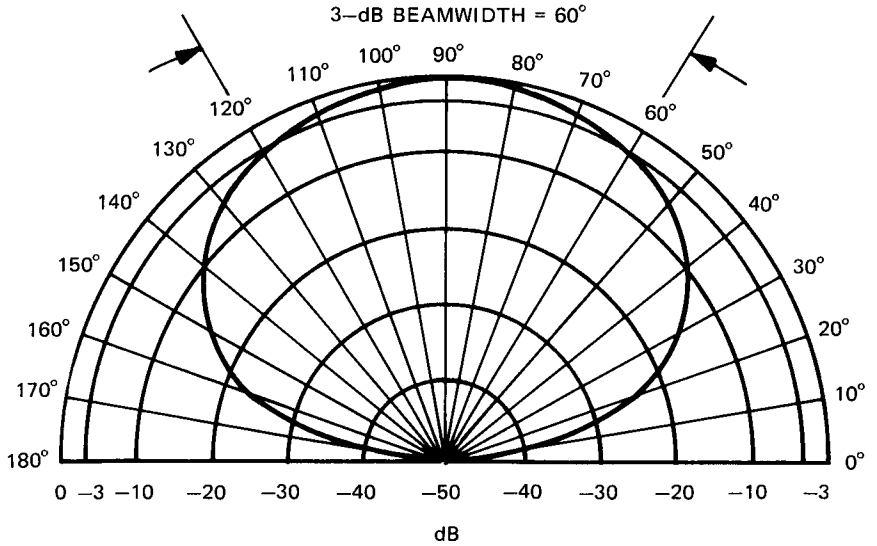
Example 1. We wish to design a two-sensor linear sonar array that will transmit most of its energy in the direction $\phi = \phi_0 = 90°$. If $F_0 = 5000$ Hz = 5 kHz and $c = 5000$ ft/s, how do we select the intersensor time delay δ and the sensor spacing l?

From Eq. (11.4) we see that $\phi = \phi_0 = 90°$ gives $\delta = \Delta_0 = 0$ s for any array length l. However, $\phi_0 = 270°$ also gives $\delta = 0$ s for any spacing l. This means that for $\delta = 0$ s this array will provide maximum radiated energy along *two* directions ($\phi_0 = 90°$ and $\phi_0 = 270°$), so it has two MRAs. Hence, we cannot focus this array in only *one* direction. Nevertheless, our design would select $\delta = 0$ s, so we should excite both sensors at the same time, say $t = 0$, with the same narrowband waveform, say $s_T(t)$. Let us now consider the selection of l.

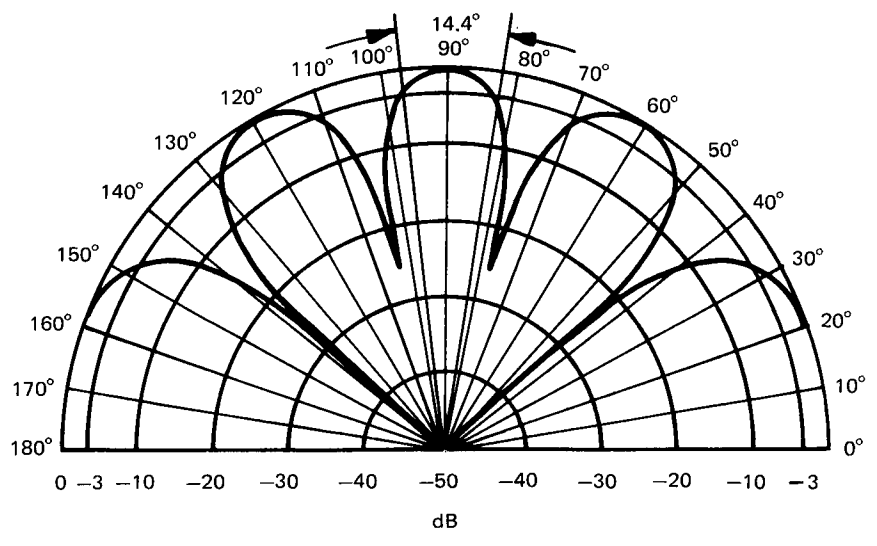
If we let $l = 0.5$ ft, the resulting beam pattern for $0° \leq \phi \leq 180°$ is shown in Fig. 11.12(a). [The beam pattern for $180° \leq \phi \leq 360°$ is the mirror image of Fig. 11.12(a).] The angular sector defined by the -3-dB point to the left of the MRA (e.g., consider the MRA at $\phi_0 = 90°$) and the -3-dB point to the right of the MRA is called the 3-dB or *half-power beamwidth*. For $l = 0.5$ ft the beamwidth is 60°. Thus, this array can focus energy in two main beams or mainlobes; one 60° wide beam is aimed in the intended direction $\phi_0 = 90°$, and the other 60° wide beam is aimed in the direction $\phi_0 = 270°$. Although these main beams are very wide, this two-sensor array still provides more spatial discrimination or spatial filtering than the single omnidirectional sensor does [compare Figs. 11.6(d) and 11.12(a)].

Suppose we increase the sensor spacing to $l = 2$ ft. The resulting beam pattern for $0° \leq \phi \leq 180°$ is shown in Fig. 11.12(b). [The beam pattern for $180° \leq \phi \leq 360°$ is the mirror image of Fig. 11.12(b).] For this case we still have two main beams or mainlobes at the MRAs $\phi_0 = 90°$ and $\phi_0 = 270°$. Notice that the beamwidth is 14.4°, which is considerably narrower than when $l = 0.5$ ft. The $l = 2$-ft pattern did produce a narrower beamwidth than the $l = 0.5$-ft pattern, but it introduced *grating lobes*. Here, a grating lobe is defined as a radiation beam or lobe in any direction, other than the intended MRA directions, that produces the same maximum radiation levels as the MRAs. Thus, the $l = 2$-ft array has grating lobes at $\phi = 0°$, 60°, 120°, 180°, 240°, and 300° [Fig. 11.12(b)]. Clearly, these grating lobes are undesirable.

From the above analysis we see that for a fixed wavelength λ the beamwidth of the mainlobe got smaller as the sensor spacing l got larger. For $l = \lambda/2 = 0.5$ ft we had a 60° wide main beam but no grating lobes. For $l = 2\lambda = 2$ ft we had a 14.4° wide main beam but several grating lobes. It turns out that the sensor spacing l



(a)



(b)

Fig. 11.12. Beam patterns. (a) Plot is for $0° \leq \phi \leq 180°$; for $180° \leq \phi \leq 360°$, we get the mirror image. Note: $l = \lambda/2 = 0.5$ ft, no grating lobes. (b) $l = 2\lambda = 2$ ft, grating lobes at $\phi = 0°$, 60°, 120°, 180°, 240°, and 300°.

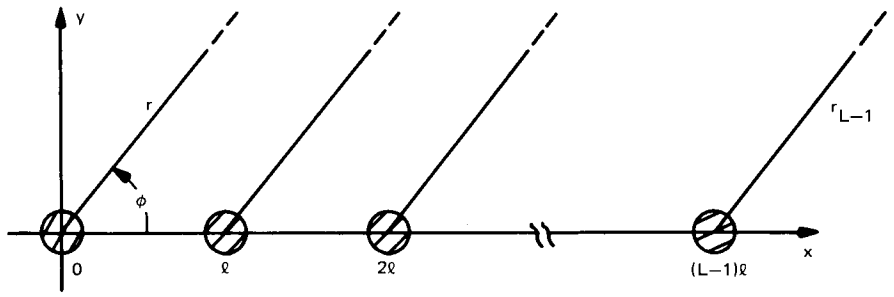


Fig. 11.13. A linear array with L equally spaced omnidirectional reciprocal sensors: l = sensor spacing; $r \simeq (L-1)l\cos\phi + r_{L-1}$.

should be chosen so that $l \leq \lambda/2$. For this choice of l the grating lobes will be eliminated. Thus, this two-sensor linear sonar array should be designed for $\delta = 0$ s and $l \leq \lambda/2$ ft. The narrowest beamwidth with no grating lobes occurs when $l = \lambda/2 = 0.5$ ft. In summary:

1. The intersensor time delay $\delta = 0$ s gives maximum radiated energy along *two* directions, $\phi_0 = 90°$ and $\phi_0 = 270°$. This linear array cannot focus its energy in *one* direction, say $\phi_0 = 90°$.
2. The sensor spacing $l = \lambda/2$ gives the narrowest 3-dB beamwidth with no grating lobes. This beamwidth is 60°.

Multisensor Arrays 2

In practice a two-sensor linear array with $l = \lambda/2$ does not provide enough spatial filtering. Thus, if the sensor spacing is constrained by $l = \lambda/2$, how can we achieve narrower beamwidths? One way is to simply add more sensors to the array. Let us elaborate.

Let us now consider an equally spaced linear array with L omnidirectional reciprocal sensors (Fig. 11.13). As before, we assume that the $x = 0$ sensor is the time reference sensor. That is, we assume that at $t = 0$ the omnidirectional sensor located at $x = 0$ is excited by $s_T(t)$. After a time delay δ_1 the omnidirectional sensor located at $x = l$ is excited by $s_T(t)$. After a time delay δ_2 the omnidirectional sensor located at $x = 2l$ is excited by $s_T(t)$, and so on. Finally, the sensor located at $x = (L-1)l$ is excited by $s_T(t)\delta_{L-1}$ s later. Here, $0 < \delta_1 < \delta_2 < \cdots < \delta_{L-1}$, and all the intersensor time delays δ_n $(n = 1, 2, \ldots, L-1)$ are referenced to $t = 0$. If all of these sensors radiate waves into a homogeneous and isotropic medium, then the resultant waveform at a distant point P can be approximated by

$$s_P(t) \simeq \frac{1}{r}s_T\left(t - \frac{r}{c}\right) + \frac{1}{r}s_T\left(t - \frac{r}{c} + \Delta - \delta_1\right) + \cdots + \frac{1}{r}s_T\left(t - \frac{r}{c} + (L-1)\Delta - \delta_{L-1}\right) \tag{11.10}$$

which is a generalization of Eq. (11.2). If the time delays δ_n satisfy the condition

$$\delta_n = n\left(\frac{l}{c}\cos\phi_0\right) = n\Delta_0 \qquad (n = 0, 1, 2, \ldots, L-1) \tag{11.11}$$

then $s_P(t)$ experiences constructive interference at a "distant" point along the line $\phi = \phi_0$. Hence, Eq. (11.11), which is a generalization of Eq. (11.4), represents the condition for maximum radiated energy in the direction ϕ_0. The angular distribution of energy is still given by Eqs. (11.5) and (11.8). However, for the L-sensor array we have the corresponding spatial filter

$$A(F, \phi) \equiv \sum_{n=0}^{L-1} a(n)e^{j2\pi nF\Delta} \tag{11.12}$$

where

$$a(n) \equiv e^{-j2\pi F\delta_n} \qquad (n = 0, 1, 2, \ldots, L-1) \tag{11.13}$$

Δ is given by Eq. (11.3), and δ_n is given by Eq. (11.11). Equations (11.12) and (11.13) are generalizations of Eqs. (11.6) and (11.7), respectively.

Example 2. Let us consider the linear sonar array in Example 1. If l/λ is constrained to be 0.5, then the element spacing is $l = 0.5$ ft, since $F_0 = 5$ kHz and $c = F_0\lambda = 5000$ ft/s. Thus, for $L = 20$ omnidirectional equally spaced sensors, the array length is $(L-1)l = 9.5$ ft. For $\phi_0 = 90°$ the timing sequence is $\delta_n = 0$ s for $n = 1, 2, \ldots, L-1$, so all the sensors are excited at $t = 0$. Figure 11.14(a) shows the beam pattern (for $0° \leq \phi \leq 180°$) that results when $\phi_0 = 90°$, $l/\lambda = 0.5$, and $L = 20$ sensors. [The beam pattern for $180° \leq \phi \leq 360°$ is the mirror image of Fig 11.14(a).]

We can easily see the effect of using $L = 20$ sensors instead of $L = 2$ sensors, all other factors being the same, by comparing Figs. 11.14(a) and 11.12(a), respectively. For the $L = 20$ case the beamwidth is 5°, whereas for the $L = 2$ case the beamwidth is 60°. Although the $L = 20$ sensor array generates a desirable 5° beamwidth, it also generates *sidelobes*. Here, a sidelobe is defined as a radiation beam or lobe in any direction other than the desired or intended direction $\phi = \phi_0$. The largest sidelobe occurs at about $\phi = 82°$. It is about 13.3 dB down from the mainlobe, which occurs at $\phi = \phi_0 = 90°$. Although sidelobes are not desirable, they are not as bad as the grating lobes encountered in Example 1.

Let us now show how we can "steer" this ($L = 20$) array in a specific direction by controlling the time delay sequence δ_n ($n = 1, 2, \ldots, 19$). Specifically, suppose we wish to steer the array or form a beam in the direction $\phi = \phi_0 = 53°$. For this case

$$\delta_n = n\left(\frac{l}{c}\cos\phi_0\right) = n60 \times 10^{-6}\ \text{s} = n60\ \mu\text{s}$$

for $n = 1, 2, \ldots, 19$. That is, at $t = 0$ we excite the $x = 0$ sensor with $s_T(t)$. At $t = 60\ \mu$s we excite the $x = l$ sensor with $s_T(t)$. At $t = 120\ \mu$s we excite the $x = 2l$ sensor with $s_T(t)$. Finally, at $t = 1140\ \mu$s we excite the $x = (L-1)l$ sen-

sor with $s_T(t)$. The resulting radiation beam pattern for $0° \leq \phi \leq 180°$ is shown in Fig. 11.14(b). [The beam pattern for $180° \leq \phi \leq 360°$ is the mirror image of Fig. 11.14(b).] Comparing Fig. 11.14(a) (b), we see how steering the array or forming beams effects the linear array's beam pattern (e.g., beamwidth and sidelobe structure).

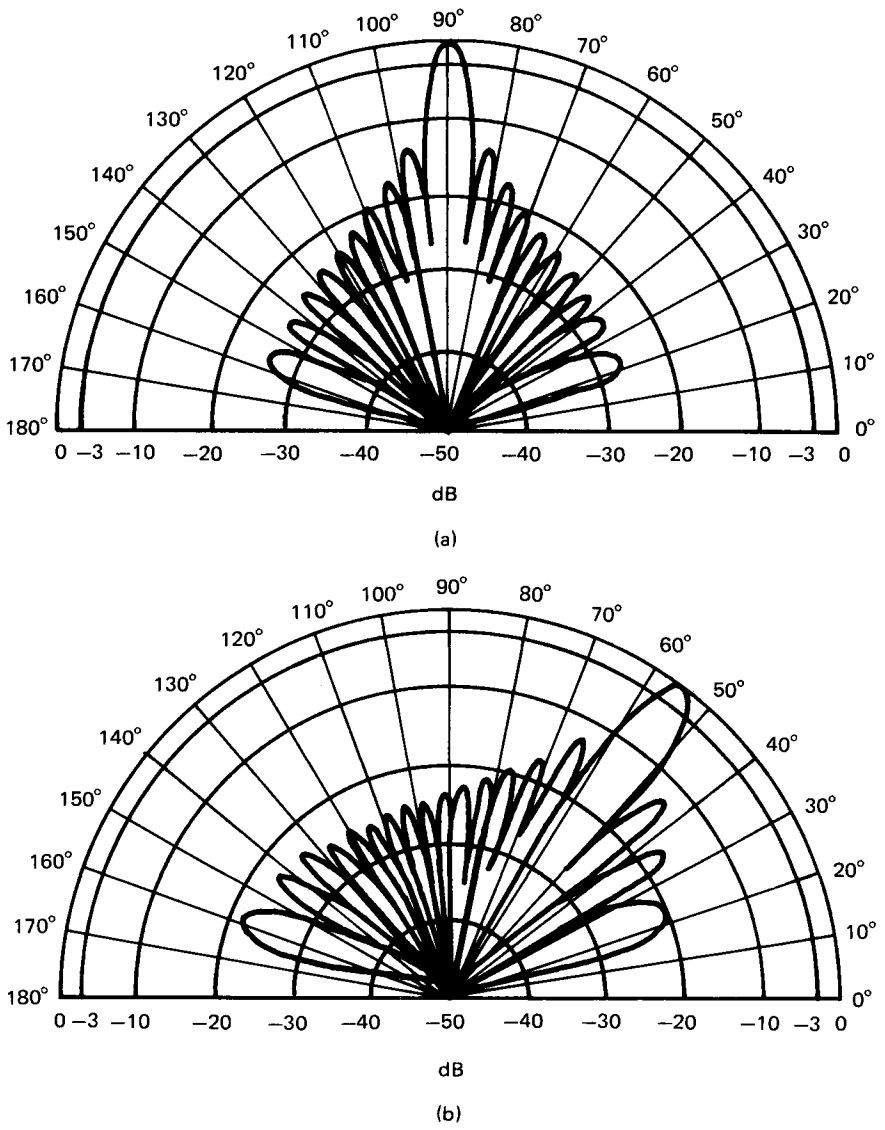


Fig. 11.14. Beam pattern for linear sonar array. (a) Plot is for $0° \leq \phi \leq 180°$; for $180° \leq \phi \leq 360°$, we get the mirror image. Note $l/\lambda = 0.5$, $\phi_0 = 90°(270°)$ = broadside, 3-dB beamwidth = 5°. (b) $l/\lambda = 0.5$, $\phi_0 = 53°(307°)$, 3-dB beamwidth 5°.

In the terminology of radar or sonar we can "form a beam" in a specific direction, say $\phi = \phi_0$, by exciting the sensors according to Eq. (11.11). As we have just seen, the intersensor or excitation time delays δ_n can be used to form beams in any direction. The ability to form beams in different directions by controlling the time delay sequence δ_n is called *beamforming*.

In sonar applications l/c is on the order of microseconds, whereas in radar applications l/c is on the order of nanoseconds, sometimes even fractions of nanoseconds. Therefore, the accurate control of δ_n in radar can be a problem. Because $s_T(t)$ is a narrowband waveform (refer to Eq. (11.1) and Fig. 11.11), we can change the phase of $s_T(t)$ rather than the time delay. Mathematically,

$$\begin{aligned} s_T(t - \delta) &= a(t - \delta)\cos[2\pi F_0(t - \delta) + \theta(t - \delta)] \\ &\simeq a(t)\cos[2\pi F_0 t + \theta(t) - 2\pi F_0 \delta] \end{aligned} \tag{11.14}$$

since $a(t)$ and $\theta(t)$ are assumed to be slowly varying waveforms. Hence, in radar, the steering or beamforming can be accomplished by *phase shifters*, which vary the phase, $-2\pi F_0 \delta$, in Eq. (11.14). By electronically phasing $s_T(t)$ at each sensor, we can effectively control the timing sequence δ_n and therefore steer or scan the beam. This idea forms the basis of *phased-array radar* [8].

In summary:

1. Adding more elements to the linear array produces a narrower beamwidth but generates *sidelobes.*
2. A linear array with $L = 20$ omnidirectional reciprocal sensors and $l/\lambda = 0.5$ can generate a 5° beam in two directions. It provides good spatial filtering in two directions but cannot provide spatial filtering in just a single direction.
3. In sonar the beamforming is generally done with time delay controls, whereas in radar the beamforming is done with electronic phase shift controls.

Example 3. In the design of radar or sonar arrays one usually specifies the analog carrier frequency F_0 (or wavelength λ) and a desired 3-dB beamwidth. These quantities will generally control the physical size of the array. The ratio l/λ for a linear array is generally set at 0.5, because $l = \lambda/2$ is the maximum sensor spacing allowed by the spatial version of Nyquist's sampling theorem. Moreover, the choice $l = \lambda/2$ also prevents grating lobes. Thus, the physical extent of a linear array will be $(L - 1)\lambda/2$, where L is determined by the design specifications F_0 and the 3-dB beamwidth.

Hence an important quantity in the design of linear arrays is the spatial filter $A(F, \phi)$ in Eq. (11.12). For a fixed analog frequency F_0 and $l = \lambda/2$, Eq. (11.12) can be rewritten as

$$A(\omega) = \sum_{n=0}^{L-1} a(n)e^{jn\omega} \tag{11.15}$$

where $\omega = \pi \cos \phi$,

$$a(n) = e^{-jn\omega_0} \qquad (n = 0, 1, 2, \ldots, L - 1) \tag{11.16}$$

and $\omega_0 = \pi \cos \phi_0$. So far we have assumed that

$$|a(n)| = \begin{cases} 1 & \text{for } 0 \leq n \leq L-1 \\ 0 & \text{otherwise} \end{cases} \tag{11.17}$$

and

$$\angle a(n) = \begin{cases} -n\omega_0 & \text{for } 0 \leq n \leq L \\ 0 & \text{otherwise} \end{cases} \tag{11.18}$$

That is, the magnitude of $a(n)$ is unity, and the phase angle of $a(n)$ is linear in n. To properly form beams for a linear array, we want the phase angle of $a(n)$ to satisfy Eq. (11.18) and the phase spectrum of $A(\omega)$ to be linear in ω. This implies that the magnitude of $a(n)$ can be any positive sequence that is symmetric over $0 \leq n \leq L-1$. Hence, $|a(n)|$ need not be limited to the positive uniform sequence Eq. (11.16). It turns out that we can judiciously select the positive symmetric sequence $|a(n)|$ to obtain reduced sidelobe levels at the expense of a modest increase in beamwidth. The values of $|a(n)|$ that accomplish this are called the *sensor shading factors.*

Except for a difference in the sign of ω, the spatial filter $A(\omega)$ can be viewed as the discrete-time Fourier transform (DTFT) of the complex sequence $a(n)$, where ω is the corresponding radian frequency. Refer to Eq. (11.15) and Chapters 1 and 2. Thus, we can use the theory of finite impulse response (FIR) linear-phase digital filters to select $|a(n)|$ or, equivalently, to shape $|A(\omega)|$. Also, the shading factors $|a(n)|$ produce the same mathematical effects as the "window functions" used in spectral analysis. Hence, the theory of window functions, as used in spectral analysis, can be useful in the design of linear arrays.

