21.2 Underwater Acoustical Signal Processing

Vyacheslav Tuzlukov, Won-Sik Yoon, and Yong Deak Kim

Background of Underwater Acoustics

Underwater acoustics entails the development and employment of acoustical methods to image underwater features, to communicate information via the oceanic wave-guide, or to measure oceanic properties. Underwater acoustics is the active or passive use of sound to study physical parameters and processes, as well as biological species and targets (for example, ships, submarines, mines, fishes, phyto- and zooplankton, etc.) at sea. In some cases, a specifically designed sound source is used to learn about the ocean and its boundaries or targets (active underwater acoustics). In other research, a natural sound or a sound generated by targets in the sea is analyzed to reveal the physical or biological characteristics of the sound source (passive underwater acoustics). Light, radar, microwaves, and other electromagnetic waves attenuate very rapidly and do not propagate any significant distance through salt water. Because sound suffers very much less attenuation than electromagnetics, it has become the preeminent tool for sensing, detection, identifying, and communicating under the ocean surface. And yet, for decades, inadequate oceanographic information about the extraordinary spatial and temporal variability of this medium has hindered underwater acousticians in their desire to predict sound propagation. It was necessary to learn more about those ocean characteristics that the traditional oceanographic instruments measure rather crudely, with great difficulty, and at great expense.

Acoustical researchers invert the problem; they use the complex nature of sound propagation to learn about the ocean. The many successes of this young science range from the identification and counting of physical and biological inhomogeneties — such as microbubbles, fragile zooplankton, fish, and mammals — to the remote measurement of distant rainfall and sea surface roughness, deep sea mountains, rocks, consolidated and suspensed sediments, as well as the shape and strength of internal waves, ocean frontal systems, and immense churning ocean eddies, hundreds of kilometers in extent. All of these unknowns can be measured by underwater acoustical techniques.

In retrospect, underwater acoustics started in 1912, when the steamship *Titanic* struck an iceberg. The subsequent loss of hundreds of lives triggered man's use of sound to sense scatterers in the sea. Within a month, two patent applications were filed by L.R. Richardson in the United Kingdom for "detecting the presence of large objects under water by means of the echo of compressional waves. . .directed in a beam. . .by a projector." The basic idea was that a precise knowledge of the speed of sound in water, and the time of travel of the sound, permits the calculation of the distance to the scatterer [1]. By 1935, acoustical sounding was used to determine the ocean depth as well as to hunt for fish schools. Much more recently it has been realized that the physical and spatial character of the scatterers can be inferred from the statistical characteristics of the scattered sound and that high-resolution images can be obtained at long range in optically opaque, turbid water.

Knowing the sound speed in water is critical to many of the applications of underwater acoustics. A value of 1435 m/sec was found, but it was soon realized that the speed in saline water is somewhat greater than this, and that in general the temperature of the water is an even more important parameter. Numerous laboratory and field measurements have now shown that the sound speed increases in a complicated way with increasing temperature, hydrostatic pressure, and the amount of dissolved salts in the water. A simplified formula for the speed in m/sec, accurate to 0.1 m/sec but good only to 1 km depth, was given by Medwin [2]:

$$c = 1449.2 + 4.6T - 0.055T^2 + 0.00029T^3 + (1.34 - 0.01T)\cdot(S - 35) + 0.016z \qquad (21.11)$$

In the above, temperature T is in degrees centigrade, salinity S is parts per thousand of dissolved weight of salts, and the depth z is in meters. The effect of salinity is quite small except near estuaries or in polar regions, where fresh water enters the sea, but microbubbles have a very large effect on the speed of propagation near the

ocean surface; frequency-dependent sound speed deviations of tens of meters per second are common in the upper ocean.

The propagation of sound in the sea has been studied intensely since the beginning of World War II when it was recognized that an understanding of this phenomenon was essential to the successful conduct of anti-submarine warfare (ASW) operations. These early measurements were quickly transformed into effective, albeit primitive, prediction tools. The study of sound propagation in the sea is fundamental to the understanding and prediction of all other underwater acoustic phenomena. Both marine seismologists and underwater acousticians have achieved advances in sound propagation, although the motivating factors have been quite different. Marine seismologists have traditionally used Earth-borne propagation or elastic waves to study the solid Earth beneath the oceans. Underwater acousticians have concentrated on the study of water-borne, compressional-wave propagation phenomena in the ocean as well as in the shallow sub-bottom layers. As research in underwater acoustics has extended to frequencies below several hundred Hertz, it has overlapped with the spectral domain of marine seismologists. Moreover, marine seismologists have become more interested in exploring the velocity–depth structure of the uppermost layers of the sea floor using higher frequencies. This area of overlapping interests has been recognized as a sub-discipline of both communities and is referred to as "ocean seismo-acoustics."

What Is Underwater Acoustical Signal Processing?

The use of acoustical signals that have propagated through water to detect, classify, and localize underwater objects is referred to as underwater acoustical signal processing or sonar. Sonar stands for "Sound Navigation Ranging" and is a method or device for detecting and locating objects, especially underwater, by means of sound waves transmitted or reflected by an object. Specific sonar application problems are submarine detection, mine hunting, torpedo homing, and bathymetric sounding. There are active and passive sonars. Active sonars transmit acoustic energy and detect targets by echolocation. Passive sonars operate by listening for acoustic emissions and can function in a covert manner.

Sonar Systems

Consider briefly the types of sonar systems. Sonar systems are used to remotely sense the interior of the ocean, what is in it, its surface, its bottom, and the structure beneath the bottom. Many specialized systems have been developed to do this. Data interpretation methods range from a simple display of "what is there and what is it doing" to statistical analysis of the pressure signals received by a system. In display and analysis there are two signal-processing classes: resolved and unresolved scatterers or reverberation. With *resolved signals*, the sources, scatterers, and so forth are separately displayed or imaged in time and space. Decades of development in sonar systems have improved the time and spatial resolution of the systems. With *unresolved signals*, the pressure signals from the sources, scatterers, and so forth are not separated in time or space and are called "reverberation." Statistical analysis, spectral analysis, and directional scattering are used in the study of unresolved scatterers in the reverberation.

Many sonar systems are almost one-task devices. The introduction of digital recording and data analysis has broadened the range of usefulness of an instrument so that a single instrument may be able to do several related tasks. Digital software has replaced many of analog operations in sonar systems, and digital signal processing has improved the adaptability of a system to new tasks. Sonar hardware and transducer configurations tend to be specialized to measurement tasks. Starting with simple sonar to echo sounder, we describe sonar configurations and their relation to remote sensing tasks. Generally, acoustic pings are used.