In summary:

1. The physical size of a radar or sonar array is controlled by F_0(orλ) and the 3-dB beamwidth.
2. The sensor spacing of a linear array is generally set at $l = \lambda/2$ to comply with the spatial version of Nyquist's sampling theorem. This choice of l also eliminates grating lobes.
3. The complex sequence $a(n)$ completely defines the linear array. The magnitude of $a(n)$ must be symmetric over $0 \leq n \leq L-1$; we can judiciously select the shading factors $|a(n)|$ to reduce the sidelobe levels at a modest increase in beamwidth. The phase angle of $a(n)$ must be linear in n; the intersensor time delays appear in $\angle a(n)$ and are used to steer the array.
4. The spatial filter (11.12) can be viewed as the DTFT of $a(n)$. Thus, the mathematics of linear-phase FIR digital filters and window functions, as used in spectral analysis, can be useful in the design of linear arrays.
5. In practice, radar and sonar arrays can be different geometrical shapes (e.g., planar, spherical, cylindrical, and linear). Thus, the intersensor time delays will depend on the shape of the array. For a more detailed treatment of array theory see [9].

The linear array (and other types of arrays) can focus its transmitted energy in a specific direction ϕ_0. The *transmit beamformer* is responsible for steering the array to ϕ_0; it does so by exciting the sensors according to Eq. (11.11). The waves radiated outward along ϕ_0 eventually strike a target and return to the array. The *receive beamformer* is responsible for delaying (or phase shifting) the backscattered waveforms received at each sensor by an amount that would maximize the received energy along ϕ_0. It turns out that the receive beamformer uses the same intersensor time delays (or phase shifts) as the transmit beamformer, namely, Eq. (11.11). Further, since all the sensors are assumed to be reciprocal, the receive beam patterns are exactly the same as the transmit beam patterns. The following example summarizes the basic function of the receiver beamformer.

Example 4. For simplicity let us assume that a stationary point target is located at some distant range $r = R$ and at some angle $\phi = \phi_0$ defined by the point $P = (r, \phi, z = 0)$. See Fig. 11.13. If the transmit beamformer of a linear array forms a beam in the direction $\phi = \phi_0$, then the waveform at P (just before it strikes the target) can be expressed as

$$s_P(t) \simeq \frac{G(\phi_0)}{R} s_T\left(t - \frac{R}{c}\right) \tag{11.19}$$

where

$$G(\phi_0) = \sum_{n=0}^{L-1} |a(n)| \tag{11.20}$$

represents the spatial gain of the linear array. The simplicity of Eq. (11.19) is due to the "constructive interference" pattern set up by the transmit beamformer. Now after $s_P(t)$ in Eq. (11.19) strikes the stationary point target, the backscattered waveform returns to the array. The approximate waveform, as seen by each sensor in the array, is [10] as follows:

Sensor Location	Backscattered Waveform (relative to $t = 0$)
$x = 0$	$\frac{AG(\phi_0)}{R^2} s_T\left(t - \frac{2R}{c}\right)$
$x = l$	$\frac{AG(\phi_0)}{R^2} s_T\left(t - \frac{2R}{c} + \Delta_0\right)$
$x = 2l$	$\frac{AG(\phi_0)}{R^2} s_T\left(t - \frac{2R}{c} + 2\Delta_0\right)$
$\vdots$	$\vdots$
$x = (L-1)l$	$\frac{AG(\phi_0)}{R^2} s_T\left(t - \frac{2R}{c} + (L-1)\Delta_0\right)$

(11.21)

Here, A is a backscattering amplitude factor and $\Delta_0 = (l/c)\cos\phi_0$.

Due to the assumed reciprocity of the array, the *receive beamformer* acts the same way as the transmit beamformer. That is, the backscattered waveform received at $x = 0$ gets zero delay (or phase shift). The backscattered waveform received at $x = l$ gets delayed (or phase shifted) by $\delta_1 = \Delta_0$. The waveform received at $x = 2l$ gets a delay of $\delta_2 = 2\Delta_0$. Finally, the waveform received at $x = (L-1)l$ gets delayed by $\delta_{L-1} = (L-1)\Delta_0$. Thus, after $s_P(t)$ strikes the stationary point target and returns to the array, the output of the receive beamformer can be expressed as

$$\begin{aligned} s_R(t) &\simeq \frac{AG(\phi_0)}{R^2}|a(0)|s_T\left(t - \frac{2R}{c}\right) + \frac{AG(\phi_0)}{R^2}|a(1)|s_T\left(t - \frac{2R}{c} + \Delta_0 - \delta_1\right) \\ &\quad + \cdots + \frac{AG(\phi_0)}{R^2}|a(L-1)|s_T\left(t - \frac{2R}{c} + (L-1)\Delta_0 - \delta_{L-1}\right) \\ &\simeq \frac{AG(\phi_0)}{R^2}s_T\left(t - \frac{2R}{c}\right)[|a(0)| + |a(1)| + \cdots + |a(L-1)|] \end{aligned} \tag{11.22}$$

or

$$s_R(t) \simeq \frac{AG^2(\phi_0)}{R^2}s_T(t - \tau) \tag{11.23}$$

Here

$$\tau \equiv \frac{2R}{c} \tag{11.24}$$

is the time delay we must estimate in order to obtain an estimate of the target's range R. Once again, the simplicity of Eq. (11.23) is due to the "constructive interference" pattern set up by the receive beamformer.

In the preceding discussion we made a very important assumption. That is, we assumed that the transmit beamformer pointed the MRA of the beam directly at the target. Thus the dominant backscattering is contained in the beam's mainlobe, the electrical signal out of the receive beamformer has the simple form [Eq. (11.23)], $G(\phi_0)$ is a maximum, and the energy in $s_R(t)$ is a maximum. If the transmit beamformer had pointed the MRA in any other direction, the dominant backscattering would have probably been contained in the beam's sidelobes, $s_R(t)$ would experience "destructive interference," and the energy in $s_R(t)$ would be less. Thus, the spatial filter $A(F, \phi)$ is "matched" when the beam's MRA is pointed directly at the target. Under this matched condition the receive beamformer signal $s_R(t)$ has maximum energy.

Using the notion of a *matched spatial filter*, we can derive a simple algorithm for obtaining a coarse estimate of the target's direction. We make the following assumptions:

1. We consider a stationary point target located in a homogeneous, isotropic, lossless medium.

2. We assume that the target is known to exist in the sector $0° \leq \phi \leq 180°$.
3. We assume no background or receiver noise.

For a linear array we propose the following algorithm:

Algorithm for Obtaining a Coarse Estimate of the Target's Direction

Step 1. Given a 3-dB beamwidth of $\Delta\phi°$, form $180°/\Delta\phi°$ beams. In each beam transmit, receive, and record the detected energy.

Step 2. The beam that has the largest detected energy is the beam that probably contains the target. The MRA angle associated with this beam gives a coarse estimate of the target's direction. Here, a coarse estimate means that the target is somewhere in the $\Delta\phi°$ beam.

This algorithm is a "common sense" type. That is, it steers the beam in steps of the 3-dB beamwidth until the MRA is matched to the target's direction. When this matching occurs, $s_R(t)$ has maximum energy, and we can assume that the matched beam is the beam that probably contains the target. In practice, many phased-array search radars use this simple concept to obtain a coarse estimate of the target's direction. Once a coarse estimate is obtained, these phased-array radars refine this estimate by using monopulse techniques [11]. Figure 11.15 gives a pictorial description of the algorithm. The "common sense" can be rigorously justified by the theory of maximum likelihood estimation [12].

We are now ready to state the time delay estimation problem for the case of any array containing numerous omnidirectional active sensors. We make the following assumptions:

1. The target is a stationary point target located in a homogeneous, isotropic, lossless medium.

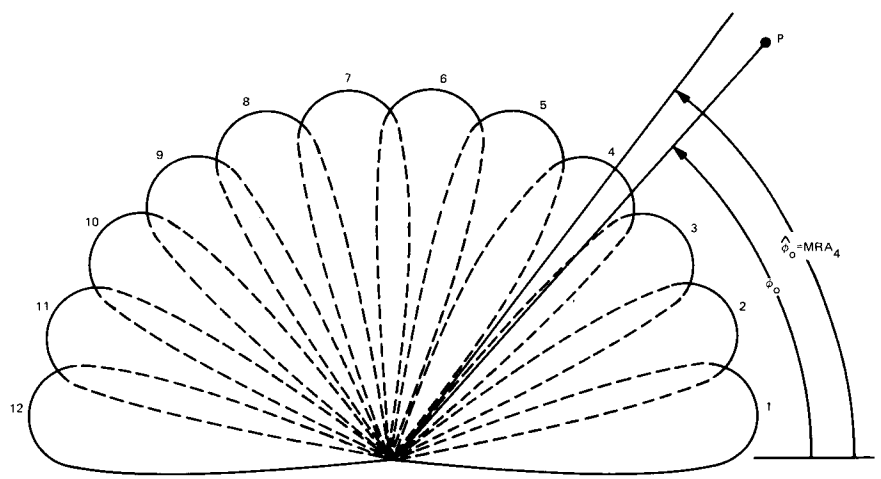


Fig. 11.5. Coarse estimate of target direction by means of matched spatial filtering.

2. The target's range R is well within the active system's detection range (e.g., the SNR is relatively large—greater than 15 dB).

3. The active system has an array of omnidirectional reciprocal sensors, so it can perform both transmit and receive beamforming; that is, it can perform spatial filtering. It transmits $s_T(t)$ in Eq. (11.1).

Let the above active system form numerous densely packed beams that fill a hemisphere. For beam j the receive beamformer output has the form

$$\underline{z}_j(t) = s_{Rj}(t) + \underline{v}_j(t) \tag{11.25}$$

where $t_p \leq t \leq t_p + t_0$ and t_0 is the length of the observation interval (Fig. 11.8). Here, $\underline{v}_j(t)$ represents the additive noise process associated with beam j, and $s_{Rj}(t)$ represents the receive beamformer output of beam j when no noise is present. For some beam $j = M$ the MRA will match (approximately) the target's direction. Under this condition

$$\underline{z}_M(t) \simeq \frac{AG^2}{R^2} s_T(t - \tau) + \underline{v}_M(t) \tag{11.26}$$

where A is a backscattering factor, G is the spatial gain of the array, R is the target range, and τ is the desired time delay (e.g., two-way travel time) given by Eq. (11.24). The problem is to estimate the target direction $\mathbf{R}/\mathrm{R}$ and target range R by forming many beams that fill the hemisphere containing the target, where $\mathbf{R}$ denotes the target range vector. As we have seen, the beam with the maximum received energy (say $j = M$) gives an estimate of $\mathbf{R}/\mathrm{R}$. In the next section we will show how Eq. (11.26) can be processed to obtain an estimate of τ. The target range R can then be obtained from Eq. (11.24).

We close this section with Figs. 11.16 and 11.17, which summarize the process of beamforming and the time delay estimation problem for an array of omnidirectional active sensors, respectively.

A Time Delay Estimation Algorithm for Active Sensors C

Stationary Point Target Backscattering Model 1

Before we discuss the details of a time delay estimation algorithm, let us examine the stationary point target backscattering model used earlier. Recall that when the beam's MRA is pointed directly at a stationary point target, the output of the receive beamformer, $s_R(t)$, has the form of Eq. (11.23). In words, $s_R(t)$ is simply a scaled, time-delayed replica of $s_T(t)$. However, in practice, this assumption is not always true, because the scattering of electromagnetic or acoustic waves off targets of different shapes is generally a more complicated process. For example, Fig. 11.18 shows what happens when a continuous wave (CW) acoustic pulse is incident on a stationary aluminum sphere embedded in water. Depending on the size parameter $ka = 2\pi(a/\lambda)$, where λ is the wavelength

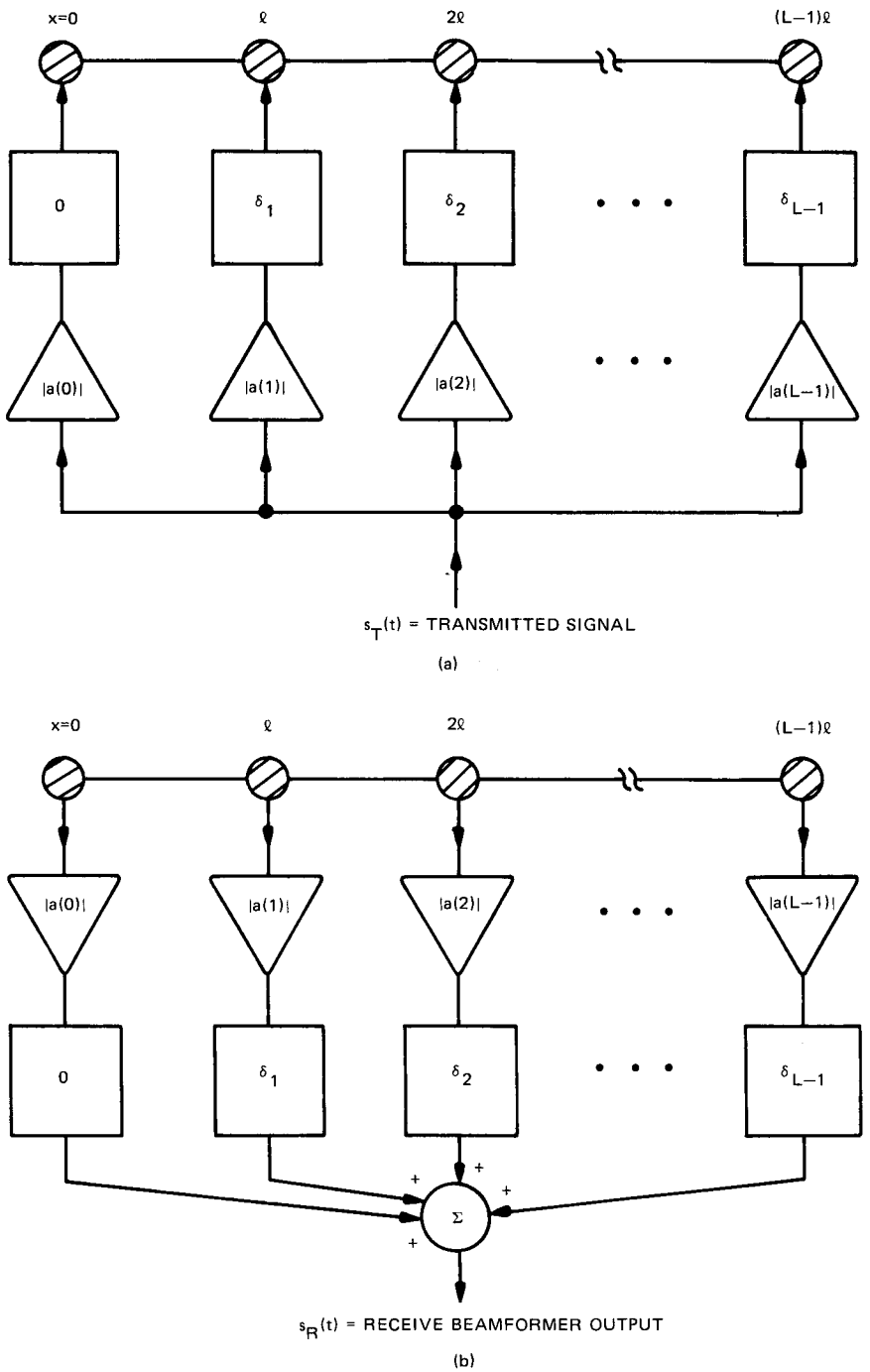


Fig. 11.16. Block diagram of the (a) transmit beamformer and (b) receive beamformer for a linear arrary.

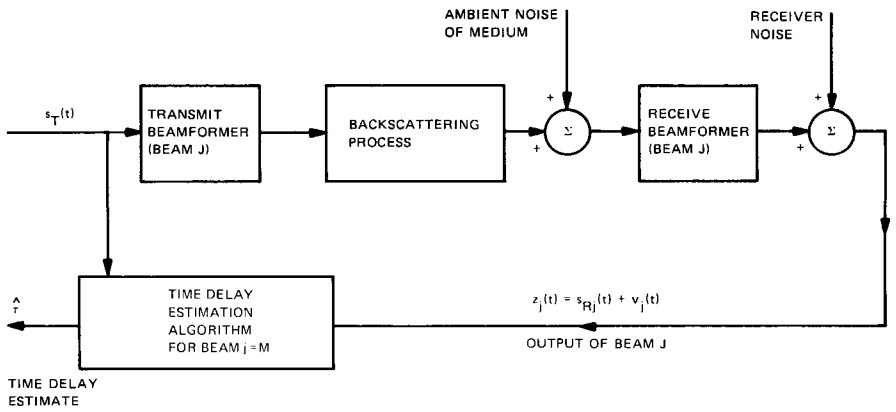


Fig. 11.17. Block diagram of the time delay estimation problem for an array of omnidirectional active sensors.

of the incident wave and a is the radius of the sphere, different backscattering occurs [13]. Figure 11.18(b), (c) shows that $s_R(t)$ is not a scaled, time-delayed version of $s_T(t)$.

To properly describe the backscattering that results from probing these differently shaped targets, we would have to solve the partial differential equations that describe the backscattering. This requires detailed knowledge of the target's size and orientation, which are generally not known. However, suppose that the backscattering process, described in the block diagrams of Figs. 11.9 and 11.17, could be described by a linear time-invariant filter, say $h(t)$. Then the noisy beamformer output $\underline{z}(t)$ could be written as

$$\underline{z}(t) = s_R(t) + \underline{v}(t) = \int h(\sigma) s_T(t - \sigma)\, d\sigma + \underline{v}(t) \tag{11.27}$$

for $t_p \leq t \leq t_p + t_0$. Thus, given $\underline{z}(t)$ for t_0 s, we could estimate $h(t)$. Once $h(t)$ is known, the backscattering process is known, and we could estimate the time delay τ. Equation (11.27) is a *convolution* process. The process of estimating $h(t)$ from Eq. (11.27) is called *deconvolution* [14].

Since deconvolution usually involves the estimation of a complete time function, algorithms that perform deconvolution can be computationally involved, depending on the nature of $h(t)$. The deconvolution problem is considered further in Chapter 10.

One attractive feature of the stationary point target backscattering model is that the time delay τ appears in $s_R(t)$ as a simple parameter. Since most of the research on time delay estimation for active systems has focused on the parameter estimation approach, we will consider that approach here.

Thus, although the stationary point target model is a gross simplification of the real received echo, it is commonly used to keep the mathematical model of the receive beamformer output as simple as possible. From a time delay estimation point of view, this makes the time delay problem mathematically tractable.

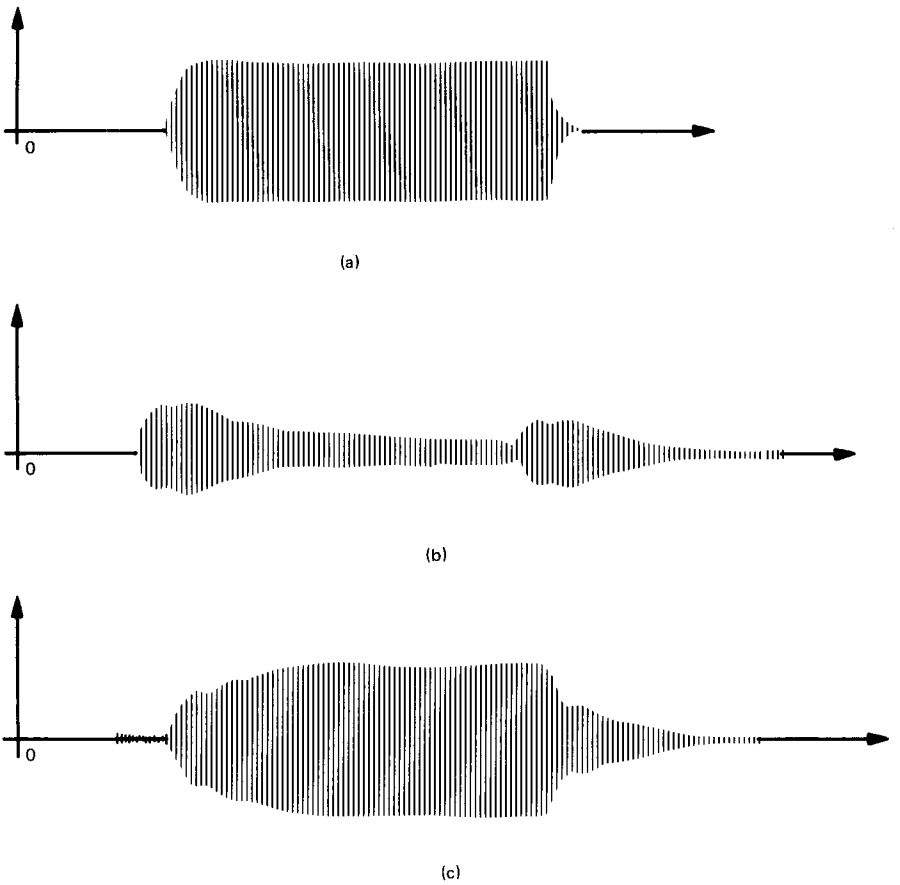


Fig. 11.18. Aluminum sphere, incident-backscattered pulse. (a) A 500 microsecond-long transmitted acoustic CW pulse, incident on a stationary aluminum sphere embedded in water. (b) The pulse reflected by an aluminum sphere when the pulse in (a) is incident for $ka = 20.78$, where $ka = 2\pi(a/\lambda)$ is the wavelength of the incident wave, and a is the radius of the sphere. (c) The backscattered pulse for $ka = 21.21$.

2 *A Least Squares Approach to Time Delay Estimation*

We assume that an active system has gone through the beamforming process and has obtained a coarse estimate of the target's direction. That is, the target is definitely located in some beam. The corresponding noisy beamformer output has the form

$$\underline{z}(t) = \frac{AG^2}{R^2} s_T(t - \tau) + \underline{v}(t) \tag{11.28}$$

for $t_p \leq t \leq t_p + t_0$. For example, see Fig. 11.8. We now wish to estimate the time delay τ.

We assume that τ is the true but unknown time delay. We further assume that τ can be treated as a time-independent parameter. To estimate τ, we form the quantity

$$J(\tilde{\tau}) = \int_{t_p}^{t_p+t_0} \left[\underline{z}(t) - \frac{AG^2}{R^2} s_T(t-\tilde{\tau}) \right]^2 dt \tag{11.29}$$

which is a function of the parameter $\tilde{\tau}$. We then vary $\tilde{\tau}$ until $J(\tilde{\tau})$ is a minimum. The value of $\tilde{\tau}$ that produces the smallest J is labeled $\underline{\hat{\tau}}$ and is called the *least squares estimate* of τ. Let us now derive a least squares algorithm for estimating τ.

Differentiating Eq. (11.29) with respect to $\tilde{\tau}$ gives

$$\frac{\partial J}{\partial \tilde{\tau}} = \int_{t_p}^{t_p+t_0} 2\left[\underline{z}(t) - \frac{AG^2}{R^2} s_T(t-\tilde{\tau}) \right]\left[-\frac{AG^2}{R^2} \frac{\partial s_T(t-\tilde{\tau})}{\partial \tilde{\tau}} \right] dt \tag{11.30}$$

Now we can obtain the minimum of Eq. (11.29) by setting Eq. (11.30) equal to zero. Thus

$$\begin{aligned} \int_{t_p}^{t_p+t_0} \underline{z}(t) \frac{\partial s_T(t-\tilde{\tau})}{\partial \tilde{\tau}} dt &= \frac{AG^2}{R^2} \int_{t_p}^{t_p+t_0} s_T(t-\tilde{\tau}) \frac{\partial s_T(t-\tilde{\tau})}{\partial \tilde{\tau}} dt \\ &= \frac{AG^2}{R^2} \int_{t_p}^{t_p+t_0} \frac{\partial}{\partial \tilde{\tau}} [\tfrac{1}{2} s_T^2(t-\tilde{\tau})]\, dt \end{aligned} \tag{11.31}$$

In practice, $t_0 \gg t_p$. That is, the observation interval t_0 is much longer than the transmitted pulse width t_p. So if we assume that $t_p \leq \tilde{\tau} \leq t_0$, then the right side of Eq. (11.31) is approximately zero:

$$\begin{aligned} \int_{t_p}^{t_p+t_0} \frac{\partial}{\partial \tilde{\tau}} [\tfrac{1}{2} s_T^2(t-\tilde{\tau})]\, dt &\simeq \frac{\partial}{\partial \tilde{\tau}} \int_{t_p}^{t_p+t_0} \tfrac{1}{2} s_T^2(t-\tilde{\tau})\, dt \\ &\simeq \frac{1}{2} \frac{\partial}{\partial \tilde{\tau}} (E_T) = 0 \end{aligned} \tag{11.32}$$

where

$$E_T = \int_{t_p}^{t_p+t_0} s_T^2(t-\tilde{\tau})\, dt = \int_0^{t_p} s_T^2(t)\, dt = \text{constant} \tag{11.33}$$

is the energy contained in the transmitted signal $s_T(t)$. Thus, it follows that Eq. (11.31) can be rewritten as

$$\int_{t_p}^{t_0+t_p} \underline{z}(t) \frac{\partial s_T(t-\tilde{\tau})}{\partial \tau} \simeq 0 \tag{11.34}$$

or

$$\frac{\partial}{\partial \tilde{\tau}} [\underline{y}(\tilde{\tau})] \simeq 0 \tag{11.35}$$

where

$$\underline{y}(\tilde{\tau}) \equiv \int_{t_p}^{t_p+t_0} \underline{z}(t) s_T(t-\tilde{\tau})\, dt \tag{11.36}$$

In summary, the value of $\tilde{\tau}$, say $\tilde{\tau} = \hat{\underline{\tau}}$, that produces a unique global maximum in Eq. (11.36) also satisfies Eq. (11.35) and therefore minimizes Eq. (11.29). The value $\hat{\underline{\tau}}$ is called the *least squares estimate* of τ.

Equation (11.36) is a key equation in the study of time delay estimation algorithms for active sensors. Mathematically $\underline{y}(\tilde{\tau})$ in Eq. (11.36) can be viewed as a cross-correlation function with independent variable $\tilde{\tau}$. That is, under the ergodic hypothesis an estimate of $\underline{y}(\tilde{\tau})$ can be interpreted as the correlation between the noise-free transmitted signal $s_{\hat{T}}(t)$ and the noisy beamformer output $\underline{z}(t)$. Thus, we can vary the *lag* $\tilde{\tau}$ for all $\tilde{\tau}$ in the interval $t_p \leq \tilde{\tau} \leq t_0$. At some lag $\tilde{\tau} = \hat{\underline{\tau}}$ the cross-correlation function $\underline{y}(\tilde{\tau})$ achieves a maximum. This maximum occurs at the least squares estimate of τ. Figure 11.19 is a block diagram of a time delay estimation algorithm that implements Eq. (11.36) as a bank of cross-correlators.

Let us now consider another way to interpret Eq. (11.36). Specifically, let us assume that $\underline{z}(t)$ is passed through a linear time-invariant filter with impulse response $s_T(-t)$. If we sample the output of this filter, say $\underline{y}(t)$, at $t = \tilde{\tau}$, we obtain the same mathematical result as in Eq. (11.36). The maximum output of this filter occurs at $t = \hat{\underline{\tau}}$.

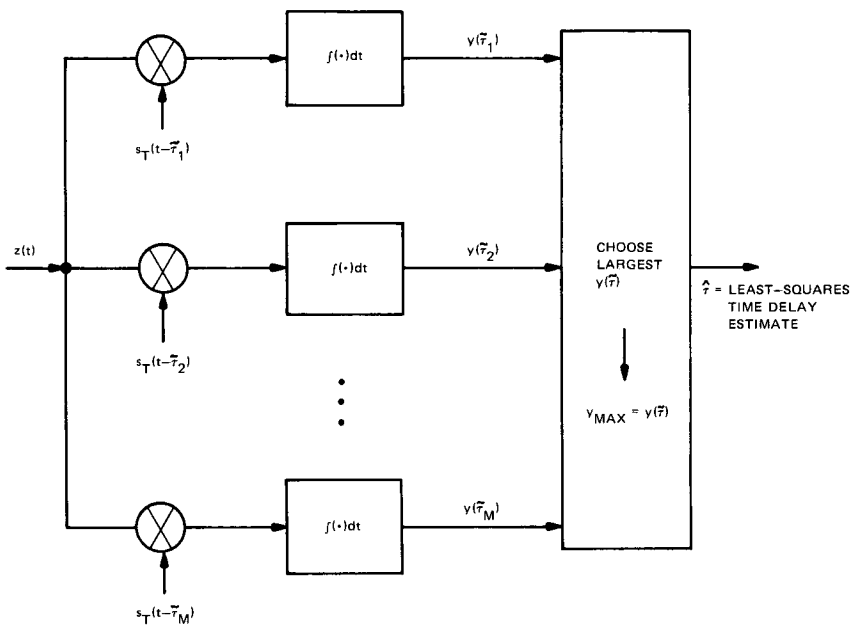


Fig. 11.19. Time delay estimation using a bank of correlators.

The preceding interpretation is possible because of the mathematical similarity of correlation and convolution. The impulse response $s_T(-t)$ not only reproduces (11.36), but it is also the only linear time-invariant filter that produces the maximum SNR at $t = \hat{\underline{\tau}}$. For $\underline{v}(t)$ a white noise process with autocorrelation function

$$E[\underline{v}(t)\underline{v}(t+\beta)] = \tfrac{1}{2}N_0\delta(\beta) \tag{11.37}$$

the maximum SNR at $t = \hat{\underline{\tau}}$ is $2E_T/N_0$, where E_T is given by (11.33). Here, $\delta(\beta)$ is

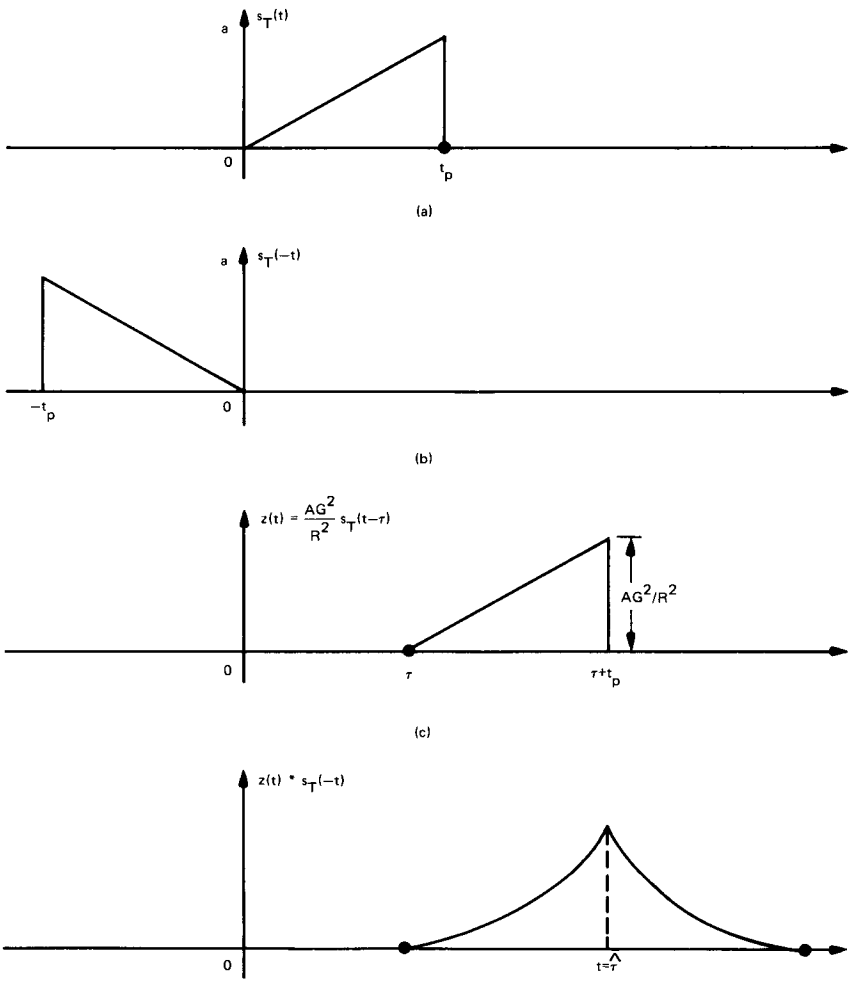


Fig. 11.20. Matched filter produces maximum SNR at $t = \hat{\underline{\tau}}$. (a) Transmitted signal $s_T(t)$; (b) matched filter; (c) receive beamformer output when no noise is present (e.g., $\underline{v}(t) = 0$); (d) output of the matched filter; (c) convolved with (b).

the familiar Dirac delta function. The filter $s_T(-t)$, which corresponds to $s_T(t)$, is called the *matched filter* for $s_T(t)$ [15] (Fig. 11.20). The matched filter not only produces the maximum SNR at $t = \hat{\tau}$, but it can also be implemented as a correlator. Matched filters or correlators also play an important role in the signal detection problem [16].

When the noise process $v(t)$ is Gaussian, Eq. (11.36) can be derived from the theory of maximum-likelihood estimation. In any event, Eq. (11.36) is a key equation. In Section V we will consider a practical implementation of Eq. (11.36) or Fig. 11.19.

In closing this section we refer the reader to Figs. 11.15–11.17 and 11.19, which summarize the important concepts involved in beamforming and time delay estimation for active systems.

III Time Delay Estimation for Passive Sensors

A The Time Delay Estimation Problem for a Passive Array with Two Omnidirectional Reciprocal Sensors

In general, a passive sonar or radar contains an array of omnidirectional sensors and a receiver. The purpose of the passive array is to sense the acoustic or electromagnetic waves that are radiated by a target (refer to Fig. 11.1). The passive receiver converts these waves into electrical signals and decides if a target was really present. Further signal processing can ultimately localize the target (i.e., estimate the target's range and direction).

Let us begin our discussion of passive arrays by considering the simplest of all passive arrays, the one with only two sensors. See Fig. 11.10. We assume that a target, located at point P, radiates the signal $s(t)$. Now if the target radiates $s(t)$ at $t = 0$, the signal received at the $x = l$ sensor has the form $s(t - r_l/c)/r_l$, and the signal received at the $x = 0$ sensor has the form $s(t - r/c)/r$. Here, we have assumed that the medium is noise free, homogeneous, isotropic, and lossless; c is the corresponding wave propagation speed. If the sensor outputs are now combined according to the rule

$$s_R(t) = \frac{|a(0)|}{r} s\left(t - \frac{r}{c}\right) + \frac{|a(1)|}{r_l} s\left(t - \frac{r_l}{c} - D\right) \tag{11.38}$$

then we have performed *receive beamforming* on the sensor outputs [refer to Fig. 11.16(b)]. Recall that $|a(0)|$ and $|a(1)|$ are the shading factors and D is a time delay parameter. When D is equal to the *intersensor time delay*

$$\Delta = \frac{r}{c} - \frac{r_l}{c} \tag{11.39}$$

Eq. (11.38) reduces to

$$s_R(t)\Big|_{D=\Delta} = s\left(t - \frac{r}{c}\right)\left[\frac{|a(0)|}{r} + \frac{|a(1)|}{r_l}\right] \tag{11.40}$$

That is, for $D = \Delta$ we experience a "constructive interference" effect, and Eq. (11.38) achieves its maximum, Eq. (11.40). Thus, we can vary D until $s_R(t)$ achieves a maximum. The values of D that maximizes the receive beamformer output $s_R(t)$ is the intersensor time delay Δ. Given an estimate of Δ, can we localize a radiating target with a two-sensor, linear, passive array? Let us proceed to answer this question.

From Fig. 11.10 we see that

$$r_l^2 = r^2 + l^2 - 2rl\cos\phi \tag{11.41}$$

where r is the target range and ϕ is the target bearing or direction. Examination of Eqs. (11.39) and (11.41) reveals that we have two equations but three unknowns, r, r_l, and ϕ. Hence, given Δ, we cannot obtain r and ϕ for all ranges and bearings. However, if the radiating target is "very far" from the passive array, such that $\phi \simeq \alpha$, then r and r_l are approximately parallel. It follows that

$$r \simeq r_l + l\cos\phi \tag{11.42}$$

and

$$\Delta \simeq \frac{l}{c}\cos\phi \tag{11.43}$$

Under this special condition (i.e., $r \to \infty$) we can use Δ to get an approximate estimate of ϕ by means of Eq. (11.43). Hence, the two-sensor, linear, passive array cannot localize a radiating target. However, when the target is very far from the array, we can estimate the target's bearing.

The two-sensor, linear, passive-array, time delay estimation problem is equivalent to the bearing estimation problem for distant targets. The basic idea is to estimate the intersensor time delay Δ and use Eq. (11.43) to estimate the bearing ϕ.

B The Time Delay Estimation Problem for a Passive Array with Three Omnidirectional Reciprocal Sensors

In the previous section we saw that the two-sensor problem produced one intersensor time delay (11.39), two equations [(11.39) and (11.41)], and three unknowns. The end result was that a radiating target could not be localized by a two-sensor, linear, passive array. Let us now see what happens to the passive localization problem when we add one more sensor to the two-sensor array in Fig. 11.10. For example, Fig. 11.21 shows a three-sensor, linear, passive array and a target located at a range r and bearing ϕ. If this target radiates the waveform $s(t)$

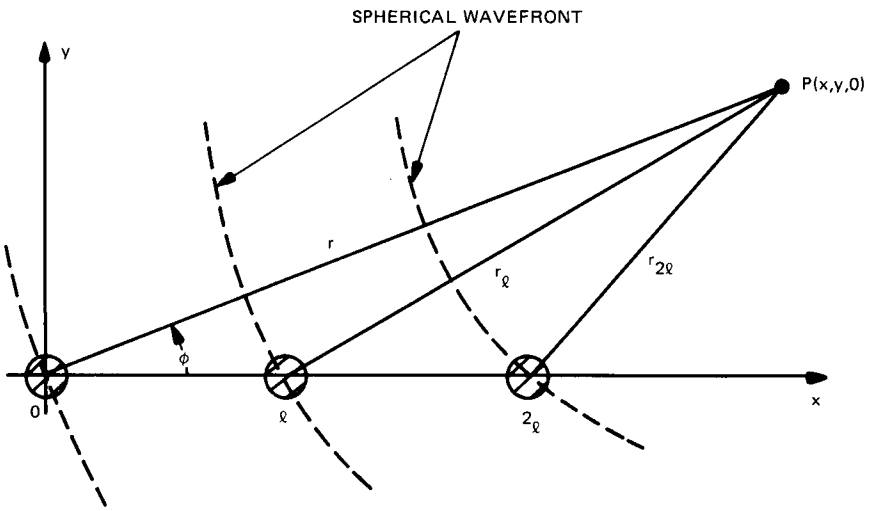


Fig. 11.21. Three-sensor linear passive array.

at $t = 0$, the waveform received at the $x = 2l$ sensor has the form $s(t - r_{2l}/c)/r_{2l}$, the waveform received at the $x = l$ sensor has the form $s(t - r_l/c)/r_l$, and the waveform received at the $x = 0$ sensor has the form $s(t - r/c)/r$. Now

$$\Delta_{01} = \frac{r}{c} - \frac{r_l}{c} \tag{11.44}$$

is the intersensor time delay between the $x = 0$ and $x = l$ sensors and

$$\Delta_{02} = \frac{r}{c} - \frac{r_{2l}}{c} \tag{11.45}$$

is the intersensor time delay between the $x = 0$ and $x = 2l$ sensors. Since the intersensor time delay between the $x = l$ and $x = 2l$ sensors can be expressed as

$$\Delta_{12} = \frac{r_l}{c} - \frac{r_{2l}}{c} = \Delta_{02} - \Delta_{01} \tag{11.46}$$

we see that the three-sensor, linear, passive array has only two independent intersensor delays. Moreover, due to the spherical nature of the wavefronts in Fig. 11.21, Δ_{02} is not, in general, an integer multiple of Δ_{01}; only when $r \to \infty$ is this approximately true. Therefore, if we attempt to perform beamforming according to the assumption that $\Delta_{02} \simeq 2\Delta_{01}$ [see Fig. 11.16(b)], we are not likely to obtain the correct result. Hence, we should consider a beamformer that combines the three sensor outputs according to the rule

$$s_R(t) = \frac{|a(0)|}{r} s\left(t - \frac{r}{c}\right) + \frac{|a(1)|}{r_l} s\left(t - \frac{r_l}{c} - D_{01}\right) + \frac{|a(2)|}{r_{2l}} s\left(t - \frac{r_{2l}}{c} - D_{02}\right) \tag{11.47}$$

where D_{01} and D_{02} are two independent time delay parameters. The beamformer output Eq. (11.47) achieves a global maximum when $D_{01} = \Delta_{01}$ and $D_{02} = \Delta_{02}$. Thus, we now have the following "intuitive" algorithm for estimating Δ_{01} and Δ_{02} under noise-free conditions:

1. Perform the beamforming operation (11.47) on the three sensor outputs.
2. Vary D_{01} and D_{02} in a two-dimensional fashion until $s_R(t)$ in (11.47) achieves a global maximum. The location of this maximum should occur at the intersensor time delays Δ_{01} and Δ_{02}.

Finally, given Δ_{01} and Δ_{02}, can the three-sensor, linear, passive array solve the passive localization problem? Let us proceed to answer this question.