Echo Sounder. The most common sonar system is the echo sounder (see Figure 21.9). It employs an electrical signal generator and amplifier, called a *transmitter*, a transducer to convert an electrical signal to sound; a transducer to convert sound to an electrical signal; an electrical receiving circuit; and a display. Separate transmitting and receiving transducers are shown. A trigger from the display or transmitter starts

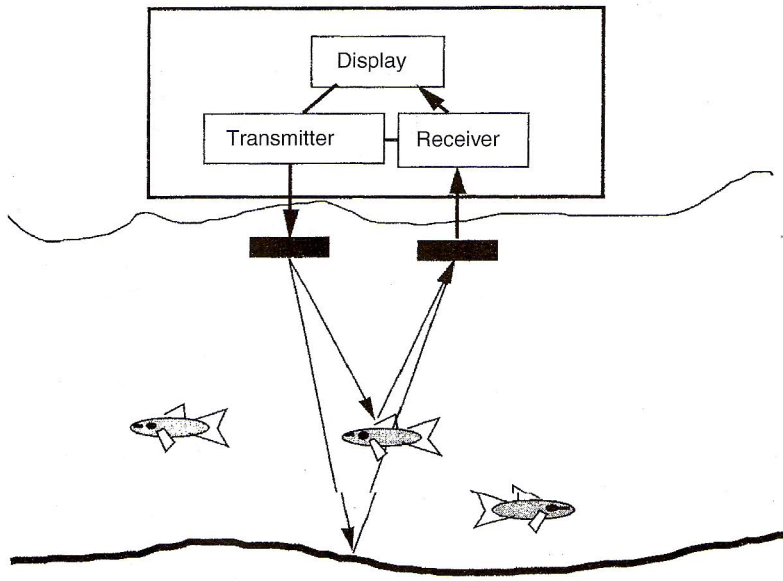

FIGURE 21.9 Echo sounder.

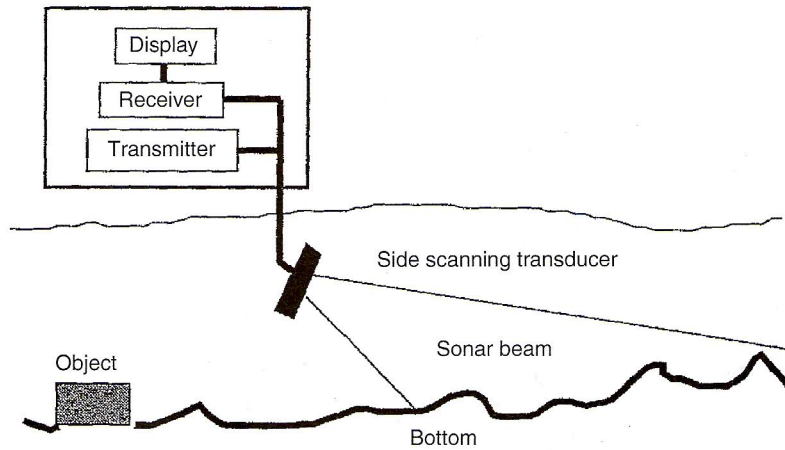

FIGURE 21.10 Side-scanning sonar.

the cycle. Many sonar systems use the same transducer for transmission and reception. Systems range in complexity from the *fish finders* that are sold in sporting goods departments to multibeam systems that are used by commercial fishermen and navies. The multibeam systems are basically combinations of many single-beam systems.

Side-Scanning Sonar. The side-scanning sonar is an echo sounder that is pointed sideways (see Figure 21.10). The sonar looks to the side of the ship and makes an echo sounding record as the ship moves. The time of return of a pulse is interpreted as the range to the bottom feature that caused the scatter. Display software converts the *raw* image to a map of features on the bottom. However, although the design concepts are the same as the simple echo sounder, the sending transducer produces a fan-shaped beam, and the receiver has a time-variable gain to compensate for range. Side-scanning sonars are used in geological studies to give images

of rough features on the sea floor. The instruments are also used to locate objects such as sunken ships on the sea floor.

Multibeam Sonar. Comparisons of mapping and object location operations that use radar in air and sonar in water demonstrate the large differences between the use of electromagnetic waves in air and sound waves in water. Radar (electromagnetic wave velocity $= 3 \times 10^8$ m/sec): the radar pulse travel time for a range of 30 km is 2×10^{-4} sec, and a simple radar systems can send, receive, and display in a very short time. The time required to make 360° image at 1° increments can be less than 0.1 sec. Thus, radar systems can use a single rotating dish to give good images. Consider the airborne radar. In 0.1 sec, an aircraft moving at a little less than the speed of sound in air (about 600 mph or 1000 kph) moves only about 30 m. The attenuation of electromagnetic waves in seawater is very large, and radar does not have a useful working range in the ocean. However, the attenuation of electromagnetic waves in glacial ice is small enough that radar soundings are used. Sonar (sound speed = 1500 m/sec): the time required for sonar to range to 30 km is 40 sec. In a sequential data acquisition system that takes one echo measurement at a time, several hours at one location would be needed to make one 360° image. A ship moving at 9 kph (2.5 m/sec) moves 100 m during the time for a single echo ranging measurement. A technological solution is to acquire sonar data in parallel by transmitting and receiving in many directions at the same time. Figure 21.11 shows an example of a multibeam sonar for sea-floor mapping. A cross-section of the ship is shown. The transmission is a broad beam. By adjusting time delays of the receiving elements, the multi-element receiving array is preformed to a set of narrow beams that look from port to starboard and measure the depths to various positions such as those shown at 1 to 7 in the figure. As the ship moves, the computer makes a contour plot of the depths. Using color coding, one gets a highly revealing picture. This system is intended to map a swath of depths along the ship track. Since these systems are usually mounted on the hull of the ship, the receiving array points in different directions as the ship

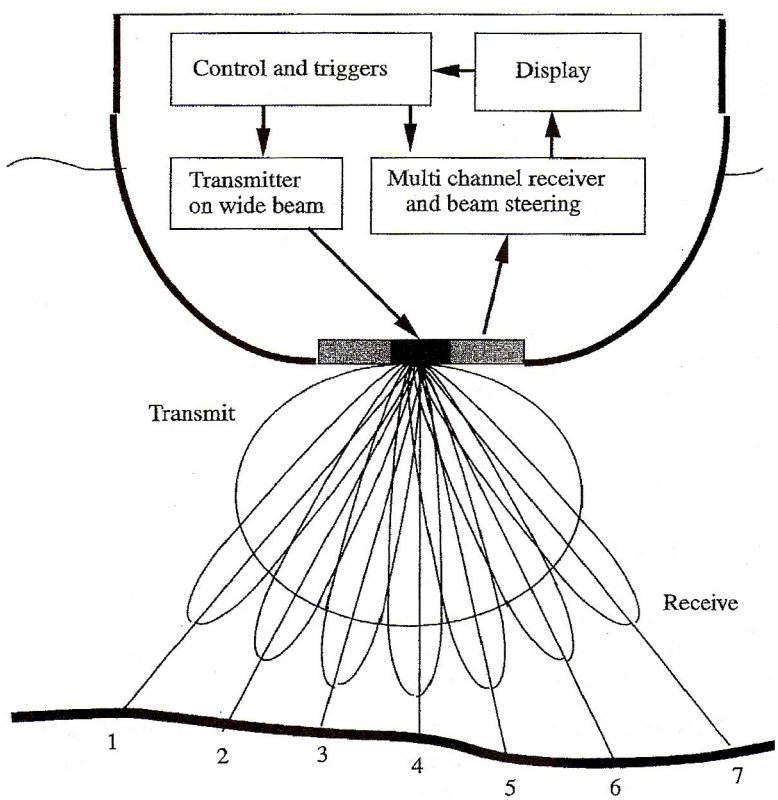

FIGURE 21.11 Multichannel sonar system using preformed beams.

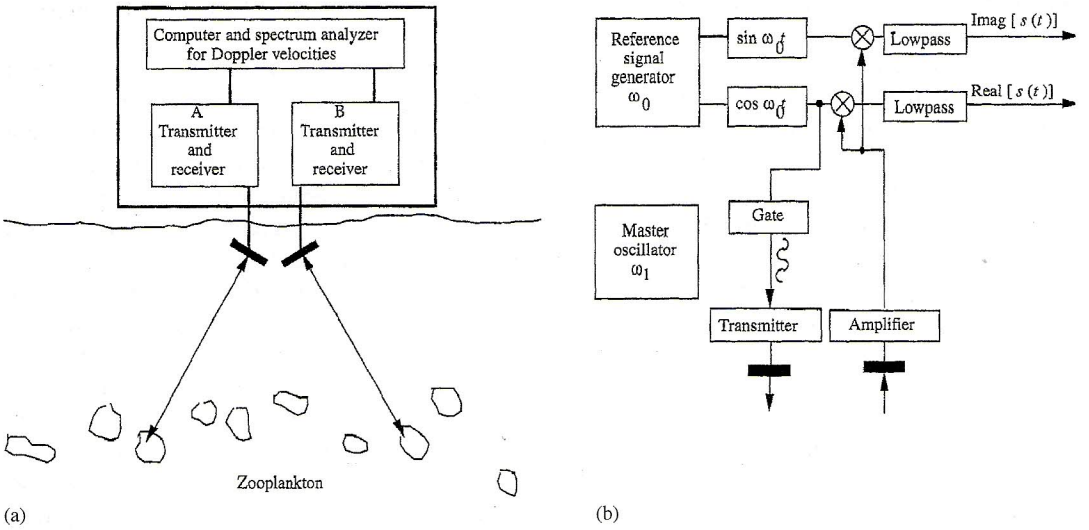

FIGURE 21.12 Doppler sonar system: (a) pings from the transmitter are backscattered from zooplankton; (b) block diagram of one channel.

pitches and rolls. The data-reduction system must compensate for the ship motions and the direction in which the receiving array is pointing when the echoes arrive.

Doppler Sonars. Doppler sonars are used to measure the velocities of ships relative to the water or the sea floor (see Figure 21.12). Commonly, Doppler sonars have four channels — two look fore and aft, and two look to starboard and port. They may also be used to measure the motion of the ocean surface or swimming objects, or internal waves, within the volume. Another application of the Doppler phenomenon is the ocean-going *portable* Doppler velocimeter [1].

Passive Sonar. Passive sonars listen to sounds in the ocean. A system may range in complexity from a single hydrophone to an elaborate, steered array of hydrophones similar to the multibeam system in Figure 21.11. There are the following noises at sea:

- Natural physical sounds: wave–turbulence interactions and oscillating bubble clouds (20 to 500 Hz); near shipping lanes the noise in the 10 to 150 Hz band is due largely to the machinery of distant ships; ocean sound on the band 500 to 20,000 Hz has been called *wind sea*, *sea state noise*, or *Knudsen noise*, because, during World War II, Vern O. Knudsen discovered that it correlated very well with wind speed [3]; the depth dependence owing to the attenuation of sea surface sound by near-surface bubble layers and bubble plumes cannot be ignored; rainfall sound.
- Natural biological sounds: noise generated by marine animals.
- Shipping noise: this noise can exhibit both spatial and temporal variabilities; the spatial variability is largely governed by the distribution of shipping routes in the oceans; the temporal variability can be introduced, for example, by the seasonal activities of fishing fleets.
- Seismoacoustic noise: microseismic band (80 mHz to 3 Hz) contains high-level microseismic noise resulting from nonlinear wave–wave interactions; noise-notch band (20 to 80 mHz) contains noise controlled by currents and turbulence in the boundary layer near the sea floor; ultra-low-frequency band (>20 mHz) contains noise resulting from surface gravity waves.

Steered Array Sonars. Transmitting or receiving arrays of transducers are steered by adding the signals from each transducer with proper time delays. The same analysis applies to send or to receive; we give the analysis for a receiving array. Consider the array of transducers in a line perpendicular to the direction $\psi = 0$ (see Figure 21.13). To steer the array, the elements at positions y_0, y_1, and so forth are given time delays τ_0, τ_1, and so forth that depend on the angle. The combination of the transducer sensitivity, analog-to-digital conversion,

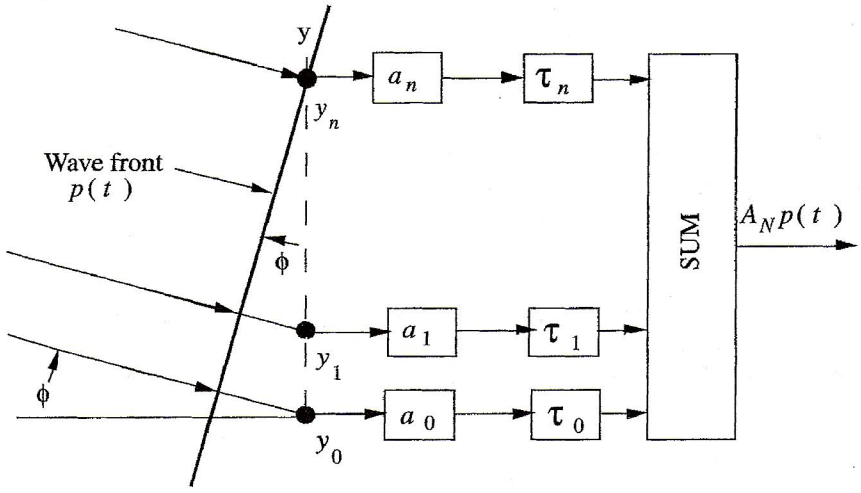

FIGURE 21.13 Electronically/digitally steered array for a plane wave entering the line of transducer array elements at angle ψ.