From Fig. 11.21 we see that

$$r_l^2 = l^2 + r^2 - 2rl\cos\phi \tag{11.48}$$

and

$$r_{2l}^2 = (2l)^2 + r^2 - 2r(2l)\cos\phi \tag{11.49}$$

Given that Δ_{01} and Δ_{02} are known (or can be estimated), we now have four independent equations [(11.44), (11.45), (11.48), (11.49)] and four unknowns (r, ϕ, r_l, r_{2l}). Thus, the three-sensor, linear, passive array shows that a solution to the passive localization problem exists. Doing some algebra we obtain the target's range

$$r = \frac{l^2 + (c\Delta_{01})^2 - \frac{1}{2}(c\Delta_{02})^2}{2c\Delta_{01} - c\Delta_{02}} \tag{11.50}$$

and the target's bearing

$$\phi = \cos^{-1}\left[\frac{l^2 - (c\Delta_{01})^2 + 2rc\Delta_{01}}{2rl}\right] \tag{11.51}$$

However, although a solution exists, it is not unique. For example, a target located at the point $(r, -\phi, z = 0)$ (e.g., below the x-axis in Fig. 11.21) would give the same time delays (and therefore the same localization solution) as a target located at the point $(r, \phi, z = 0)$. Hence, we have an ambiguity in target bearing. Recall that this ambiguity also existed for the array with active sensors. It turns out that this bearing ambiguity is directly related to the fact that each sensor has an omnidirectional spatial characteristic instead of a directional one. Thus, to resolve the bearing ambiguity, (i) we must have a priori knowledge that the target exists in the sector $0° \leq \phi \leq 180°$ (or $180° \leq \phi \leq 360°$), or (ii) we must use directional or baffled sensors.

In summary, the three-sensor, linear, passive array provides a solution to the passive localization problem. In the next section we consider various signal processing strategies for estimating the two independent time delays Δ_{01} and Δ_{02}.

C A Time Delay Estimation Algorithm for Passive Sensors

1 *An Alogrithm for the Two-Sensor Linear Array*

Although we cannot solve the passive localization problem with a two-sensor, linear, passive array, we can obtain the bearing of a distant target by estimating the intersensor time delay Δ_{01}. Let us elaborate.

Assume that the point P in Fig. 11.10 represents a distant, stationary target that radiates a real, zero-mean, wide-sense stationary (WSS), random process $\underline{s}(t)$. Further, we assume that the passive receiving array is stationary and that the position of the sensors is exactly known. Hence, the signal received at the $x = 0$ sensor has the form

$$\underline{z}_0(t) = \frac{1}{r}\underline{s}\left(t - \frac{r}{c}\right) + \underline{v}_0(t) \tag{11.52}$$

and the signal received at the $x = l$ sensor has the form

$$\underline{z}_1(t) = \frac{1}{r_l}\underline{s}\left(t - \frac{r_l}{c}\right) + \underline{v}_1(t) \tag{11.53}$$

Here, $\underline{v}_0(t)$ and $\underline{v}_1(t)$ are assumed to be real, uncorrelated, WSS noise processes. If we now perform receive beamforming on these two sensor, the beamformer output has the form

$$\underline{z}(t) = |a(0)|\underline{z}_0(t) + |a(1)|\underline{z}_1(t - D_{01}) \tag{11.54}$$

where $|a(0)|$ and $|a(1)|$ are the shading factors and D_{01} is a time delay parameter. Since $\underline{s}(t)$, $\underline{v}_0(t)$, and $\underline{v}_1(t)$ are assumed to be real, zero-mean, WSS, random processes, it follows that $\underline{z}(t)$ is also a real, zero-mean, WSS random process.

Based on our discussions in Section III.A, we could vary D_{01} until $\underline{z}(t)$ achieves a maximum. However, since $\underline{z}(t)$ is a random process, this procedure does not make sense. Rather than maximizing the beamformer output [Eq. (11.54)], let us consider maximizing the mean-squared value of the beamformer output, namely,

$$J(D_{01}) = E[\underline{z}(t)]^2 = E[|a(0)|\underline{z}_0(t) + |a(1)|\underline{z}_1(t - D_{01})]^2 \tag{11.55}$$

The value of D_{01} that maximizes Eq. (11.55) will be our estimate of the intersensor time delay Δ_{01}. Now Eq. (11.55) can be rewritten as

$$J(D_{01}) = |a(0)|^2 E[\underline{z}_0(t)]^2 + 2|a(0)||a(1)|R_{01}(D_{01}) + |a(1)|^2 E[\underline{z}_1(t - D_{01})]^2 \tag{11.56}$$

where

$$R_{01}(D_{01}) \equiv E[\underline{z}_0(t)\underline{z}_1(t - D_{01})] \tag{11.57}$$

is the statistical cross-correlation function between the real processes $\underline{z}_0(t)$ and $\underline{z}_1(t)$. Here, D_{01} is the correlation or lag parameter. Since $\underline{z}_0(t)$ and $\underline{z}_1(t)$ were

assumed to be zero-mean, WSS random processes, it follows that

$$E[\underline{z}_0(t)]^2 = \text{var}[\underline{z}_0(t)] = \text{constant}$$
$$E[\underline{z}_1(t - D_{01})]^2 = E[\underline{z}_1(t)]^2 = \text{var}[\underline{z}_1(t)] = \text{constant} \quad (11.58)$$

Consequently, from Eq. (11.55) through Eq. (11.58) we conclude that maximizing Eq. (11.55) is equivalent to maximizing the cross-correlation function Eq. (11.57). Hence, locating the peak of the cross-correlation function is equivalent to estimating the intersensor time delay Δ_{01}.

Notice that $R_{01}(D_{01})$, as defined in Eq. (11.57), requires ensemble averaging. However, if the zero-mean, WSS processes $\underline{z}_0(t)$ and $\underline{z}_1(t)$ are ergodic (see Chapter 1) and are observed for $t_p \leq t \leq t_p + t_0$, then Eq. (11.57) can be approximated by

$$\hat{\underline{R}}_{01}(D_{01}) = \int_{t_p}^{t_p + t_0} \underline{z}_0(t)\underline{z}_1(t - D_{01})\, dt \quad (11.59)$$

provided that the observation interval t_0 is very long. Here, $\hat{\underline{R}}_{01}(D_{01})$ denotes an estimate of $R_{01}(D_{01})$. Figure 11.22 summarizes a cross-correlation time delay estimation algorithm for the two-sensor, linear, passive array. Once Δ_{01} is known, the target's bearing follows from Eq. (11.43).

We can also implement the cross-correlation algorithm of Fig. 11.22 in the frequency domain. For example, the cross-spectral density associated with Eq. (11.57) is defined by the Fourier transform

$$G_{01}(F) \equiv \int_{-\infty}^{\infty} R_{01}(D_{01}) e^{-j2\pi F D_{01}}\, dD_{01} \quad (11.60)$$

Substituting Eq. (11.59) into Eq. (11.60) gives the estimate

$$\hat{\underline{G}}_{01}(F) = \underline{Z}_0(F)\underline{Z}_1^*(F) \quad (11.61)$$

provided that t_0 is very long. Here, $\underline{Z}_0(F)$ and $\underline{Z}_1(F)$ are the Fourier transforms of $\underline{z}_0(t)$ and $\underline{z}_1(t)$, respectively, and $\hat{\underline{G}}_{01}(F)$ denotes an estimate of $\underline{G}_{01}(F)$. Figure 11.23 shows the frequency domain implementation of Fig. 11.22. Notice that the sensor outputs are not prefiltered. That is, the received waveforms at $x = 0$ and $x = l$ go directly into a Fourier transform.

Let us now consider prefiltering $\underline{z}_0(t)$ with a linear time-invariant filter $h_0(t)$ and prefiltering $\underline{z}_1(t)$ with a linear time-invariant filter $h_1(t)$. Since the time delays show up in the phase spectra, let us assume that both prefilters have the same phase spectrum. This will prevent any distortion due to prefiltering. Hence, this idea generalizes the cross-correlation algorithms in Figs. 11.22 and 11.23. Figure 11.24 shows a time domain implementation of a generalized cross-correlation (GCC) algorithm, whereas Fig. 11.25 shows a frequency domain implementation of a GCC algorithm. In either case $\hat{\underline{R}}_{01}^h(D_{01})$ is commonly called the *generalized cross-correlation function.*

We have introduced the ideas of prefiltering and GCC (see Figs. 11.24 and 11.25) as natural, intuitive extensions of Figs. 11.22 and 11.23. A more rigorous justification of the GCC method is given in [17]. Further, there has been a great

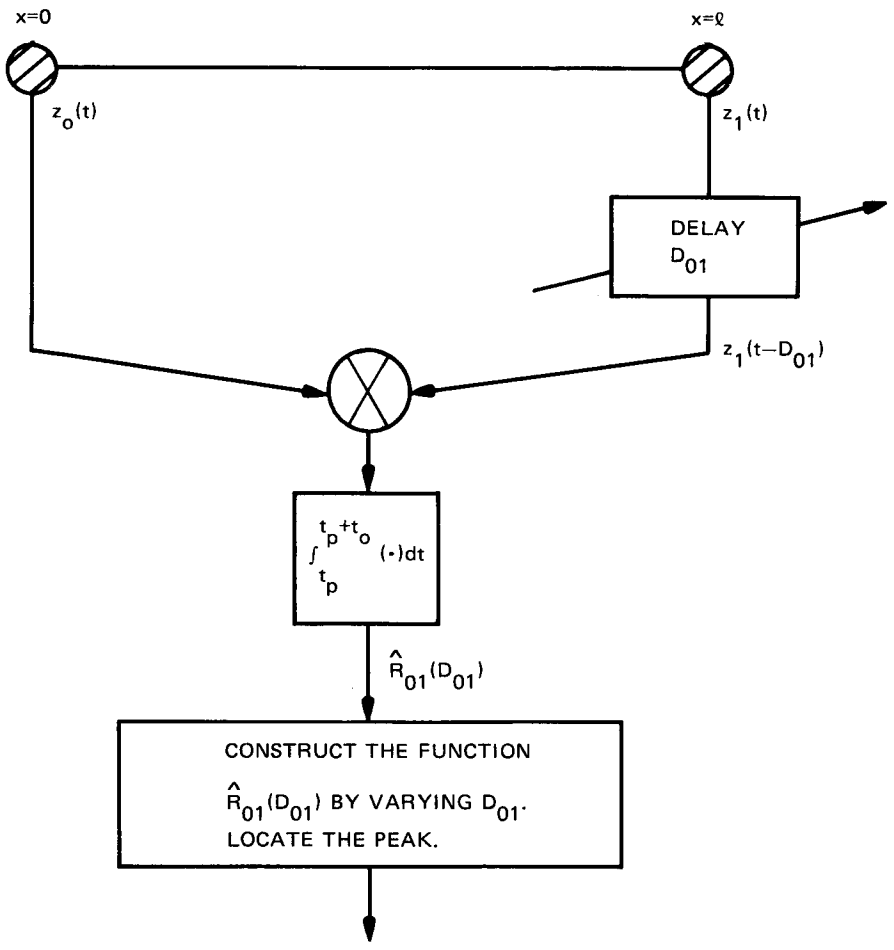


Fig. 11.22. Estimating Δ_{01} by a cross-correlation algorithm.

deal of research conducted on the selection of an "optimal" frequency weighting function $W(F) = |H_0(F)| \cdot |H_1(F)|$ for use in the frequency domain implementation of the GCC method [17]. (See Fig. 11.25.) For example, when the underlying spectral densities of $\underline{z}_0(t)$ and $\underline{z}_1(t)$ are known, when $\underline{s}(t)$, $\underline{v}_0(t)$, and $\underline{v}_1(t)$ are Gaussian and mutually uncorrelated, and when $\hat{G}_{01}(F)$ is obtained by averaging periodogram-type estimates of the cross-spectral density according to [18], then we can choose $W(F)$ to minimize the variance of the time delay estimate $\hat{\underline{\Delta}}_{01}$. The resulting frequency weighting function is

$$W(F) = \frac{1}{G_{ss}(F)}\left[\frac{|\gamma_{01}(F)|^2}{1-|\gamma_{01}(F)|^2}\right] \tag{11.62}$$

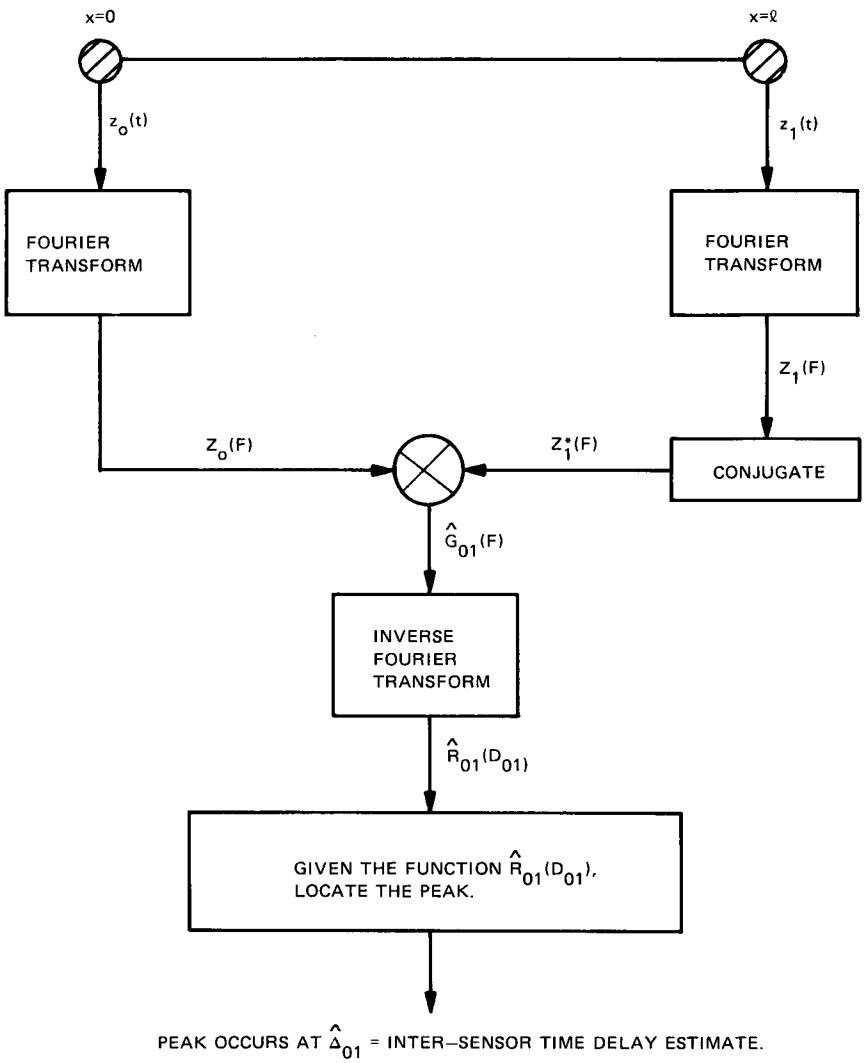


Fig. 11.23. Frequency domain implementation of the cross-correlation algorithm in Fig. 11.22. The sensor outputs $\underline{z}_0(t)$ and $\underline{z}_1(t)$ are not prefiltered—they go directly into a Fourier transform.

where

$$\gamma_{01}(F) = \frac{G_{01}(F)}{\sqrt{G_{00}(F)G_{11}(F)}} \tag{11.63}$$

Here, $G_{ss}(F)$ is the spectral density of $\underline{s}(t)$, $G_{00}(F)$ is the spectral density of $\underline{z}_0(t)$, $G_{11}(F)$ is the spectral density of $\underline{z}_1(t)$, and $G_{01}(F)$ is the cross-spectral density between $\underline{z}_0(t)$ and $\underline{z}_1(t)$. The quantity $\gamma_{01}(F)$ is called the *coherence function* between $\underline{z}_0(t)$ and $\underline{z}_1(t)$. Under the above assumptions and the assumption that t_0

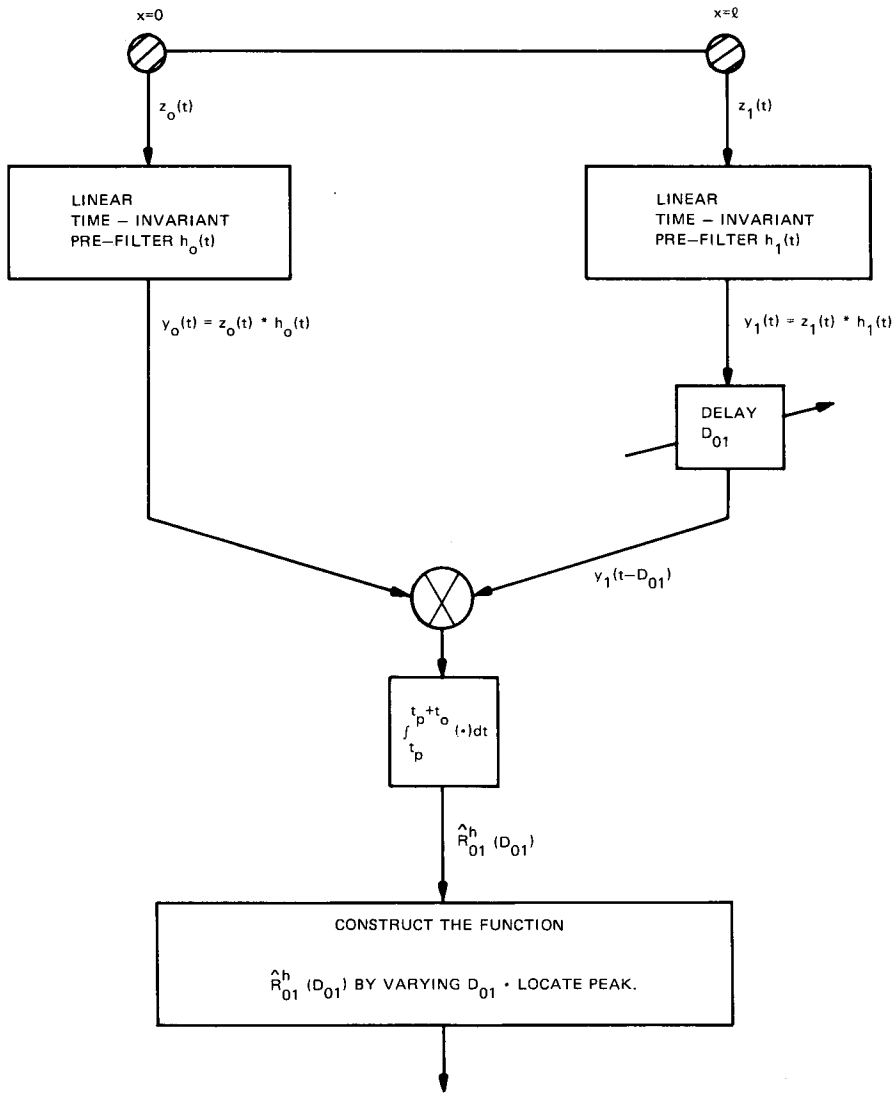


Fig. 11.24. Generalized Cross Correlation (GCC) method.

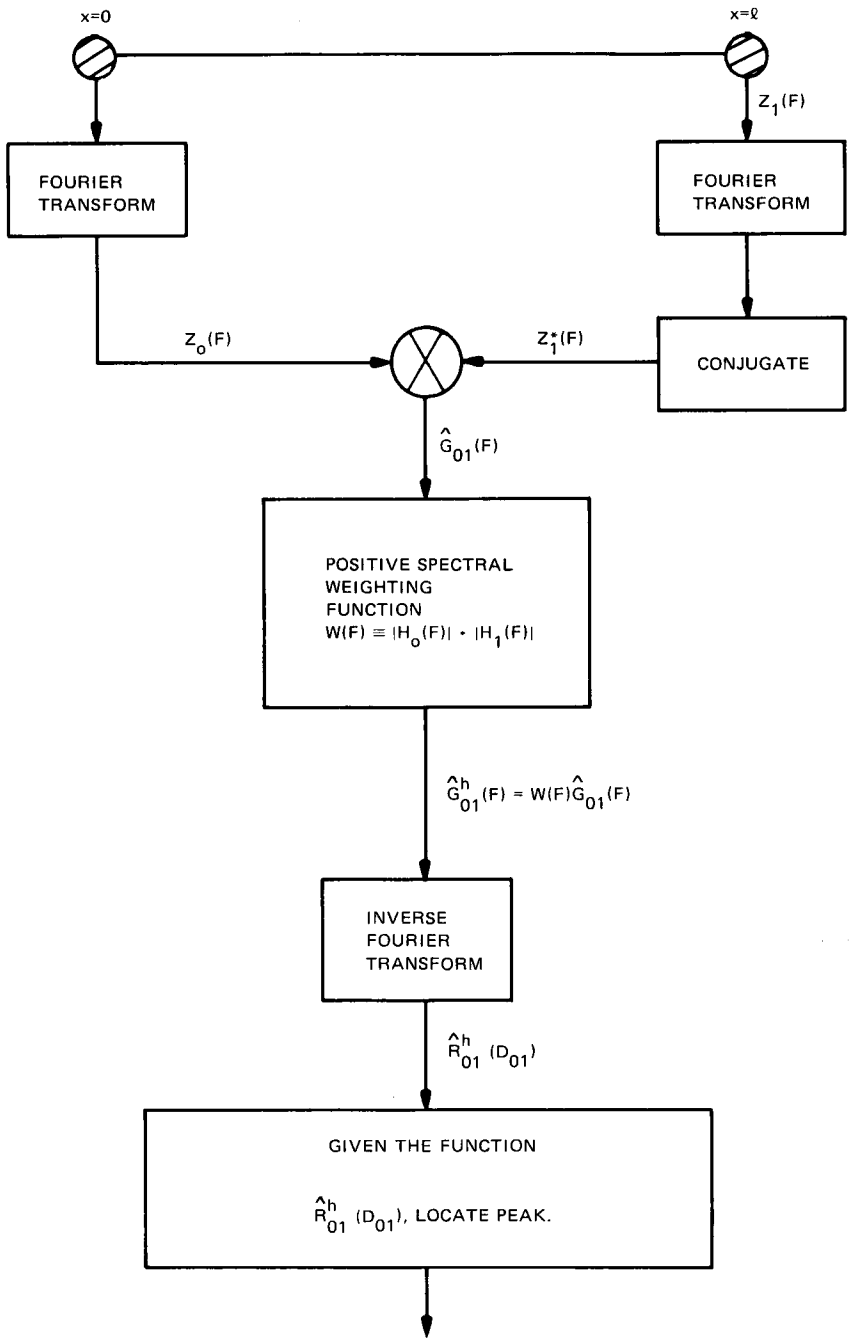


Fig. 11.25. Frequency domain implementation of the GCC method.

is a very long observation time, we can show that the GCC algorithm (Fig. 11.25) that uses Eqs. (11.62) and (11.63) is the *maximum-likelihood estimate* of Δ_{01} [17, 19].

There are other reasons for choosing $W(F)$. For example, the weighting function

$$W(F) = \frac{1}{\sqrt{G_{00}(F)G_{11}(F)}} \tag{11.64}$$

can be used to desensitize the GCC method to the bandwidth properties of $\underline{s}(t)$ [20]. That is, Eq. (11.64) has a prewhitening effect on $\hat{G}_{01}(F)$. Reference [21] provides a good discussion on several commonly used weighting functions and the reasons for their selection.

2 An Algorithm for the Three-Sensor Linear Array

As we saw in Section III.B, the three-sensor, linear, passive array solves the passive localization problem; that is, we can obtain the target's range r and bearing ϕ. However, we do have an ambiguity in target bearing that can be resolved by using (i) a priori knowledge or (ii) directional, rather than omnidirectional, sensors. Nevertheless, we are now faced with the problem of estimating the two independent time delays Δ_{01} and Δ_{02}. The third time delay, Δ_{12}, can be obtained from Eq. (11.46).

Let us assume that the radiating target and receiving passive array are both stationary. Further, we assume that the sensor positions are exactly known. We will also assume that the signal $\underline{s}(t)$ and the three noise processes $\underline{v}_0(t)$, $\underline{v}_1(t)$, and $\underline{v}_2(t)$ are all real, zero-mean, Gaussian, mutually uncorrelated, WSS random processes. Now the waveform received at the $x = 0$ and $x = l$ sensors are given by Eqs. (11.52) and (11.53), respectively, and the waveform received at the $x = 2l$ sensor is given by

$$\underline{z}_2(t) = \frac{1}{r_{2l}} s\left(t - \frac{r_{2l}}{c}\right) + \underline{v}_2(t) \tag{11.65}$$

As we discussed in Section III.B, to obtain the target's range, we cannot assume that $\Delta_{02} \simeq 2\Delta_{01}$, which is equivalent to assuming that $r \to \infty$. In other words, we cannot perform the receive beamforming according to Fig. 11.16(b). Thus, we should perform the beamforming according to the rule

$$\underline{z}(t) = |a(0)|\underline{z}_0(t) + |a(1)|\underline{z}_1(t - D_{01}) + |a(2)|\underline{z}_2(t - D_{02}) \tag{11.66}$$

where $|a(0)|$, $|a(1)|$, and $|a(2)|$ are the shading factors and D_{01} and D_{02} are the two independent time delay parameters. The values $\hat{D}_{01} = \hat{\Delta}_{01}$ and $\hat{D}_{02} = \hat{\Delta}_{02}$ that maximize the mean-squared value of the beamformer output [Eq. (11.66)] will be taken as our estimates of Δ_{01} and Δ_{02}, respectively.

Now the mean-squared value of Eq. (11.66) is

$$J(D_{01}, D_{02}) = E[\underline{z}(t)]^2$$
$$= E[|a(0)|\underline{z}_0(t) + |a(1)|\underline{z}_1(t - D_{01}) + |a(2)|\underline{z}_2(t - D_{02})]^2 \quad (11.67)$$

which represents a surface or two-dimensional function. Expanding Eq. (11.67), we get

$$J(D_{01}, D_{02}) = |a(0)|^2 E[\underline{z}_0(t)]^2 + |a(1)|^2 E[\underline{z}_1(t)]^2$$
$$+ |a(2)|^2 E[\underline{z}_2(t)]^2 + 2|a(0)||a(1)|R_{01}(D_{01})$$
$$+ 2|a(0)||a(2)|R_{02}(D_{02}) + 2|a(1)|a(2)|R_{12}(D_{02} - D_{01}) \quad (11.68)$$

which follows from the WSS assumption on $\underline{z}_0(t)$, $\underline{z}_1(t)$, and $\underline{z}_2(t)$. The statistical cross-correlation function $R_{01}(D_{01})$ is defined in Eq. (11.57). The other two cross-correlation functions are defined by

$$R_{02}(D_{02}) = E[\underline{z}_0(t)\underline{z}_2(t - D_{02})] \quad (11.69)$$

and

$$R_{12}(D_{02} - D_{01}) = E[\underline{z}_1(t - D_{01})\underline{z}_2(t - D_{02})] \quad (11.70)$$

where $\underline{z}_0(t)$, $\underline{z}_1(t)$, and $\underline{z}_2(t)$ are real processes. If these processes are ergodic and observed over a very long observation interval of length t_0, then the ensemble averages in Eqs. (11.57), (11.69), and (11.70) can be approximated by the time averages Eq. (11.59),

$$\hat{\underline{R}}_{02}(D_{02}) = \int_{t_p}^{t_p + t_0} \underline{z}_0(t)\underline{z}_2(t - D_{02})\, dt \quad (11.71)$$

and

$$\hat{\underline{R}}_{12}(D_{02} - D_{01}) = \int_{t_p}^{t_p + t_0} \underline{z}_1(t - D_{01})\underline{z}_2(t - D_{02})\, dt \quad (11.72)$$

respectively. Here, $\hat{\underline{R}}_{01}(D_{01})$, $\hat{\underline{R}}_{02}(D_{02})$, and $\hat{\underline{R}}_{12}(D_{02} - D_{01})$ are estimates of $R_{01}(D_{01})$, $R_{02}(D_{02})$, and $R_{12}(D_{02} - D_{01})$, respectively,

The first three terms in Eq. (11.68) are constants. Thus, instead of maximizing Eq. (11.68) we could maximize the two-dimensional function

$$M(D_{01}, D_{02}) = 2|a(0)||a(1)|R_{01}(D_{01}) + 2|a(0)||a(2)|R_{02}(D_{02})$$
$$+ 2|a(1)||a(2)|R_{12}(D_{02} - D_{01}) \quad (11.73)$$

which is really the last three terms of Eq. (11.68). To locate the peak of Eq. (11.73), which is equivalent to maximizing the mean-squared value of the beamformer output, we could use the following algorithm:

Step 1. Select a point (D_{01}, D_{02}) from the set of all points in the $D_{01}D_{02}$ plane.

Step 2. Compute the three cross-correlation estimates (11.59), (11.71), and (11.72).

Step 3. Compute (11.73) for a given set of shading factors.

Step 4. Repeat steps 1–3 until (11.73) becomes a surface.

Step 5. Locate the peak of this surface. Say the peak occurs at the point $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$. Take $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$ as the estimate of $(\Delta_{01}, \Delta_{02})$.

Step 6. Use (11.50) and (11.51) to convert $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$ into range and bearing estimates.

In Fig. 11.26 we show how to implement steps 1–6 by using three cross-correlators. We use the output of these correlators to construct the surface Eq. (11.73), and after locating the peak of this surface we label it $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$. Given that $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$ is an estimate of the intersensor time delays $(\Delta_{01}, \Delta_{02})$, we can find the target's range and bearing by evaluating Eqs. (11.50) and (11.51), respectively, at $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$.

The algorithm described in Fig. 11.26 searches, in an unconstrained fashion, through all possible time delay values (D_{01}, D_{02}) before selecting the optimum pair $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$. With $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$, it then transforms these time delay estimates into range and bearing estimates by Eqs. (11.50) and (11.51), respectively. An alternative approach to range-bearing estimation would be to constrain the time-delay values (D_{01}, D_{02}) in Fig. 11.26 to follow the range-bearing Eqs. (11.50) (11.51). That is, we would first pick a range-bearing pair (r, ϕ), use Eqs. (11.50) and (11.51) to convert this pair to (D_{01}, D_{02}), and then use Fig. 11.26 to maximize the mean-squared value of the beamformer output [Eq. (11.68)]. Thus, the time delay pair that maximized Eq. (11.68) would immediately imply a range-bearing estimate; that is, we would not have to use Eqs. (11.50) and (11.51) after performing the time delay estimation algorithm is Fig. 11.26. In the literature this latter approach is known as the *focused beamformer*, because it constrains the time delay parameters (D_{01}, D_{02}) to "focus" on a specific (r, ϕ) pair [22]. The focused beamformer and Fig. 11.26 provide theoretically equivalent approaches to range-bearing estimation. However, for practical applications, it is sometimes more convenient to perform an unconstrained search (e.g., Fig. 11.26), rather than a constrained search (e.g., the focused beamformer), in the time delay parameters [23].

The time delay estimation algorithm in Fig. 11.26 requires a two-dimensional (2-D) peak detector. Although this represents an optimum algorithm, it would be convenient for practical reasons to reduce the algorithm to several one-dimensional (1-D) peak detectors. For example, reference [24] considers a time delay estimation procedure that processes each sensor pair (e.g., three pairs in Fig. 11.26) by a GCC algorithm. See Figs. 11.24 and 11.25. The time delay estimates from each GCC pair are then judiciously combined to produce the overall time delay estimate $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$ [24]. Thus, this procedure replaces the 2-D peak detection algorithm of Fig. 11.26 by several (e.g., three) judiciously combined GCC algorithms (i.e., 1-D peak detectors).

Let us now discuss a conventional time delay estimation algorithm used for estimating range and bearing with a passive sonar array. As we saw in Section III.B, the three-sensor, linear, passive array solves the range and bearing

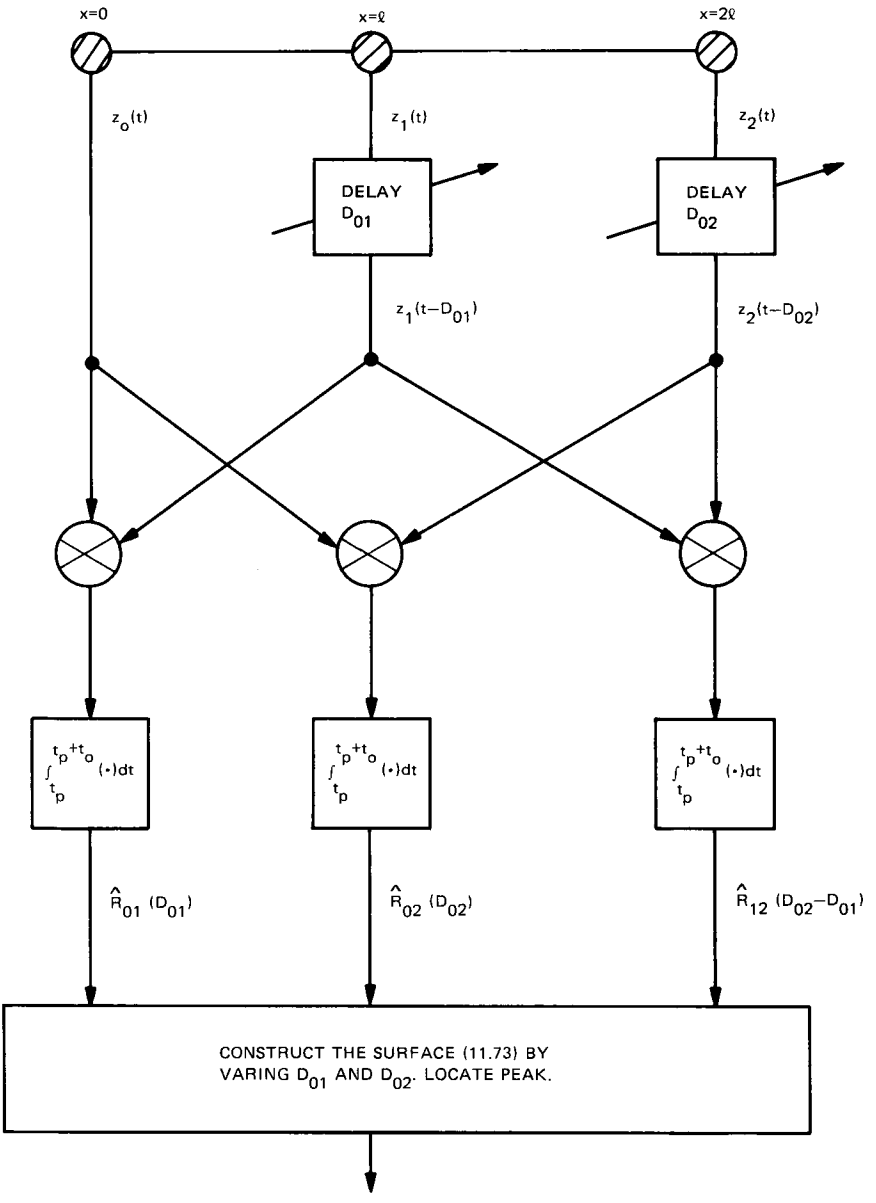


Fig. 11.26. Optimum 2-D algorithm for estimating intersensor time delays.

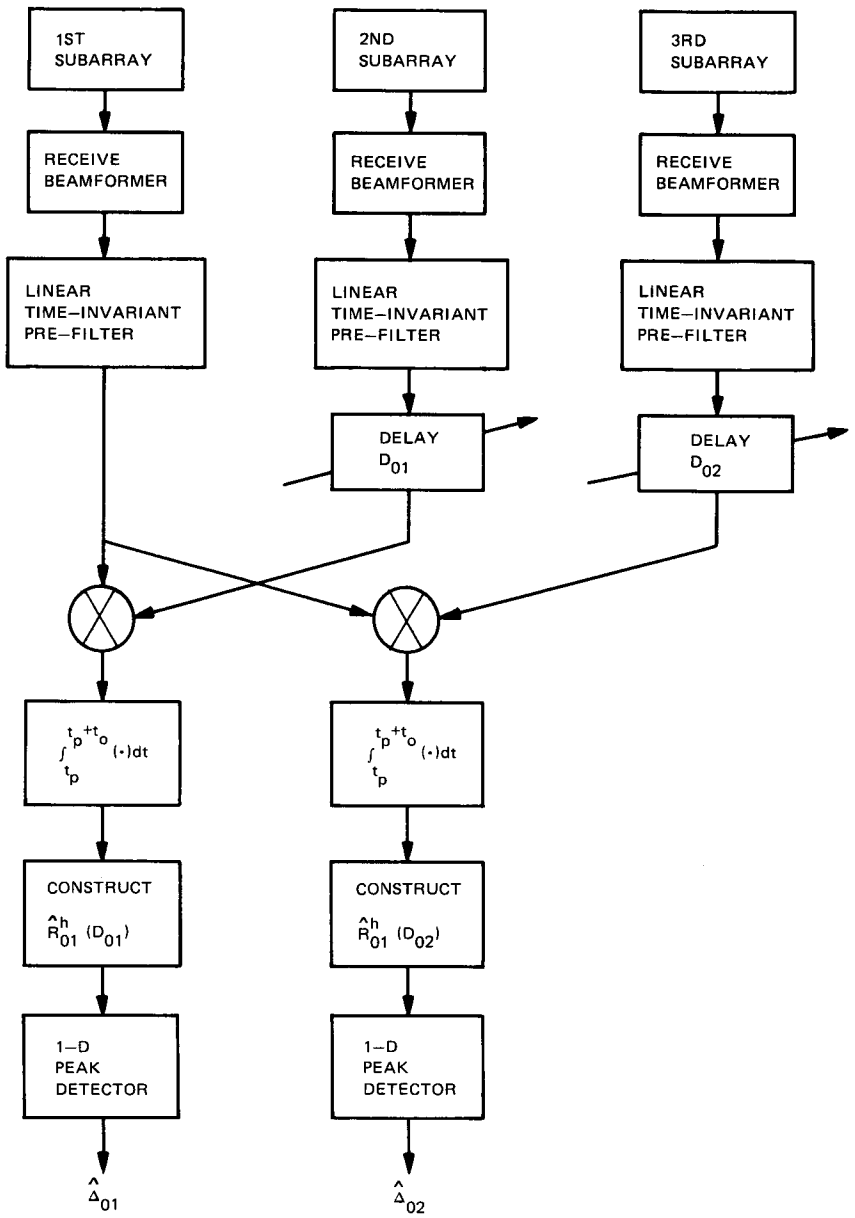


Fig. 11.27. Conventional suboptimum realization for estimating intersensor time delays.

estimation problem. Now instead of using three omnidirectional sensors, let us consider the idea of using three *subarrarys*. (Here, a subarray generally contains several omnidirectional sensors.) First, we will perform receive beamforming on each subarray. Next, we will perform a GCC algorithm on the beamformer outputs associated with the first and second subarrays. The resulting time delay estimate is approximately $\hat{\underline{\Delta}}_{01}$. Then we will perform a GCC algorithm on the beamformer outputs associated with the first and third subarrays. The resulting time delay estimate is approximately $\hat{\underline{\Delta}}_{02}$. Given $(\hat{\underline{\Delta}}_{01}, \hat{\underline{\Delta}}_{02})$, we use Eqs. (11.50) and (11.51) to obtain the range and bearing estimates, respectively. This conventional time delay estimation algorithm is shown in Fig. 11.27.

Although the conventional algorithm of Fig. 11.27 is not the theoretically optimum algorithm (e.g., the GCC version of Fig. 11.26 and the focused beamformer), it is very practical and easy to implement. That is, it can be implemented at a lower cost than the optimum algorithm and performs almost as well [1].

Recall that there are only two independent time delays (e.g., Δ_{01} and Δ_{02}). The remaining time delay, Δ_{12}, can be obtained from Eq. (11.46). Thus, we need not perform a GCC algorithm on the second and third subarrays in Fig. 11.27. However, in practice, this is done for redundancy; that is, it can be used to verify that the algorithm is working properly.

As we have seen, cross-correlation and time delay estimation go hand in hand. In the next section we will discuss the theory of cross-correlation and its relationship to the time delay estimation problem.

IV CROSS-CORRELATION AND ITS RELATIONSHIP TO THE TIME DELAY ESTIMATION PROBLEM

A Cross-Correlation: A Measure of Similarity

Two geometrical vectors, say $\mathbf{x}$ and $\mathbf{y}$, two continuous-time waveforms, say $x(t)$ and $y(t)$, and two discrete-time sequences, say $x(n)$ and $y(n)$, are commonly classified as similar or dissimilar. For example, if two geometrical vectors have the same magnitude and direction, they are similar; specifically, they are equivalent. Sometimes, two vectors could have the same magnitude but different directions. Then the vectors are similar but not equivelent. To quantify the term "similar" for two geometrical vectors, we frequently use the inner product, scalar product, or dot product. For two geometrical vectors $\mathbf{x}$ and $\mathbf{y}$ the *dot product* is defined by

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos \theta \tag{11.74}$$

where θ is the angle between the vectors. When two geometrical vectors have the

same magnitude and are pointed in the same direction, we have $\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}|^2$. For this case the dot product is a maximum, and the two vectors are identical. When two geometrical vectors have the same magnitude but are perpendicular (i.e., $\theta = 90°$), we have $\mathbf{x} \cdot \mathbf{y} = 0$. For this case the dot product is zero, and the two vectors are dissimilar. In any case the *dot product* quantifies the terms *similar* and *dissimilar* for two *geometrical vectors.*

Let us now consider the two real, continuous-time waveforms $x(t)$ and $y(t)$. If the two waveforms have the same amplitude and time distributions, they are *equivalent.* In the time delay estimation problem the two waveforms generally have the same amplitude distributions but different time distributions; that is, the two waveforms can be expressed as $x(t)$ and $y(t) = x(t - \tau)$. To quantify the term "similar" for two continuous-time waveforms $x(t)$ and $y(t)$, we define the inner product

$$\langle x(t), y(t) \rangle = \int_a^b x(t)y(t)\,dt \tag{11.75}$$

where $a \leq t \leq b$ defines the time interval of interest. If $y(t) = x(t - \tau)$, then

$$\langle x(t), y(t) \rangle = \int_a^b x(t)x(t - \tau)\,dt \tag{11.76}$$

When $\tau = 0$ the waveforms $x(t)$ and $x(t - \tau)$ are aligned. For this case the inner product [Eq. (11.76)] is a maximum, and the two waveforms are similar; specifically, they are identical. When either the waveform $x(t)$ or $x(t - \tau)$ does not partially occupy the time interval $[a, b]$ (i.e., they are disjoint in time), then Eq. (11.76) is zero. For this case the waveforms are dissimilar. The inner product Eq. (11.76), plotted as a function of τ, is the autocorrelation function of $x(t)$. Thus, a measure of *similarity* for two continuous-time waveforms is the *inner product* Eq. (11.75).