and amplifiers for the individual elements has the value a_n. The concept works for sources as well as receivers. To electronically/digitally steer the array, we insert appropriate time delays in each y_n-channel. Let the signal at the 0th hydrophone be $p(t)$ and the channel amplification factor be a_0. From the geometry in Figure 21.13, the plane wave front arrives at the nth hydrophone at advance Δt_n before reaching the 0th hydrophone, where

$$\Delta t_n = \frac{y_n \sin\psi}{c} \tag{21.12}$$

The signal at hydrophone n is $p(t - \Delta t_n)$. The analog/digital conversion and amplification is in the a_n. The time delay τ_n is inserted to give the signal $p(t - \Delta t_n + \tau_n)$. The sum signal for N channels is

$$A_N p_N(t) = \sum_{n=0}^{N-1} A_n p(t - \Delta t_n + \tau_n) \tag{21.13}$$

where A_n is an amplitude factor. Now, if τ_n is chosen to equal Δt_n, then the signals add in phase for that direction ψ, and we have

$$A_N p_N(t) = p(t) \sum_{n=0}^{N-1} A_n \quad \text{for } \Delta t_n = \tau_n \tag{21.14}$$

This method of array steering is called *delay and sum*. The only assumption is that the signals in each channel are the sum except for their time delays. Delay and sum processing works for any $p(t)$. The directional response of a steered array in other directions can be computed by choosing an incoming angle ψ' and letting

$$\tau_n = \frac{y_n \sin\psi'}{c} \tag{21.15}$$

Then

$$\Delta t_n - \tau_n = \frac{y_n(\sin\psi - \sin\psi')}{c} \tag{21.16}$$

The directional response of the array as a function of ψ depends on the value of ψ'. We have assumed that the incident sounds are plane waves. This is equivalent to assuming that the curvature of the wave front is small over the dimensions of the array, i.e., less than 0.125λ, where λ is the acoustic wavelength. The plane-wave assumption is effective for small arrays and distant sources. Arrays are built in many configurations: cylinders, spheres, and so on. The multibeam sonar described before is one example. By using time delays, almost any shape can be steered to receive signals of any curvature from any direction. However, when the arrays are built around a structure, diffraction effects can cause the performance to deteriorate.

Quantitative measurements of sound scattered from an object require that we understand the specifications and use of the sonar system as well as the physics of the scattering process. Sonars are often designed and adapted to the physics of a particular type of measurements. However, the operating characteristics of the sonar and the physics of the reflection and scattering processes are really independent. Now consider a generic sonar system.

Generic Sonar. The generic sonar, shown in Figure 21.14, is a combination of analog and digital components. The generic sonar system has an automatic transmit/receive switch. Some systems have separate transducers for sending and receiving. The trigger, from a clock that is internal and/or external, initiates the transmission and reception cycle. The receiver includes the electronic and digital signal processing. The transducers may be mounted on the ship or in a tow body, the *fish*. The signals may be recorded on an analog tape, a digital tape, or a compact disc. Typical displays are paper chart recorders and video display terminals. Sonars that are used for surveys and research usually include control of ping duration, choice of the time-varying-gain (TVG) function, calibration signals, displays, and analog signal outputs. TVG as a function of range for two settings for a generic sonar is shown in Figure 21.15. The operation is used to compensate for range dependence. Unresolved overlapping echoes (volume scatter) have a pressure amplitude proportional to t^{-2}, and the amplitude compensation is proportional to t and $20 \log R$, where R is a range. Isolated echoes from individuals have echo amplitudes proportional to t^{-2} and compensation is t^2 or $40 \log R$. As shown in Figure 21.15, for each transmission and reception cycle, the TVG starts at low gain and increases as a function of time. All sonar receivers have a minimum output that is related to the noise and a maximum output/limit when the amplifier overloads. Digital systems have equivalent minimum and maximum limits. The TVG is chosen to keep the output electrical signal amplitudes approximately the same for near and distant scatterers and to keep the signals above the minimum and less than the maximum limits.

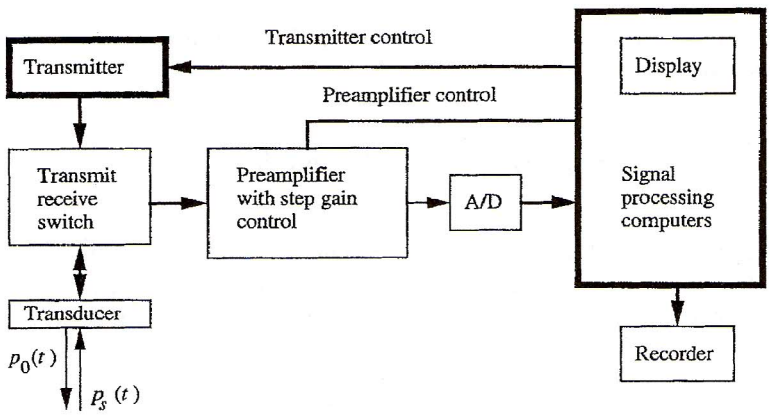

FIGURE 21.14 Block diagram of a generic sonar.

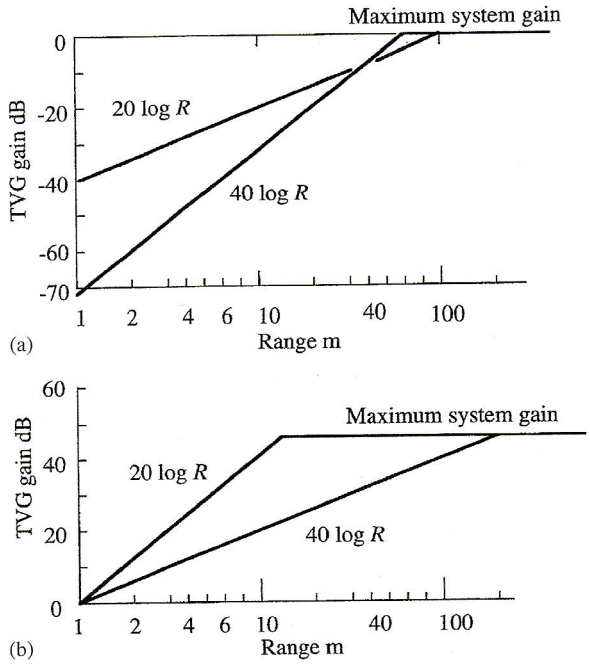

FIGURE 21.15 TVG as a function of range for two settings for a generic sonar: (a) the final TVG reference gain is unity (0 dB); (b) the initial reference gain is unity (0 dB). The maximum TVG action is limited by the maximum gain of the amplifier.