For discrete-time sequences we define the inner product

$$\langle x(n), y(n) \rangle = \sum_{n=0}^{N-1} x(n)y(n) \tag{11.77}$$

where $0 \leq n \leq N - 1$ is the interval of interest. Again, for time delay problems we have $y(n) = x(n - m)$. For this case Eq. (11.77) can be written as the discrete-time autocorrelation function of $x(n)$. Hence, a measure a *similarity* for two discrete-time sequences is the *inner product* Eq. (11.77).

The cross-correlation function between ergodic sequences $x(n)$ and $y(n)$ is defined by

$$r_{xy}(m) = \sum_{n=\infty}^{\infty} x(n)y^*(n - m) \tag{11.78}$$

where $x(n)$ and $y(n)$ are usually complex sequences. Here, the independent variable m is called the lag variable. Notice that Eq. (11.78) is really the inner

product $\langle x(n), y(n-m)\rangle$, which is a function of the lag m. Thus, the cross-correlation function, like the other inner products, provide a general measure of similarity between two waveforms.

Figure 11.28 shows two different waveforms and the corresponding cross-correlation function. Notice that the cross-correlation function is, in general, not symmetrical and not unimodal. Figure 11.29 shows the cross-correlation function between a waveform $x(n)$ and a delayed version of this waveform, $x(n-D)$. Here, D is an integer value of time delay. For this case the cross-correlation function is always symmetrical and always unimodal, provided the SNR is large. Thus, the time delay estimation problem allows $r_{xy}(m)$ to be unimodal. Hence time delay estimation algorithms that locate the peak of $r_{xy}(m)$ will always select one maximum or peak, which occurs at the time delay D.

In summary, the cross-correlation function is an inner-product-type function that provides a measure of similarity between two waveforms. For time delay estimation problems the cross-correlation function is equivalent to the

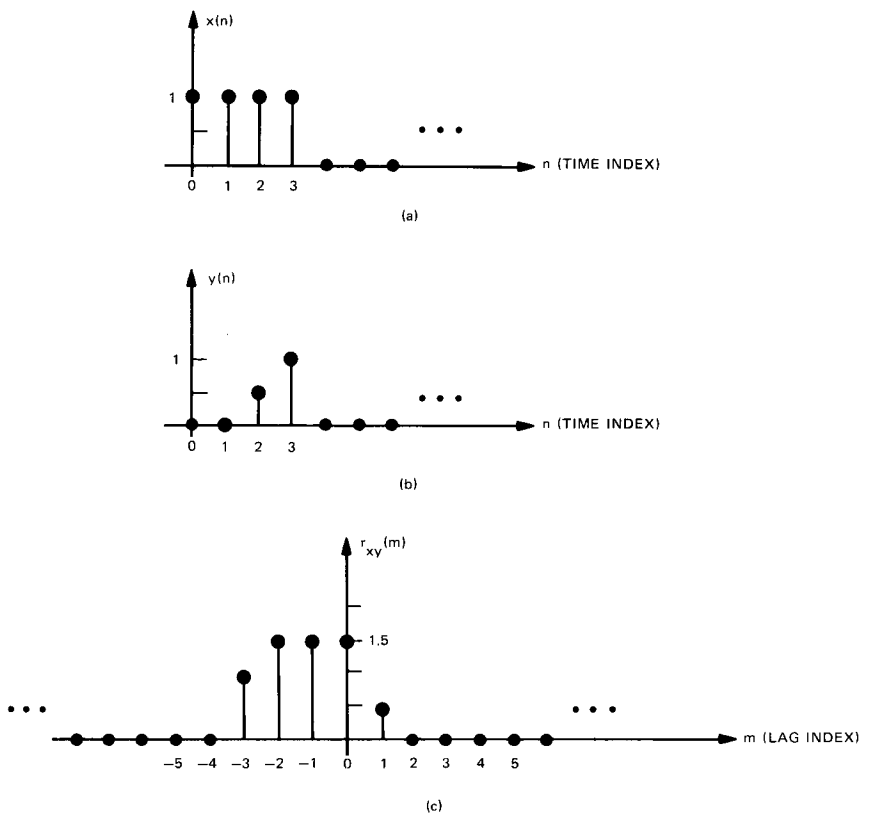


Fig. 11.28. Cross-correlation function. (a) Right-sided waveform $x(n)$; (b) right-sided waveform $y(n)$; (c) $R_{xy}(m)$ = cross correlation function between x(n) and $y(n)$.

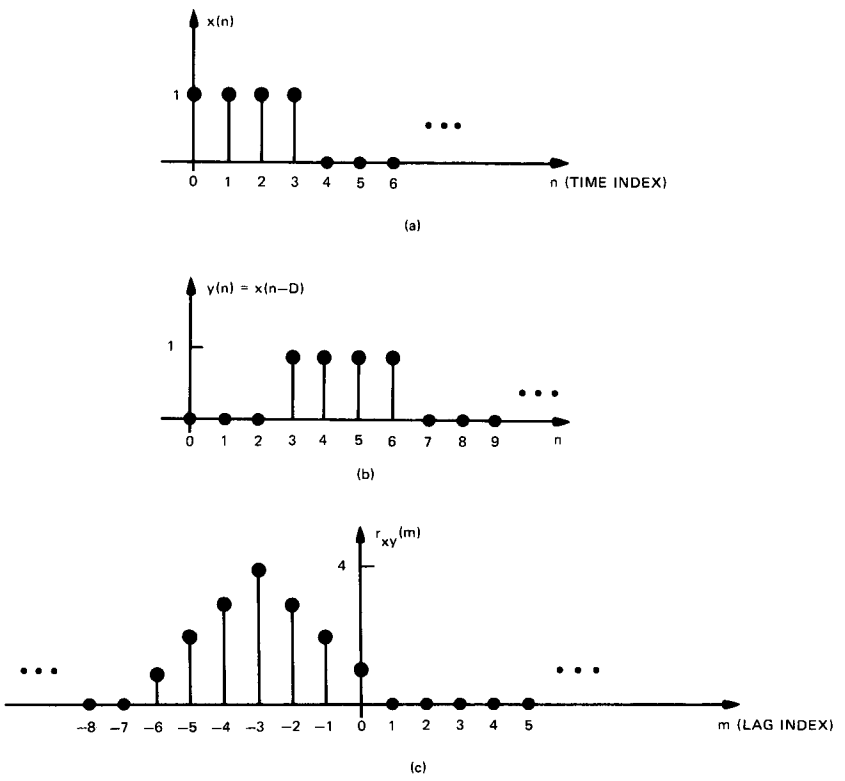


Fig. 11.29. Cross-correlation function. (a) Right-sided waveform $x(n)$; (b) waveform in (a) delayed by D = 3 units; (c) cross correlation function $r_{xy}(m)$ = autocorrelation function $r_{xx}(m + 3)$.

autocorrelation function centered at the true time delay. Hence time delay estimation algorithms that locate the peak of $r_{xy}(m)$ will always obtain a unique global peak (provided the SNR is large) due to the unimodal, symmetrical nature of the autocorrelation function.

B Cross-Correlation and Mean-Squared Criteria

The cross-correlation function arises quite naturally when one attempts to minimize or maximize mean-squared objective functions. For example, in the active sensor problem we assumed that the optimal time delay estimate was the one that minimized Eq. (11.29). Equation (11.29) represents the mean-squared (i.e., average-squared) error between the observation $z(t)$ and the model $(AG^2/R^2)s_T(t - \tau)$. The end result was the cross-correlation function Eq. (11.36); the location of the peak of Eq. (11.36) is the active time delay estimate.

For the passive sensor problem we assumed that the optimal intersensor time delay estimates were the ones that maximized the mean-squared value of the

beamformer output. Refer to Eqs. (11.55) and (11.67). The end result was the cross-correlation functions between the sensor outputs (or subarray beamformer outputs); the peaks of these cross-correlation functions were related to the intersensor time delay estimates.

In general, if one has a mean-squared objective function of the form

$$\begin{aligned} J(D) &= E[x(n) + ay(n - D)]^2 \\ &= E[x^2(n) + 2ax(n)y(n - D) + a^2y^2(n - D)] \end{aligned} \tag{11.79}$$

if $x(n)$ and $y(n)$ are zero mean WSS processes, if a is a constant, and if D is a time delay parameter, then

$$J(D) \propto r_{xy}(D) \tag{11.80}$$

That is, the cross term in the mean-squared objective function [Eq. (11.79)] gives rise to a cross-correlation function. Hence, locating the peak of a cross-correlation function is equivalent to maximizing Eq. (11.79).

In summary, mean-squared objective functions and cross-correlation functions go hand in hand. However, if one chooses to select an objective function that is not of the mean-squared type, then there is no obvious reason why cross-correlation functions should appear in the resulting time delay estimation algorithms. Nevertheless, the mean-squared value of the beamformer output makes physical sense, so maximizing this quantity is not only reasonable but wise in terms of the ease of implementation. For the mathematical properties of cross-correlation functions, see Chapter 1.

V THE IMPLEMENTATION OF SOME TIME DELAY ESTIMATION ALGORITHMS USING THE FAST FOURIER TRANSFORM (FFT)

A From Theory to Implementation

In previous sections we covered the basic theory and functional flow diagrams for both the active and passive time delay estimation algorithms. We now discuss some practical implementation issues.

With the advent of high-speed digital electronics, there has been a steady transition from analog signal processing to digital signal processing. Today, many radar and sonar systems are moving toward computer-controlled, automated software-based systems. Both radar and sonar signal processing have been heavily influenced by the commerical availability of gate arrays, CMOS devices, high-speed digital signal processing chips, microprocessors, microcomputers, and FFT algorithms. The implementation of many time delay estimation algorithms rely quite heavily on the FTT. Thus, we will discuss the role of the FFT in both the active and passive problems.

B Implementation Issues for the Active Case

The front end of an active radar or sonar receiver generally contains a receive beamformer, analog filters, amplifiers, and dynamic range controllers [e.g., automatic gain control (AGC)]. These devices prepare the radar or sonar analog signal for digital signal processing. For a stationary point target located in a homogeneous, isotropic, lossless medium, the analog signal, after front-end processing, has the form

$$\underline{z}(t) = Ms_T(t - \tau) + \underline{v}(t) \tag{11.81}$$

for $t_p \leq t \leq t_p + t_0$. Here, t_p is the duration of the transmitted pulse, t_0 is the length of the observation interval, $s_T(t)$ is the transmitted waveform [refer to Eq. (11.1)], M is a constant, $\underline{v}(t)$ is a zero-mean, white Gaussian noise process, and τ is the desired time delay.

Equation (11.81) is an analog signal with a corresponding analog spectrum $\underline{Z}(F)$. In active radar and sonar $\underline{z}(t)$ generally has a narrowband spectrum. That is, $\underline{Z}(F)$ is centered about some carrier frequency F_0 (Hz) and has a bandwidth W (Hz) that is much smaller than F_0 [25]. For example, Figure 11.30 shows a typical magnitude spectrum of a narrowband process $\underline{z}(t)$. Let us now discuss the analog-to-digital conversion (ADC) of this narrowband process.

Suppose $\underline{z}(t)$ was associated with an X-band radar whose carrier frequency was $F_0 = 9 \times 10^9$ Hz $= 9$ GHz and whose bandwidth was $W = 1 \times 10^6$ Hz $=$ 1 MHz. Naively, we could sample this process at the Nyquist rate of $2(F_0 + W/2)$ Hz or 18 GHz. However, the current ADC technology cannot support this sampling requirement. Further, if the observation interval was $t_0 = 0.75 \times 10^{-3}$ s, then sampling $\underline{z}(t)$ at 18 GHz would result in 13.5×10^6 digital samples, an enormous amount of data.

To circumvent these problems and make the ADC practical, we first translate the narrowband process $\underline{z}(t)$ down to baseband (i.e., a lowpass spectrum centered

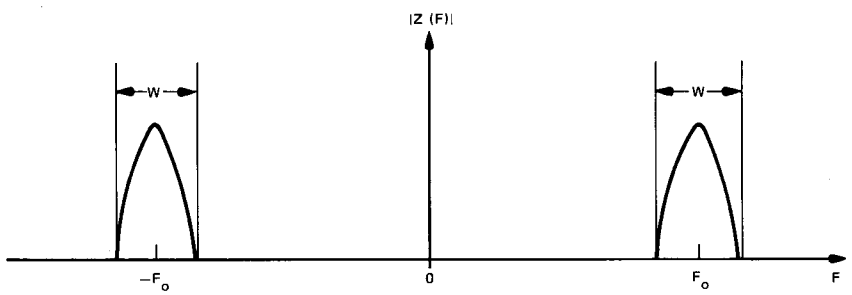


Fig. 11.30. Magnitude spectrum of a narrowband process $\underline{z}(t)$.

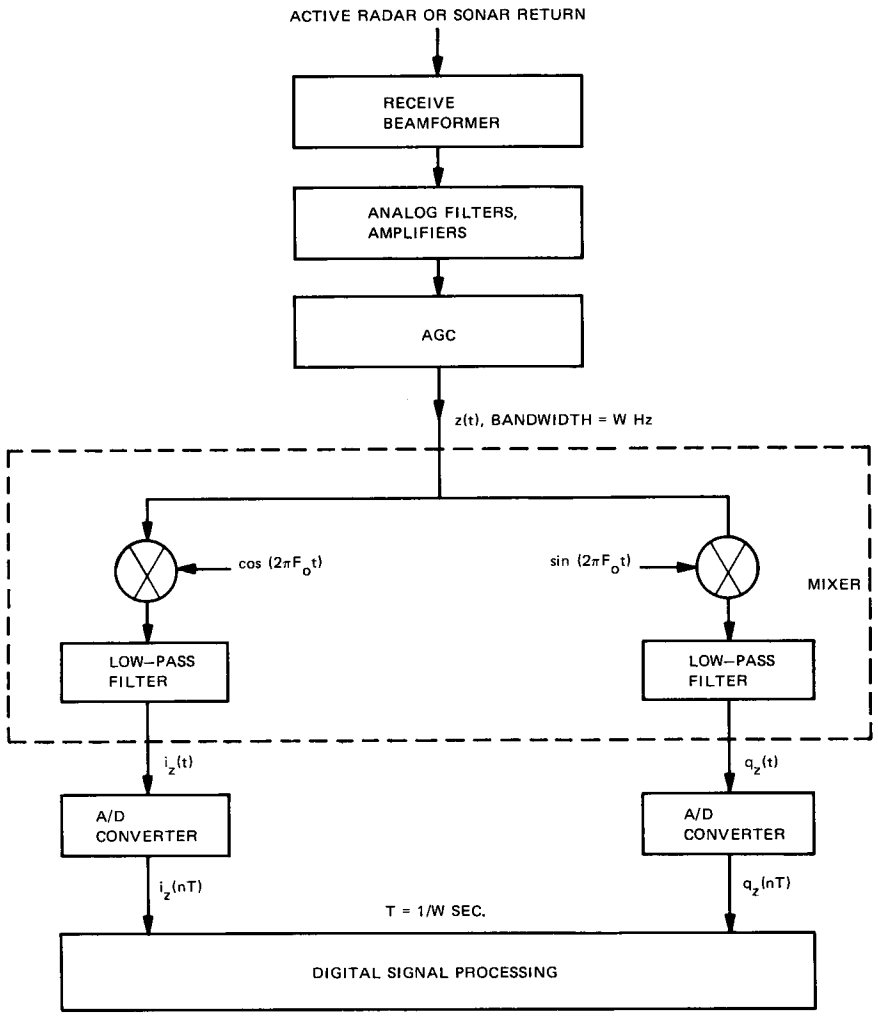


Fig. 11.31. Block diagram of the narrowband-to-baseband translation process.

at $F = 0$), then sampled at a much lower rate. The narrowband-to-baseband translation, common practice in active radar and sonar, is generally done with analog multipliers and filters. Figure 11.31 is a block diagram description of this.

Figure 11.31 shows that the narrowband-to-baseband translation, commonly referred to as *mixing*, decomposes $\underline{z}(t)$ into an *in-phase* (I) *component* $\underline{i}_z(t)$ and a *quadrature* (Q) *component* $\underline{q}_z(t)$. Mathematically, the mixing process generates a lowpass *complex analog signal*

$$\underline{\tilde{z}}(t) = \underline{i}_z(t) + j\underline{q}_z(t) \tag{11.82}$$

for $t_p \leq t \leq t_p + t_0$. The magnitude of $\underline{\tilde{z}}(t)$ is

$$|\underline{\tilde{z}}(t)| = \{[\underline{i}_z(t)]^2 + [\underline{q}_z(t)]^2\}^{1/2} \tag{11.83}$$

and the phase angle associated with $\underline{\tilde{z}}(t)$ is

$$\arg[\underline{\tilde{z}}(t)] = \tan^{-1}\left\{\frac{\underline{q}_z(t)}{\underline{i}_z(t)}\right\} \tag{11.84}$$

Since each analog component of $\underline{\tilde{z}}(t)$ has a lowpass spectrum with cutoff frequency $W/2$ Hz, this means that $i_z(t)$ and $q_z(t)$ can be sampled at WHz. Thus, the above X-band radar would require 10 MHz ADCs in the I and Q paths. The current ADC technology can easily support this sampling requirement. Further, for $t_0 = 0.75$ ms, the 10-MHz sampling rate generates 7500 I samples and 7500 Q samples instead of the 13.5×10^6 z-samples generated by the 18-GHz sampling rate. This represents a data compression ratio of approximately 1000 to 1.

At this point the active radar or sonar return has been front-end filtered, mixed, and converted to discrete-time form. We assume that the time origin has been shifted so that $0 \leq t \leq t_0$. Hence, if $t_0 = (N-1)T$, then we have the N complex samples $\underline{\tilde{z}}(nT)$ $(n = 0, 1, \ldots, N-1)$. Equivalently, we now have the N in-phase samples $\underline{i}_z(nT)$ $(n = 0, 1, \ldots, N-1)$ and the N quadrature samples $\underline{q}_z(nT)$ $(n = 0, 1, \ldots, N-1)$. The radar or sonar return is now ready for digital signal processing (Fig. 11.31).

Recall the discussion in Section II.C.2. For real data $z(t)$ and a stationary point target in an ideal medium, we showed that the least squares solution to the active time delay estimation problem was to locate the peak of the correlator output. Refer to Eq. (11.36) and Fig. 11.19. For complex data $\underline{\tilde{z}}(t)$ [refer to Eq. (11.82) and Fig. 11.31], Eq. (11.36) becomes

$$\underline{\tilde{y}}(\tilde{\tau}) = \underline{i}_y(\tilde{\tau}) + j\underline{q}_y(\tilde{\tau}) = \int_{t_p}^{t_p+t_0} \underline{\tilde{z}}(t)\tilde{s}_T^*(t-\tilde{\tau})\,dt \tag{11.85}$$

Here, $\underline{\tilde{z}}(t)$ and

$$\tilde{s}_T(t) = s_{Ti}(t) + js_{Tq}(t) \tag{11.86}$$

are obtained by mixing $\underline{z}(t)$ and $s_T(t)$, respectively. See Fig. 11.31. The quantities $\underline{i}_y(\tilde{\tau})$ and $\underline{q}_y(\tilde{\tau})$ are the real and imaginary components, respectively, of the complex correlator output [Eq. (11.85)]. If we implement Eq. (11.85) on a digital computer than $\underline{\tilde{y}}(\tilde{\tau})$ can be approximated by

$$\underline{\tilde{y}}(mT) = \sum_{n=0}^{N-1} \underline{\tilde{z}}(nT)\tilde{s}_T^*(nT - mT) \tag{11.87}$$

where $\tilde{\tau} = mT$ $(m = 0, \pm 1, \ldots, \pm M \mp 1)$, $t = nT$ $(n = 0, 1, \ldots, N-1)$, and T is the uniform sampling increment. Generally, the length of $\tilde{s}_T(nT)$, say S, is shorter than N, so $M > N > S$.

Many active radars and sonars compute Eq. (11.87) by the FFT. Specifically, if

$$\underline{\tilde{Z}}(k) = \sum_{n=0}^{NL-1} \underline{\tilde{z}}(n)e^{-j2\pi nk/NL} \qquad (k = 0, 1, \ldots, NL-1) \tag{11.88}$$

represents the discrete Fourier transform (DFT) of $\underline{\tilde{z}}(n) = \underline{\tilde{z}}(nT)$ and

$$\tilde{S}_T(k) = \sum_{n=0}^{NL-1} \tilde{s}_T(n)e^{-j2\pi nk/NL} \qquad (k = 0, 1, \ldots, NL-1) \tag{11.89}$$

represents the DFT of $\tilde{s}_T(n) = \tilde{s}_T(nT)$, then

$$\underline{\tilde{Y}}(k) = \underline{\tilde{Z}}(k)S_T^*(k) \qquad (k = 0, 1, \ldots, NL-1) \tag{11.90}$$

represents the discrete cross-spectrum between $\underline{\tilde{z}}(n)$ and $\tilde{s}_T(n)$, provided that t_0 is "large." The DFT

$$\underline{\tilde{y}}(m) = \frac{1}{NL}\sum_{k=0}^{NL-1} \underline{\tilde{Y}}(k)e^{j2\pi nk/NL} \qquad (m = 0, 1, \ldots, NL-1) \tag{11.91}$$

gives the correlator output Eq. (11.87). We note that the size of the DFTs, namely NL, is usually much larger than N because $\underline{\tilde{z}}(n)$ and $\tilde{s}_T(n)$ are appended with $NL - N$ and $NL - S$ zeros, respectively. This zero-padding is necessary to avoid the circular convolution effects that one encounters when using DFTs to perform convolution or correlation. Refer to Chapter 1, Section VII.D. Now once the zero-padding and DFT size NL have been established, the FFT algorithm (see Chap. 7) is used to implement the DFTs (11.88) and (11.89). Next, the cross-spectrum Eq. (11.90) is computed; then the inverse FFT algorithm is used to implement the DFT Eq. (11.91). At this point the complex correlator output $\underline{\tilde{y}}(m)$ is available for further processing.