It is preferable to choose a TVG function that keeps the electrical signals in a good recording range rather than trying to match some preconceived notion of what the TVG ought to be. Standard preprogrammed choices are gain proportional to t^2 or $40 \log R$ and gain proportional to t or $20 \log R$. How the TVG operates on a voltage signal in an instrument depends on the engineering design. Some designs start with a very low gain and then increase to the amplifiers limit (see Figure 21.15a). Others start with an initial gain of unity (0 dB) and increase to the amplifier's limit (see Figure 21.15b). If one has individual echoes from many small, isolated targets, then the echo pressure amplitudes (the echoes) decrease as $\approx t^{-2}$, and a TVG of $40 \log R$ compensates for the spherical divergence.

When the objects are close together, as in a cloud of scatterers, and their echoes overlap, the sum of all of the unresolved echo pressures tends to decrease as t^{-1}, and a TVG of $20 \log R$ gives compensation. This is a characteristic of volume scatter. In acoustical surveys, both clouds of fish and individuals can be present, and neither choice of TVG fits all. It is better to use one TVG choice that keeps the voltage levels in a good range for recording. Whatever the TVG choice, an appropriate range dependence can be included in digital signal processing. Digital signal processing technology has enabled the mass manufacture of small sports-fisherman's sonars. These inexpensive instruments contain preprogrammed computers and are actually very sophisticated. The small, portable sonars can identify echoes from individual fish, display relative fish sizes, look sideways and separate echoes of large fish from reverberation, and show water depth.

Sonar with Band-Shifting or Heterodyning Operations. The ping from sonar may have a carrier frequency of 100 kHz and duration of 1 msec. Examples of pings having the same envelope and different carrier frequencies are sketched in Figure 21.16 and Figure 21.17. Since the carrier frequency f_c is known, sampling the envelope and measuring the relative phase of the carrier frequency can describe the ping. This can simplify the signal processing operations and greatly reduce memory requirements in sonar systems.

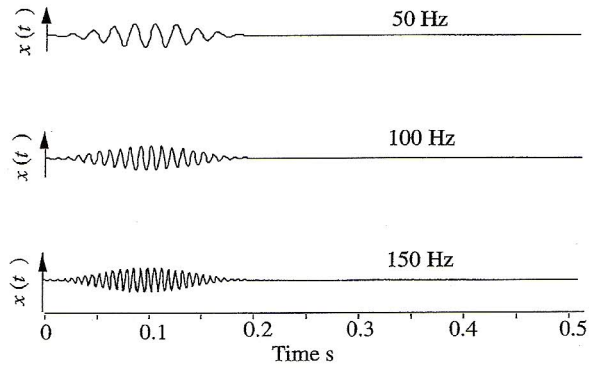

FIGURE 21.16 Signals and their spectra. All of the pings have the same duration. Time-domain presentation of pings $x(t)$ with carrier frequencies of 50, 100, and 150 Hz.

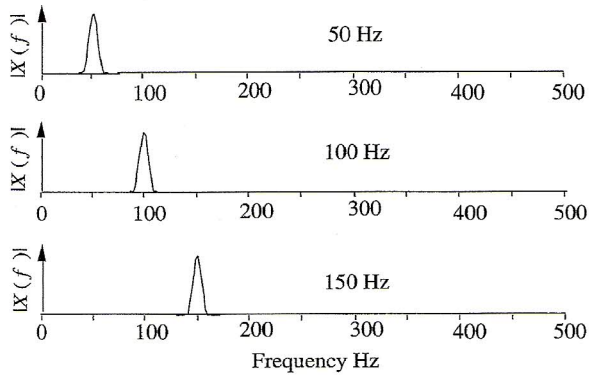

FIGURE 21.17 Signals and their spectra. All of the pings have the same duration. Modulus of the spectral amplitudes $|X(f)|$.

The ping is determined by:

$$x(t) = \begin{cases} e(t)\sin 2\pi f_c t & \text{for } 0 < t < t_p \\ 0 & \text{otherwise} \end{cases} \tag{21.17}$$

where the envelope of the ping is

$$e(t) = 0.5\left|1 - \cos\frac{2\pi t}{t_p}\right| \tag{21.18}$$

The spectrum of $x(t)$ is $X(f)$, which is sketched in Figure 21.18. The two multiplication operations, shown in Figure 21.19, are the first step:

$$x_H(t) = e(t)\sin 2\pi f_c t \cos 2\pi f_H t \tag{21.19}$$

and

$$x_Q(t) = e(t)\sin 2\pi f_c t \sin 2\pi f_H t \tag{21.20}$$

We make the approximation that the time dependence of the envelope $1 - \cos\frac{2\pi}{t_p}$ can be ignored because

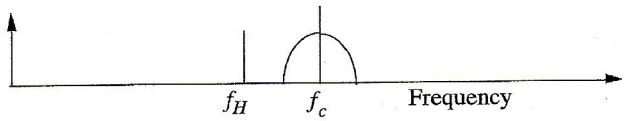

FIGURE 21.18 Band-shift or heterodyne operations: spectrum of the input signal.

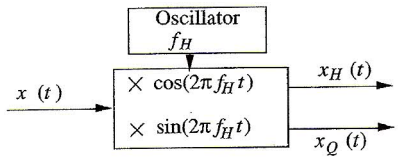

FIGURE 21.19 Band-shift or heterodyne operations: diagram of the band-shift operation for both cosine (H) and the sine quadrature (Q) multiplications.

the frequencies are very small compared with f_c. Expansion of the products of the sine and cosine terms gives

$$2 \sin 2\pi f_c t \cos 2\pi f_H t = \sin[2\pi(f_c - f_H)t] - \sin[2\pi(f_c + f_H)t] \tag{21.21}$$

and

$$2 \sin 2\pi f_c t \sin 2\pi f_H t = \cos[2\pi(f_c - f_H)t] - \cos[2\pi(f_c + f_H)t] \tag{21.22}$$

and the H and Q components of the band-shifted pings are

$$x_H(t) \approx 0.5e(t)\{\sin[2\pi(f_c - f_H)t] + \sin[2\pi(f_c + f_H)t]\} \tag{21.23}$$

and

$$x_Q(t) \approx 0.5e(t)\{\cos[2\pi(f_c - f_H)t] - \cos[2\pi(f_c + f_H)t]\} \tag{21.24}$$

The spectra of the band-shifted pings are sketched in Figure 21.20. Band-shifting operations are also known as *heterodyning*. The final steps are to low-pass-filter the results to select the $f_c - f_H$ bands. It is sufficient to sample the heterodyned signal at more than twice the frequency bandwidth of the envelope. For example, let a ping have the duration of 1 msec and a carrier frequency of 100 kHz. The bandwidth of the 1 msec ping is approximately 1 kHz, and thus the minimum sampling frequency is 2 kHz. At 4 kHz, the envelope would be sampled four times.

The relative phases of heterodyned signals are preserved in the heterodyning operation. Let the ping given by Equation (21.17) have a relative phase η:

$$x(t, \eta) = e(t) \sin(2\pi f_c t + \eta) \tag{21.25}$$

The shift of the envelope is very small. Repeating the steps that gave Equation (21.21) to Equation (21.24), the

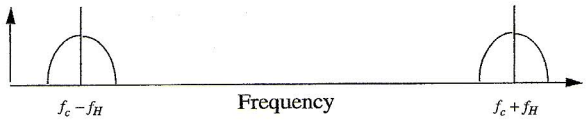

FIGURE 21.20 Band-shift or heterodyne operations: output spectra.

low-pass heterodyned signals with a phase shift are

$$x_H(t,\eta) \approx 0.5e(t)\sin(2\pi f_d t + \eta) \tag{21.26}$$

where $f_d = f_c - f_H$, and

$$x_Q(t,\eta) \approx 0.5e(t)\cos(2\pi f_d t + \eta) \tag{21.27}$$

The reference signal is $x(t)$, the phase-shifted signal is $x(t,\eta)$, and we want to measure the relative phase. The cross-correlations of the signals are less sensitive to noise than direct comparisons of the phases. Consider the cross-correlation of $x_H(t)$ and $x_Q(t,\eta)$:

$$<x_H(t)x_Q(t,\eta)> \equiv \frac{1}{\text{Norm}}\int_0^{t_p} e^2(t)\sin 2\pi f_d t\cos(2\pi f_d t + \eta)dt \tag{21.28}$$

where

$$\text{Norm} \equiv \int_0^{t_p} e^2(t)\sin^2 2\pi f_d t dt \tag{21.29}$$

The expansions of the product of the sine and cosine are

$$\sin 2\pi f_d t\cos(2\pi f_d t + \eta) = -0.5\sin\eta + 0.5\sin(4\pi f_d t + \eta) \tag{21.30}$$

and

$$\sin^2 2\pi f_d t = 0.5 - 0.5\cos 4\pi f_d t \tag{21.31}$$

The substitution of Equation (21.30) and Equation (21.31) into Equation (21.28) gives the sum of two integrals. The integral that includes the $\sin(4\pi f_d t + \eta)$ term tends to 0. The remaining integral is

$$<x_H(t)x_Q(t,\eta)> \approx -\frac{1}{2\text{Norm}}\int_0^{t_p} e^2(t)\sin\eta dt \tag{21.32}$$

The evaluation of *Norm*, using Equation (21.30) and Equation (21.31), reduces Equation (21.32) to

$$<x_H(t)x_Q(t,\eta)> \approx -\sin\eta \tag{21.33}$$

The other cross-correlations are determined as follows:

$$<x_Q(t)x_H(t,\eta)> \approx \sin\eta \tag{21.34}$$

$$<x_H(t)x_H(t,\eta)> \approx \cos\eta \tag{21.35}$$

$$<x_Q(t)x_Q(t,\eta)> \approx \cos\eta \tag{21.36}$$

The following expression gives the tg η:

$$\mathrm{tg}\,\eta = -\frac{<x_H(t)x_Q(t,\eta)> - <x_Q(t)x_H(t,\eta)>}{<x_H(t)x_H(t,\eta)> + <x_Q(t)x_Q(t,\eta)>} \tag{21.37}$$

The cross-correlation method of measuring the phase difference is useful when two pressure signals (pings) are received on a pair of transducers and their phase difference is used to calculate the direction of the incident pressure signal [4].

Echo Identification Rules. The echo amplitude display from echo sounders shows an interesting phenomenon as the instrument moves slowly over isolated fish. Normally, TVG compensates for range so that the echo amplitudes are the same for the same sizes of fish at different ranges. Figure 21.21 shows the details of echoes from one fish in a sonar beam. First for simplicity of drawing, the sonar is fixed and the fish swims through the center of the sonar beam. Pings measure the sequence of ranges to the fish at positions "a" to "i." The graphic recorder plots the echoes beneath the time of the ping and the sequence of echoes form a crescent. The echo amplitudes are given by the position of the fish in the transducer-response pattern. The graphic record of echoes would be the same if the fish was fixed and the sonar transducer was moved from left to right ("a" to "i") over the fish. An echo crescent is formed as the scattering object (fish) moves through the echo sounder's beam pattern. Starting with the object at the left edge of the beam, the echo is weak, and, from the geometry, the range is a little larger than at position "e." As the object moves into the center of the beam ("e"), the echo is larger, and the range is smaller. The width of the crescent and its amplitude depend on depth and whether the object goes through the center of the beam or is off to one side. The procedures to use these effects in the analysis of echoes from fish are suggested in [5]. The computer identification of an echo requires acceptance rules. The envelopes of a single echo, reverberations, and several echoes are shown in Figure 21.22. The sonar return has been compensated for spherical spreading. The echo envelope is $e(t)$. A single echo has shape parameters. The ping has a duration of t_p, and ideally the duration of the echo is t_p. A specific example

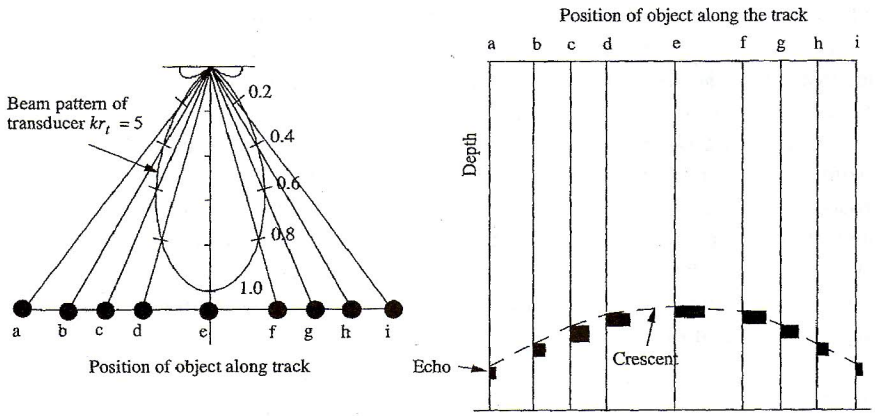

FIGURE 21.21 Simulation of a graphic display of echoes from a single fish.