The time delay τ appears in both the magnitude and phase of $\underline{\tilde{y}}(m)$. However, for active radar and sonar the phase of $\underline{\tilde{y}}(m)$ is a very sensitive function of τ; that is, the smallest change in τ produces large phase fluctuations. Therefore the phase of $\underline{\tilde{y}}(m)$ is generally not stable, so it is not used for time delay estimation. On the other hand, the magnitude of $\underline{\tilde{y}}(m)$ is less sensitive to small changes in τ. Consequently, the magnitude of $\underline{\tilde{y}}(m)$ is generally used for time delay estimation. Thus, the FFT processing gives $\underline{\tilde{y}}(m)$, and the peak location of

$$|\underline{\tilde{y}}(m)| = |\underline{i}_y(m) + j\underline{q}_y(m)| = \left|\sum_{n=0}^{N-1} \underline{\tilde{z}}(n)\tilde{s}_T^*(n-m)\right| \tag{11.92}$$

gives the final active time delay estimate. Figure 11.32 summarizes the above discussion.

When no noise is present [e.g., $v(t) = 0$], the magnitude of $\tilde{y}(\tilde{\tau})$ can be written as

$$|\tilde{y}(\tilde{\tau})| = \left|\int_{-\infty}^{\infty} M\tilde{s}_T(t-\tau)\tilde{s}_T^*(t-\tilde{\tau})\,dt\right| \tag{11.93}$$

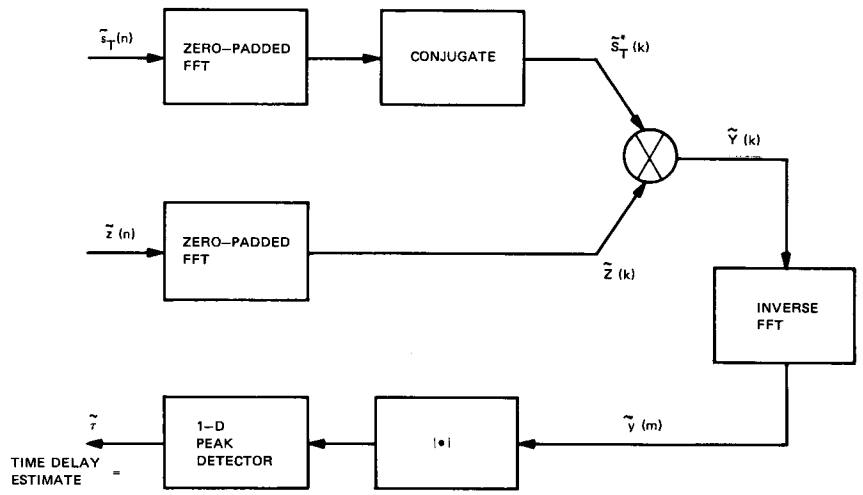


Fig. 11.32. FFT processing.

Under the no-noise assumption, Eq. (11.93) is entirely dependent on the properties of the transmitted waveform $s_T(t)$. Further, $\tilde{y}(\tilde{\tau})$ achieves its maximum value when $\tilde{\tau} = \tau$ (i.e., when $\tilde{\tau}$ equals the true time delay τ). Notice that this maximum value is proportional to the energy E_T in $s_T(t)$.

Up to now we have assumed that the active radar or sonar operated against a stationary point target. For constant-velocity point targets (see [26]), Eq. (11.93) can be generalized to

$$|\tilde{y}(\tilde{\tau}, \tilde{F}_d)| = \left| \int_{-\infty}^{\infty} M\tilde{s}_T(t-\tau)\tilde{s}^*(t-\tilde{\tau})e^{-j2\pi(F_d-\tilde{F}_d)t}\,dt \right| \tag{11.94}$$

where F_d is the *Doppler shift*. If Eq. (11.94) is normalized such that

$$A(\tilde{\tau} = \tau, \tilde{F}_d = F_d) = \frac{|\tilde{y}(\tau, F_d)|}{|M|E_T = 1} \tag{11.95}$$

we obtain

$$A(\tilde{\tau}, \tilde{F}_d) = \left| \int_{-\infty}^{\infty} \tilde{s}_T(t-\tau)\tilde{s}_T^*(t-\tilde{\tau})e^{-j2\pi(F_d-\tilde{F}_d)t}\,dt \right| \tag{11.96}$$

The square of Eq. (11.96) is commonly referred to as the *ambiguity function*, which was originally introduced by Ville [27]. The ambiguity function represents the magnitude squared of the complex correlator output or matched filter output under the constant-velocity point target and no-noise assumptions. Notice that $A(\tilde{\tau}, \tilde{F}_d)$ is a 2-D surface with a global maximum at the point (τ, F_d). Thus, we can perform time delay and Doppler estimations simultaneously by locating the global peak of Eq. (11.96). Further, since $s_T(t)$ controls the shape of the ambiguity function, one can judiciously select an "optimal" transmitted waveform by means

of Eq. (11.96). In any event, the ambiguity function plays a major role in radar and sonar signal processing. References [25] and [28] provide good discussions on the ambiguity function as well as numerous 3-D pictures.

C Implementation Issues for the Passive Case

The front end of a passive radar or sonar receiver generally contains a receive beamformer, analog filters, amplifiers, and AGC circuitry. Just like the active case these devices prepare the analog signals for digital signal processing.

As we saw in Section III.C, the GCC algorithm plays an important role in the passive time delay estimation problem. Figure 11.25 is a block diagram for implementing the GCC algorithm in the frequency domain. Since we are interested in an FFT implementation, our discussion will be centered around Fig. 11.25.

The analog signals $\underline{z}_0(t)$ and $\underline{z}_1(t)$ that experience that front-end processing are generally *broadband*. That is, the spectra $\underline{Z}_0(F)$ and $\underline{Z}_1(F)$ are generally spread out over a fairly wide bandwidth. Although these bandwidths are wide, they are usually compatible with practical ADCs. Thus, the samples $z_0(nT)$ $(n = 0, 1, \ldots, N-1)$ and $z_1(nT)$ $(n = 0, 1, \ldots, N-1)$ are usually obtained by direct analog-to-digital conversion of $\underline{z}_0(t)$ and $\underline{z}_1(t)$, respectively. For example, a passive sonar receiver with a 0–10 kHz bandwidth would require at least a 20-kHz ADC. For an observation interval of $t_0 = 25.6$ msc, this means that the length of $\underline{z}_0(nT)$ and $\underline{z}_1(nT)$ is $N = 512$ points. Here, $T = \frac{1}{20}$ kHz $= 50$ μs is the uniform sampling increment.

Given that

$$\underline{Z}_0(k) = \sum_{n=0}^{NL-1} \underline{z}_0(n) e^{-j2\pi nk/NL} \qquad (k = 0, 1, \ldots, NL-1) \tag{11.97}$$

is the DFT of $\underline{z}_0(n) = \underline{z}_0(nT)$ and

$$\underline{Z}_1(k) = \sum_{n=0}^{NL-1} \underline{z}_1(n) e^{-j2\pi nk/NL} \qquad (k = 0, 1, \ldots, NL-1) \tag{11.98}$$

is the DFT of $\underline{z}_1(n) = \underline{z}_1(nT)$, then the cross-spectrum between $\underline{z}_0(n)$ and $\underline{z}_1(n)$ is

$$\hat{\underline{G}}_{01}(k) = \underline{Z}_0(k)\underline{Z}_1^*(k) \qquad (k = 0, 1, \ldots, NL-1) \tag{11.99}$$

For a given spectral weighting function $W(k)$ $(k = 0, 1, \ldots, NL-1)$, the GCC algorithm modifies the cross-spectrum according to the multiplicative rule

$$\hat{\underline{G}}_{01}^h(k) = W(k)\hat{\underline{G}}_{01}(k) \qquad (k = 0, 1, \ldots, NL-1) \tag{11.100}$$

Finally, the IDFT

$$\hat{R}_{01}^h(n) = \frac{1}{NL} \sum_{k=0}^{NL-1} \hat{\underline{G}}_{01}^h(k) e^{j2\pi nk/NL} \qquad (n = 0, 1, \ldots, NL-1) \tag{11.101}$$

gives the GCC function. The location of the peak of Eq. (11.101) gives the "optimal" time delay estimate.

For an FFT implementation of Fig. 11.25, the DFTs Eqs. (11.97) and (11.98) can be implemented by an FFT algorithm. Similarly, Eq. (11.101) can be implemented by an IFFT algorithm.

As we saw, Fig. 11.27 can be used to solve the passive localization problem. The practical aspect of Fig. 11.27 is that it involves two independent GCC algorithms. Since we have just shown how a GCC algorithm can be implemented by FFTs, it follows that Fig. 11.27 has an FFT implementation.

In summary, we can use the FFT algorithm to solve the active and passive time delay estimation problems. When applying the FFT to correlation or convolution problems, be careful to avoid the effects of circular convolution. This is why the FFT input sequences are zero-padded. That is, the FFT size NL is composed of N data points and a sufficient number of zeros.

VI ALGORITHM PERFORMANCE

A Performance for the Active Case

In Section II we discussed the time delay estimation problem for active sensors. Recall that we considered a simple backscattering model (i.e., a stationary point target model) and an additive white noise model. Although these models represent a gross simplification of a real-world target echo, they keep the mathematics tractable, so we can obtain a feel for the problem.

Let us now make the following assumptions:

1. The transmitted signal $s_T(t)$ has the form (11.1).
2. The backscattering process is described by a single stationary point target located in a homogeneous, isotropic, lossless medium. Thus, the noise-free target echo is simply a scaled time-delayed version of (11.1), there are no other scatterers, and no multipath.
3. The additive noise model is a zero-mean, WSS, Gaussian white noise process with autocorrelation function (11.37).
4. The transmitted pulse width t_p is much smaller than the observation interval t_0.
5. The MRA of the transmit beam is pointed directly at the target, so the target's direction is exactly known.
6. The time delay estimate $\hat{\tau}$ is obtained by locating the peak of the magnitude of the cross-correlation function. The phase of the cross-correlation function is not used.

Based on these assumptions and the additional assumption of a large SNR

(e.g., greater than 15 dB), the variance of the time delay estimate $\hat{\tau}$ is [16, 26]

$$\text{var}(\hat{\tau}) = \frac{1}{d^2\beta^2} \tag{11.102}$$

where

$$d^2 \equiv \frac{2E_T}{N_0} = \frac{E_T}{N_0/2} \tag{11.103}$$

$$E_T = \int_0^{tp} s_T^2(t)\,dt = \int_{-\infty}^{\infty} |S_T(F)|^2\,dF \tag{11.104}$$

and

$$\beta^2 = \frac{\displaystyle\int_{-\infty}^{\infty} (2\pi F)^2 |S_T(F)|^2\,dF}{\displaystyle\int_{-\infty}^{\infty} |S_T(F)|^2\,dF} \tag{11.105}$$

Here, E_T is the energy of the transmitted signal, $N_0/2$ is the magnitude of the white noise power spectral density (PSD), and $S_T(F)$ is the Fourier (analog) spectrum of $s_T(t)$. It is common practice to interpret d^2 as a measure of SNR and β^2 as a mean-squared measure of transmitted signal bandwidth [16]. Thus, if the square root of $\text{var}(\hat{\tau})$ represents the time delay estimation accuracy, then large SNRs (e.g., greater than 15 dB) and large bandwidths produce accurate time delay estimates.

Such results are intuitively pleasing. For example, Fig. 11.8(b) shows a stationary point target echo in additive white noise. We see that when the signal amplitude is large compared to the noise amplitudes and when the rise time of the pulse is small (which translates into a large signal bandwidth), we can intuitively argue that the time delay estimate $\hat{\tau}$ will be more accurate. Equation (11.102) [or the square root of Eq. (11.102)] quantifies these intuitive notions.

Performance for the Passive Case B

In Section III we discussed the time delay estimation problem for passive sensors. Recall that to perform target localization by a passive array, we must be able to accurately estimate the intersensor time delays Δ_{01} and Δ_{02} [see Eqs. (11.44), (11.45) and Fig. 11.21].

Let us now make the following assumptions:

1. The passive array receives a radiated signal $\underline{s}(t)$ from a stationary target. When appropriate, the radiated signal $\underline{s}(t)$ belongs to a zero mean, WSS Gaussian process with power spectral density $G_{ss}(F)$.

2. The passive array is stationary, and the elements or sensors are all collinear.
3. The additive noise $\underline{v}(t)$ in the received signal model is a zero-mean, WSS Gaussian white noise process with autocorrelation function (11.37) and power spectral density $G_{vv}(F)$.
4. The radiating target and linear passive array are located in a homogeneous, isotropic, and lossless medium. Further, there is no multipath corruption of the received signal.
5. The radiated signal and additive noise process are uncorrelated.

Based on these assumptions, the variance of the time delay estimate $\hat{\underline{\tau}}$ between two sensors is [17]

$$\text{var}(\hat{\underline{\tau}}) = \left\{ 2t_0 \int_0^{\infty} (2\pi F)^2 \left[\frac{|\gamma(F)|^2}{1 - |\gamma(F)|^2} \right] dF \right\}^{-1} \tag{11.106}$$

where

$$|\gamma(F)|^2 = \frac{G_{ss}^2(F)}{[G_{ss}(F) + G_{vv}(F)]^2} \tag{11.107}$$

and t_0 is the length of the observation interval.

For the active case, $\text{var}(\hat{\underline{\tau}})$, given by Eq. (11.102), was self-explanatory. That is, given a transmitted signal $s_T(t)$, d^2 and β^2 clearly defined the notions of SNR and bandwidth. However, for the passive case it is not clear how Eq. (11.106) relates to SNR and bandwidth. Thus, to make things clear, we will consider a simple example.

Example 5. Let us assume that both the signal and noise PSDs are constant over a finite-length frequency band and zero otherwise. Refer to Fig. 11.33. Under this condition Eq. (11.106) reduces to

$$\text{var}(\hat{\underline{\tau}}) = \frac{3}{8\pi^2 t_0} \frac{1 + 2(S_0/N_0)}{(S_0/N_0)^2 (F_2^3 - F_1^3)} \tag{11.108}$$

Since the signal power S is

$$S = \int_{-\infty}^{\infty} G_{ss}(F)\, dF = S_0(F_2 - F_1) \tag{11.109}$$

and the noise power N is

$$N = \int_{-\infty}^{\infty} G_{vv}(F)\, dF = N_0(F_2 - F_1) \tag{11.110}$$

Eq. (11.108) can be rewritten in terms of signal-power-to-noise-power ratios; that is,

$$\text{var}(\hat{\underline{\tau}}) = \frac{3}{8\pi^2 t_0} \frac{1 + 2(S/N)}{(S/N)^2 (F_2^3 - F_1^3)} \tag{11.111}$$

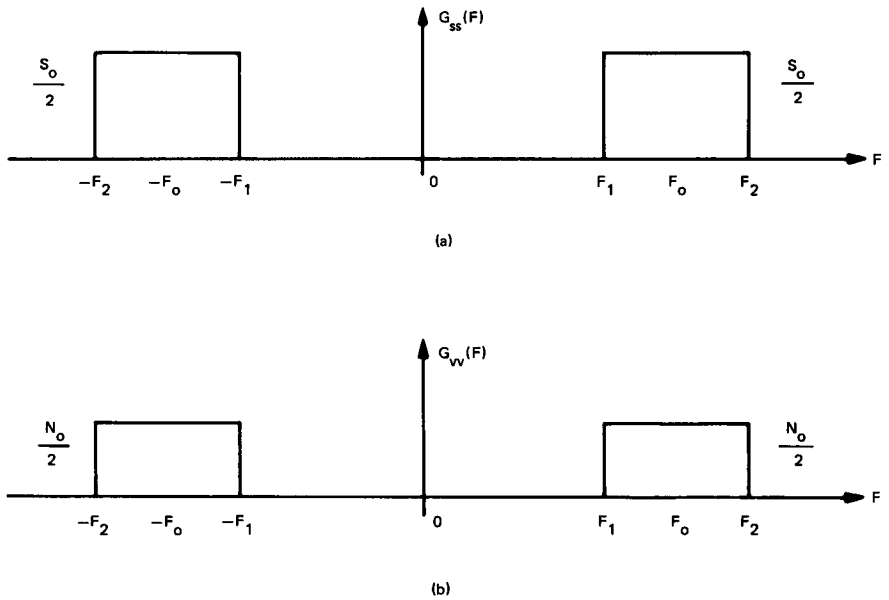


Fig. 11.33. Assumed PSDs for (a) the signal $s(t)$ and (b) the noise $\underline{v}(t)$.

For low SNRs (i.e., SNR $\ll$ 1) we have

$$\text{var}(\hat{\underline{\tau}}) \simeq \frac{3}{8\pi^2 t_0} \frac{1}{(S/N)^2(F_2^3 - F_1^3)} \tag{11.112}$$

whereas for high SNRs (i.e., SNR $\gg$ 1) we have

$$\text{var}(\hat{\underline{\tau}}) \simeq \frac{3}{8\pi^2 t_0} \frac{2}{(S/N)(F_2^3 - F_1^3)} \tag{11.113}$$

In summary:

1. Like the active case, large SNRs and large bandwidths produce accurate time delay estimates.
2. For the low SNR case the accuracy of $\hat{\underline{\tau}}$ [or the square root of (11.112)] is inversely proportional to SNR.
3. For the high SNR case the accuracy of $\hat{\underline{\tau}}$ [or the square root of (11.113)] is inversely proportional to $\sqrt{\text{SNR}}$.

Note that Eq. (11.102) for the active case and Eq. (11.106) for the passive case represent a lower bound on the variance of the time delay estimate $\hat{\underline{\tau}}$. That is, under the given assumptions, Eqs. (11.102) and (11.106) represent the best we can do. In the theory of maximum-likelihood estimation this lower bound is commonly referred to as the *Cramèr–Rao lower bound.* The derivation of the Cramèr–Rao lower bound can be found in [12, 16, 29].

In terms of the passive localization problem, $\text{var}(\hat{\underline{\Delta}}_{01}) = \text{var}(\hat{\underline{\Delta}}_{02}) = \text{var}(\hat{\tau})$ [Eq. (11.106)]. Recall that $\hat{\underline{\Delta}}_{01}$ was the time delay estimate between sensors 0 and 1, and $\hat{\underline{\Delta}}_{02}$ was the time delay estimate between sensors 0 and 2. The variance of the range r [Eq. (11.50)] and of the bearing ϕ [Eq. (11.51)] can be found in [1]. It turns out that an accurate range estimate requires very accurate time delay estimates, whereas an accurate bearing estimate is less sensitive to time delay accuracy.

The performance of time delay estimation algorithms has been the subject of many research papers. For a good overview of the active and passive cases, see [30] and the references there.

C A Numerical Example

For the active case we showed that Fig. 11.32 was an FFT implementation of Fig. 11.19. Let us now show how to compute the necessary quantities shown in Fig. 11.32 and the time delay estimate $\hat{\underline{\tau}}$.

First, we compute the complex signal

$$\tilde{s}_T(n) = s_{Ti}(n) + js_{Tq}(n) \tag{11.114}$$

which is the complex demodulated version of the transmitted waveform $s_T(t)$. Figure 11.34(a) shows a ficticious transmitted pulse [i.e., $s_T(t)$] of duration $t_p = 678.9$ μs and rise time 234.5 μs. The narrowband spectrum of $s_T(t)$ is shown in Fig. 11.34(b). Figure 11.34(c) shows the complex signal $\tilde{s}_T(n)$ [see Eq. (11.114)]. The narrowband-to-baseband translation of $s_T(t)$ [i.e., the complex demoduation of $s_T(t)$] was done digitally, with the FIR linear-phase lowpass filter shown in Fig. 11.35 (see Chapters 2 and 3). The sampling frequency was 512 kHz.

Next, bandlimited Gaussian white noise was added to the waveform in Fig. 11.34(a) to produce the received noisy echo $\underline{z}(t)$. Refer to Fig. 11.36(a). The duration of this noisy echo, or the length of the observation interval t_0, was $t_0 = 1000$ μs, or 1 ms. Figure 11.36(b) shows the narrowband spectrum of $\underline{z}(t)$, and Fig. 11.36(c) shows the complex demodulated signal

$$\tilde{\underline{z}}(n) = z_i(n) + j\underline{z}_q(n) \tag{11.115}$$

Again, the narrowband-to-baseband translation of $z(t)$ was done digitally, with the FIR filter in Fig. 11.35; the sampling frequency was 512 kHz.