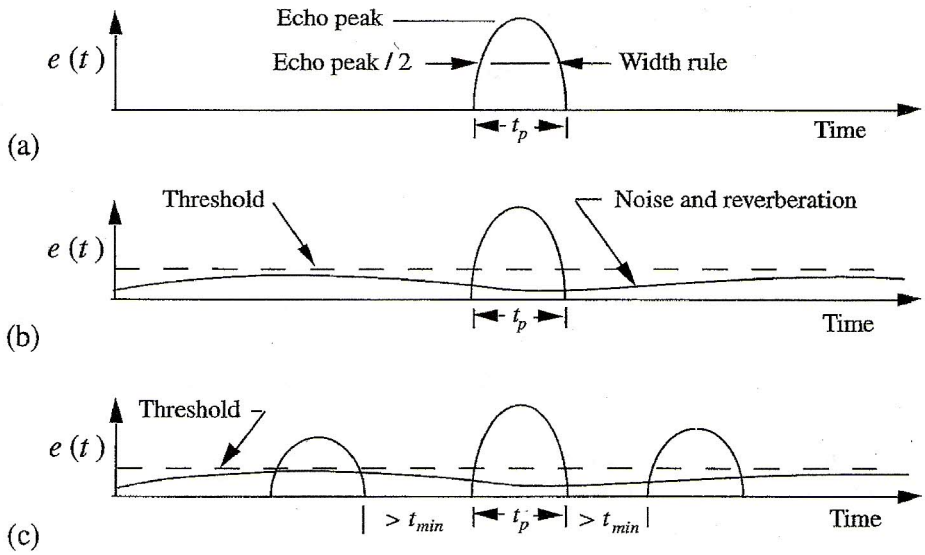

FIGURE 21.22 Identification of echoes: (a) at half-echo amplitude, the width must be less than 1.5 t_p and greater than 0.5 t_p; (b) the echo amplitude must be greater than a threshold and the reverberation noise; (c) the echo is less than the threshold before and after the echo and the minimum time t_{min} is greater than the reciprocal frequency bandwidth of the receiver.

of acceptance rules follows. Single echo identification and acceptance rules are important parts of the signal processing codes that are used in data analysis. Acceptance rules are used in the fish-finding sonars for sportsmen.

Signal Processing

Signals are the messages that we want to receive at our hydrophone. Noises are everything that we do not want to receive. The types of signal messages include impulses and continuous wave (CW) tones of short or long duration and constant or varying frequency; they also include complicated coded messages and random sequences. The form of noise can run the gamut. We all know the popular saying "Beauty (or ugliness) is in the eye of the beholder." One can propose a comparable acoustical maxim: "Signals (or noise) are in the ear of the listener." There are many examples: the sonar operator searching for the *signal* of a submarine will call the sounds of whales and dolphins *noise*. Needless to say, the marine mammals seeking to communicate or locate food would characterize man-made sounds as *noise*. Some sounds called *noise* for many years are now recognized as bearing information that qualifies them as *signals*; for example, the sound of rainfall at sea is now used to measure the size and number of raindrops per square meter per second. Flow noise at a transducer, electrical circuit noises, and the 60- or 50-Hz electrical interferences from power lines are generally regarded as noise by everybody.

Traditionally, underwater ambient noise has been specified in terms of the sound measured at a convenient hydrophone, some distance from the sources. The origins of the sound are often a mystery. We initiate a different approach. We will survey the acoustic power, source pressure, directionality, and intermittency of physical, biological, and man-made ocean sounds at their source. When this information is known, one's knowledge of propagation in the ocean allows us to calculate the ambient sound at any location. In addition to our survey of many of the more common sound sources at sea, we need operations that allow the listener to sort out the signal from the noise

Sampling Rules

Practically all underwater sound signals and noise are recorded digitally, and the results of analysis are displayed on computer terminals. The acoustician uses signal acquisition and digitizing equipment, signal

processing algorithms, and graphic display software to make this happen. The noisy signal that comes from our hydrophone is an electrical voltage that is a continuous function of time. It is called an analog signal. Hydrophone signals must be sampled to convert them from analog to digital format in order to enter a digital computer. We must sample signals properly or we get garbage. The sampling rules are general. They apply to either temporal or spatial sampling of the oceanic environment. When the rules are obeyed, the original signal can be recovered from the sampled signal with the aid of an interpolation procedure. If the rules are not obeyed, and sampling is too sparse, the original signal cannot be recovered. The Nyquist sampling rules are [6]:

- *Space-Domain Rule.* The spatial sampling interval must be less than half of the shortest wavelength of the spatial variation. Spatial sampling is sometimes described in terms of the spatial wave number, $k_s = \frac{2\pi}{\lambda_s}$, where λ_s is the distance between samples.
- *Time-Domain Rule.* The time interval between samples must be less than half of the shortest period in the signal. Sampling is defined in terms of the sampling frequency $f_s = t_0^{-1}$, where t_0 is the time between samples. Otherwise stated, the sampling frequency must be greater than twice the highest frequency component in the signal.

Spatial sampling. In traditional marine biology, samples are taken by towing a net through the water. A net is lowered to the depth, opened, and towed for a specified distance. Net sampling takes a lot of work.

Temporal sampling. The electrical signal $x(t)$ is sampled by an analog-to-digital converter to create a sequence of numbers (see Figure 21.23). The clock (Figure 21.23(a)) gives the sampling instruction. The sample is the

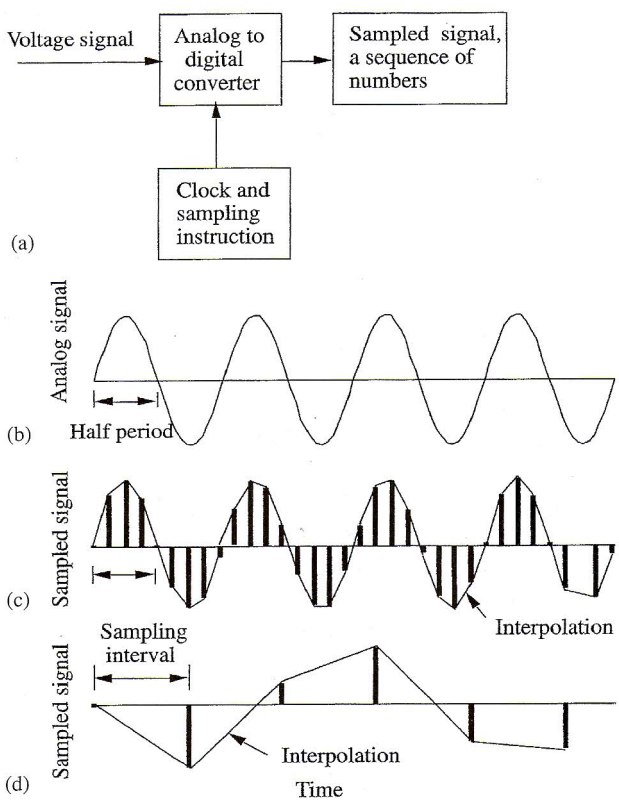

FIGURE 21.23 Temporal sampling of a simple signal: (a) sampling system to change an analog voltage into a sequence of numbers; (b) input analog sinusoidal voltage; (c) the result of sampling four times during each half-period (the vertical lines represent the magnitudes of the sampled voltages and the straight lines between ends of the vertical lines are interpolations); (d) the result of sampling at times greater than the half-period.

instantaneous value of the signal voltage at the clock time. No information is recordered about the signal voltage between samples where straight lines are drawn. Examples of data taken at two different sampling intervals are shown. In Figure 21.23(c) there are four samples in a half-period. In Figure 21.23(d) the sampling interval is larger than the half-period. Reconstruction of the inadequately sampled signal in Figure 21. 23(d) does not resemble the original signal, whereas Figure 21.23(c) does. A practical rule of thumb is to sample at intervals less than the period/3 for approximate reconstruction of the original signal.