To form the cross-spectrum,

$$\underline{\tilde{Y}}(k) = \tilde{S}_T^*(k)\underline{\tilde{Z}}(k) \tag{11.116}$$

we must compute the FFT of the complex sequence $\tilde{s}_T(n)$ and the FFT of the complex sequence $\tilde{\underline{z}}(n)$. These FFTs must be zero-padded in order to avoid the effects of circular convolution, as explained in Section V.B. Thus, 1 ms of data gives 512 data points, so a zero-pad of 512 zeros requires a 1024-point FFT.

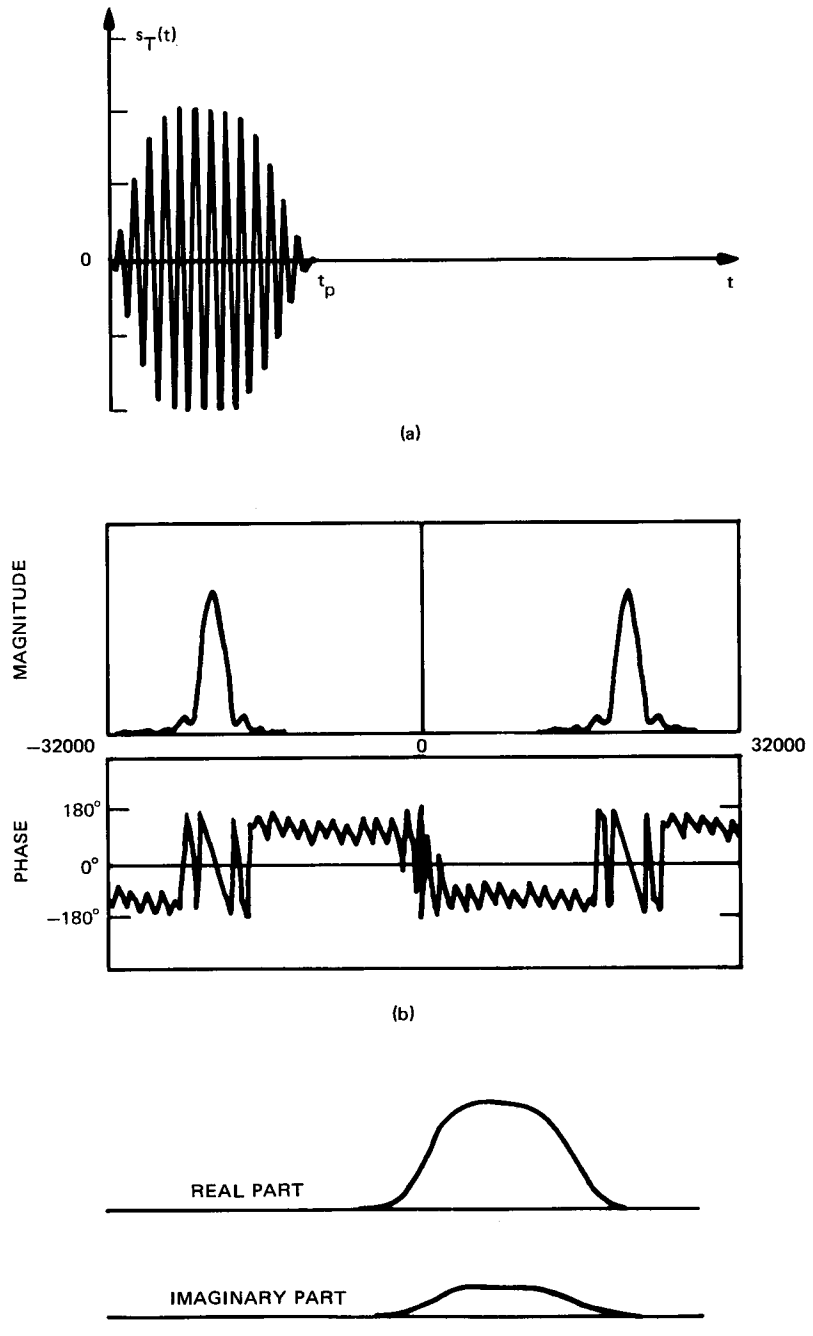
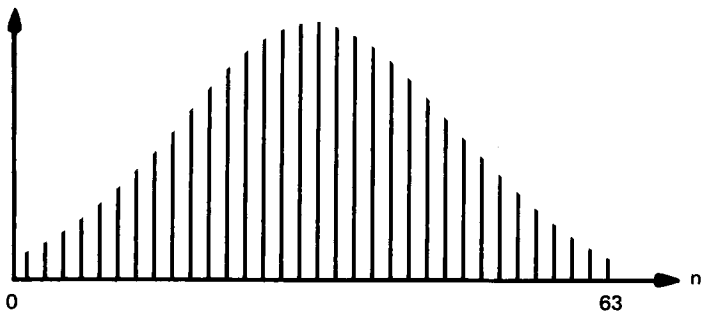
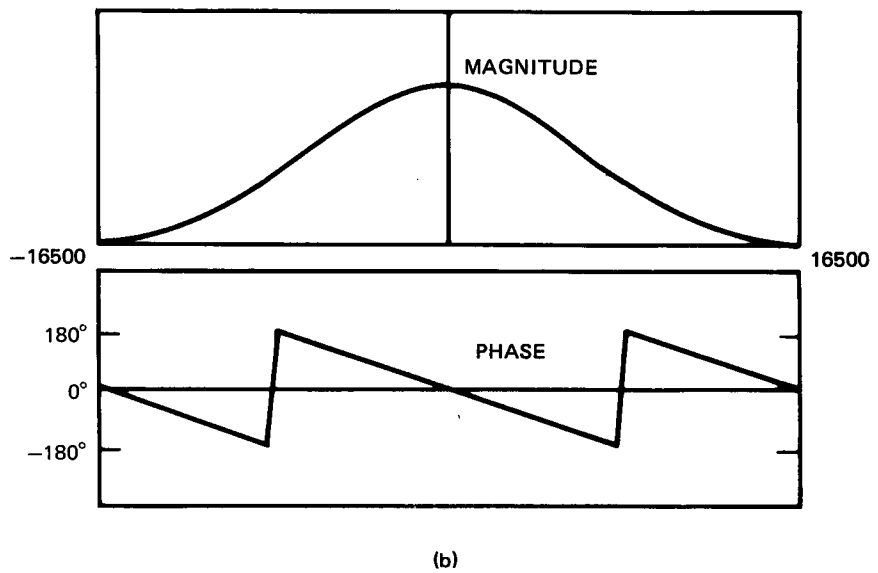


Fig. 11.34. (a) Transmitted pulse. (b) Magnitude and phase spectra. (c) Complex demodulation.



(a)



(b)

Fig. 11.35. (a) A 64 tap, linear-phase FIR lowpass digital filter. (b) Frequency response for (a).

Figure 11.37(a) shows a 1024-point FFT of $\tilde{s}_T(n)$, and Fig. 11.37(b) shows a 1024-point FFT of $\tilde{z}(n)$. The cross-spectrum $\tilde{Y}(k)$ is shown in Fig. 11.37(c).

Now that the FFTs have been properly zero-padded, the IFFT of $\tilde{Y}(k)$ [Fig. 11.37(c)] gives the complex cross-correlation $\tilde{y}(m)$ [Refer to Eq. (11.91)]. The magnitude and phase of $\tilde{y}(m)$ are shown in Fig. 11.38. The peak of $|\tilde{y}(m)|$ is located at $\hat{\tau} = 248.047$ μs, which is the time delay estimate of τ.

The SNR for this numerical example was 15 dB. The error in our estimate was -1.247 μs. When the SNR increases, the error decreases.

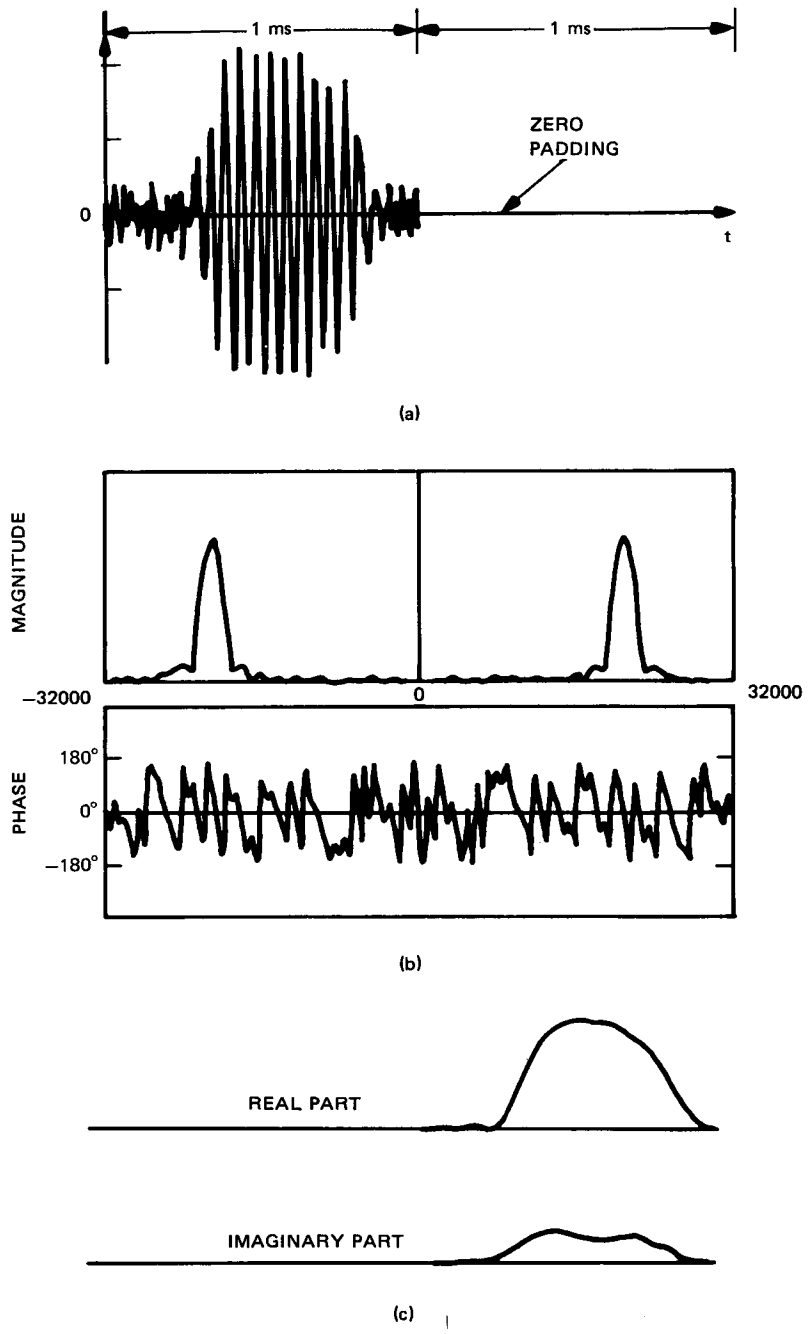


Fig. 11.36. (a) Noisy echo; zero-padding for FFT. (b) Magnitude and phase spectra for $\underline{z}(t)$. Baseband translation of $\underline{z}(t)$.

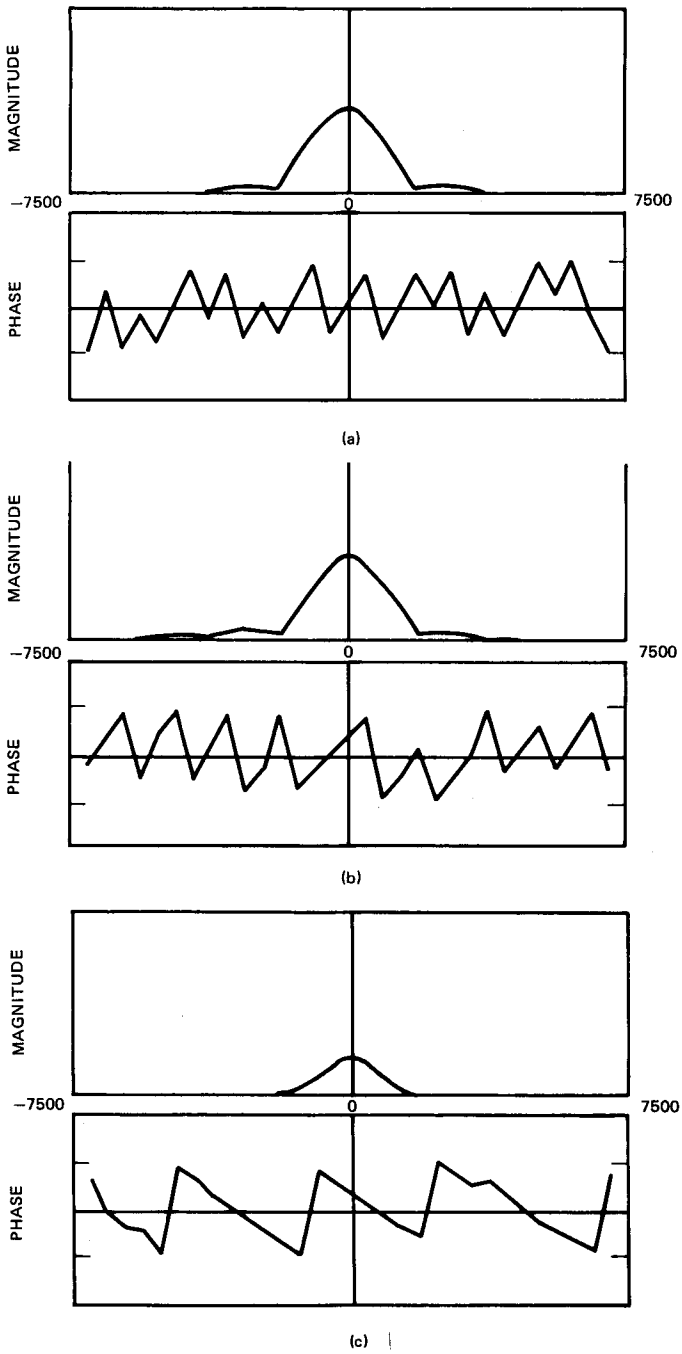


Fig. 11.37. Spectrum and cross-spectrum. (a) 1024-point FFT of $\tilde{x}_T(n)$; (b) 1024-point FFT of $\tilde{z}(n)$; (c) cross spectrum $\tilde{Y}(k)$ 0.

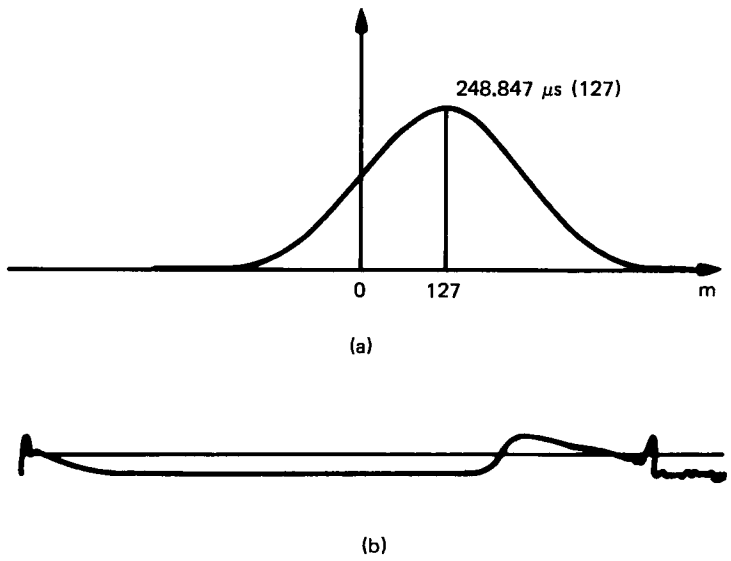


Fig. 11.38. (a) The magnitude and (b) phase angle of the complex cross-correlation function.

SUMMARY VII

In this chapter we presented an overview of the time delay estimation problem. We discussed the active and passive radar and sonar problems, beamforming, and active and passive localization. We then discussed the theory and algorithms involved in estimating the intersensor time delays.

In real-world applications we do not have stationary point target models and simple backscattering; in addition, there are multipath, finite-length observations, and nonstationary, non-Gaussian noise. However, this chapter will give the reader a good understanding of the basic principles; then the more difficult problems will be easier to analyze. For example, a good reading of this chapter will allow the reader to understand some current real-world problems, namely, the fundamental limitations of the passive theory when narrowband energy is present and the problem of tracking a time-variant time delay [31, 32].

REFERENCES

1. G. C. Carter, Time Delay Estimation for passive sonar signal processing, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29,** 463–470 (June 1981).
2. B. P. Bogert, M. J. R. Healy, and J. W. Tukey, The quefrency analysis of time series for echoes, *Proc. Symp. on Time Series Analysis*, edited by M. Rosenblatt, Wiley, New York, 1963, pp. 209–243.
3. M. M. Gibson, Delay estimation of disturbances on the basilar membrane, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29,** 621–623 (June 1981).

4. M. A. Rodriguez, R. H. Williams, and T. J. Carlow, Signal delay and waveform estimation using unwrapped phase averaging, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29,** 508–513 (June 1981).
5. D. J. Torrieri, Statistical theory of passive location systems, *IEEE Trans. Aerospace Electronic Systems* **AES-20,** 183–198 (March 1984).
6. J. P. Van Etten, LORAN C system product and development, *ITT Electrical Communication* **45,** 100–115 (1970).
7. P. M. Morse and H. Feshback, *Methods of Theoretical Physics*, Part I, McGraw-Hill, New York, 1953, pp. 838–841.
8. M. I. Skolnik, *Introduction to Radar Systems*, 2nd ed., McGraw-Hill, New York, 1980, pp. 278–337.
9. W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, Wiley, New York, 1981, pp. 537–550.
10. M. T. Silvia and A. B. Weglein, Method for obtaining a near-field inverse scattering solution to the acoustic wave equation, *J. Acoust. Soc. Am.* **69,** 478–482 (February 1981).
11. S. M. Sherman, *Monopulse Principles and Techniques*, Artech House, Dedham, Mass., 1984, pp. 1–21.
12. H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part III, Wiley, New York, 1971, pp. 167–187.
13. W. G. Neubauer, R. H. Vogt, and L. R. Dragonette, Acoustic reflection from elastic spheres. I: Steady-state signals, *J. Acoust. Soc. Am.* **55,** 1123–1129 (1974).
14. M. T. Silvia and E. A. Robinson, *Deconvolution of Geophysical Time Series in the Exploration for Oil and Natural Gas*, Elsevier, New York, 1979, pp. 1–45.
15. G. L. Turin, An introduction to matched filters, *IRE Trans. Information Theory* **IT-6,** 311–329 (June 1960).
16. C. W. Helstrom, *Statistical Theory of Signal Detection*, Pergamon Press, New York, 1968, pp. 112–115.
17. C. H. Knapp and G. C. Carter, The generalized correlation method for estimation of time delay, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-24,** 320–327 (August 1976).
18. G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*, Holden Day, San Francisco, 1968, pp. 363–421.
19. E. J. Hannan and P. J. Thomson, Estimating group delay, *Biometrika* **60,** 241–253 (1973).
20. G. C. Carter, A. H. Nuttall, and P. G. Cable, The smoothed coherence transform, *Proc. IEEE* **61,** 1497–1498 (1973).
21. J. C. Hassab and R. E. Boucher, Optimum estimation of time delay by a generalized correlator, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-27,** 373–380 (August 1979).
22. G. C. Carter, Variance bounds for passively locating an acoustic source with a symmetric line array, *J. Acoust. Soc. Amer.* **62,** 922–926 (1977).
23. C. Lawrence Ng and Y. Bar-Shalom, "Optimum multisensor, multitarget time delay estimation," *NUSC Technical Report* 65757, Naval Underwater Systems Center, New London, Conn. April 20, 1983, pp. 57–75.
24. W. R. Hahn, Optimum signal processing for passive ssonar range and bearing estimation, *J. Acoust. Soc. Amer.* 201–207 (1975).
25. *T. H. Glisson, C. I. Black, and A. P. Sage, On sonar signal analysis, IEEE Trans. Aerospace Electronic Systems* **AES-6,** 37–49 (January 1970).
26. J. V. DiFranco and W. L. Rubin, *Radar Detection*, Prentice-Hall, Englewood Cliffs, N.J., 1968, pp. 623–625.
27. J. Ville, Theorie et application de la notion de signal analytique, *Cables et Transmission* **2,** 61–74 (1948).
28. C. E. Cook and M. Bernfeld, *Radar Signals: An Introduction to Theory and Applications*, Academic Press, New York, 1967.
29. A. D. Whalen, *Detection of Signal in Noise*, Academic Press, New York, 1971, pp. 321–362.

30. A. H. Quazi, An overview on the time delay estimate in active and passive systems for target localization, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29,** 527–533 (June 1981).
31. E. Weinstein and A. J. Weiss, Fundamental limitations in passive time delay estimation. Part II: Wide-band systems, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-32,** 1064–1078 (October 1984).
32. J. O. Smith and B. Friedlander, Adaptive interpolated time delay estimation, *IEEE Trans Aerospace Electronic Systems* **AES-21,** 180–199 (March 1985).