Filter Operations

Electrical filters were originally introduced into electronic systems by radio and telephone engineers to separate the signals they wanted from those that they did not. Instrumentation manufacturers built analog filters or *black boxes* for research laboratories. These filters had switches on the front to make frequency bandpass choices. We use the filters that are built into our radio or television set when we select a channel that tunes in our desired station and rejects others. Audio amplifiers have filters (equalizers, bass, and treble controls) to modify the amplitudes of the input frequencies and to enhance the quality of the sound coming from the speakers. Digital communication engineers have developed the digital equivalent of the black box analog filter. An analog-to-digital converter digitizes the incoming analog electrical signal, and a computer does the filter operations. In many systems, the filtered digital sequence of numbers is converted back to an analog signal for listening and display. For simplicity in our discussion of operations on signals and noise, we use the neutral symbols x, h, and y to represent source, filters, and filter outputs, respectively.

Finite Fourier Transformations. Digital computers, digitized data, and efficient algorithms have made the numerical computations of Fourier transformations practical. The Fourier transformation of a finite number of data points is called the finite Fourier transformation (FFT) or discrete Fourier transformation (DFT). The time intervals between all digitally sampled data points are t_0. For N data points, the Fourier transformation pairs are

$$X_{\text{fft}}(m) = \sum_{n=0}^{N-1} x(n) \cdot e^{-2j\frac{\pi mn}{N}} \tag{21.38}$$

$$x(n) = \frac{1}{N} \sum_{m=0}^{N-1} X_{\text{fft}}(m) \cdot e^{2j\frac{\pi mn}{N}} \tag{21.39}$$

where $x(n)$ is the nth digital input signal amplitudes, and $X_{\text{fft}}(m)$ is the mth spectral component.

Equation (21.38) changes a time-dependent series of terms into a frequency spectrum. The companion Equation (21.39) or inverse finite Fourier transformation or (IFFT), changes a spectrum into a time-dependent expression. Real $x(n)$ transforms to complex $X_{\text{fft}}(m)$ and vice versa. A simple example demonstrates the periodic properties. Let $N = 64$ and suppose that the original digitized signal is real and exists between $n = 0$ and $n = 63$ as shown in Figure 21.24(a). The digitized signal is assumed to be an isolated event (see Figure 21.24(a)). The FFT method assumes that the signal has the period $N = 64$, as shown in Figure 21.24(b). Evaluations of the FFT give real and imaginary components of $X_{\text{fft}}(m)$. Inspection of Figure 21.24(c) shows that the real components are symmetric about 0, 0.5N, and N. In Figure 21.24(d) the imaginary components are antisymmetric about 0, 0.5N, and N. The modulus $|X_{\text{fft}}(m)|$, Figure 21.24(e), is

$$|X_{\text{fft}}(m)| = \sqrt{\{\text{Re}X_{\text{fft}}(m)\}^2 + \{\text{Im}X_{\text{fft}}(m)\}^2} \tag{21.40}$$

which is symmetric about 0, 0.5N, and N. If we had started with $\text{Re}X_{\text{fft}}(m)$ and $\text{Im}X_{\text{fft}}(m)$, we would have the periodic $x(n)$. Figure 21.24 shows most of the important properties of the FFT.

Filter-Response Measurements. As shown in Figure 21.25, the frequency response of the filter is the ratio of the output/input voltages for a long-duration sinusoidal input signal (the oscillator). Two measurements are sketched in Figure 21.25. To record signals digitally, we need a low-pass filter to prepare the signal for the digitization operation. The filter shown in Figure 21.25(b) is a *low-pass, antialiasing filter.* It is adjusted to pass

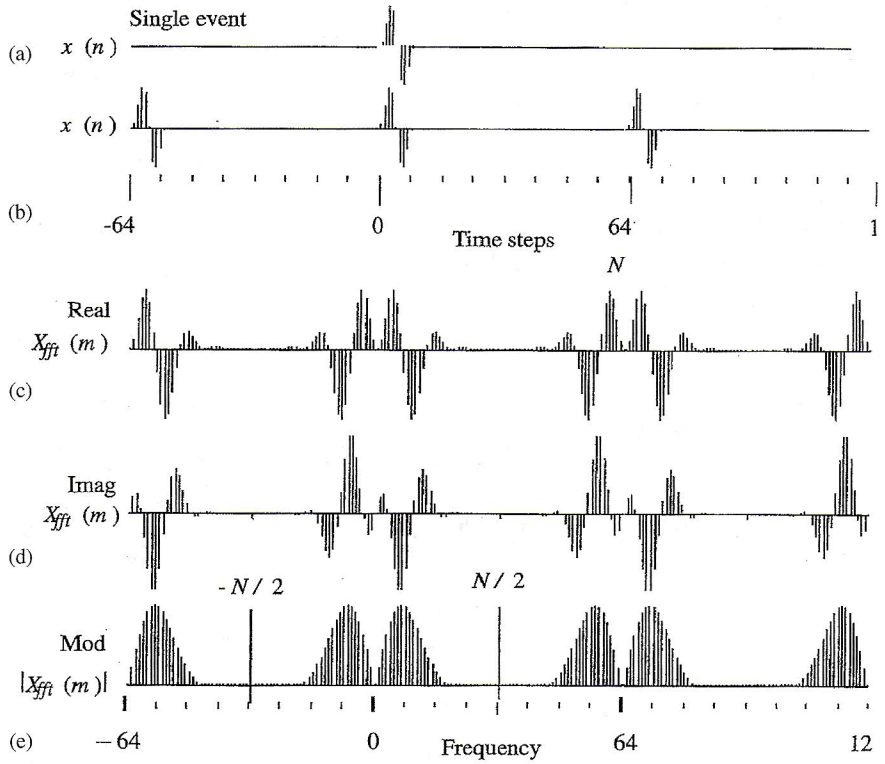

FIGURE 21.24 Signal, periodic Fourier series and its spectrum. Even if the orthogonal signal is not periodic, expansion in the Fourier series creates a new signal that is periodic. In applications, many zeros are appended to the signal to move the next cycle out of the way: (a) original signal; (b) periodic signal; (c) real component of $X_{fft}(m)$; (d) imaginary component of $X_{fft}(m)$; and (e) modulus of $X_{fft}(m)$,$|X_{fft}(m)|$.

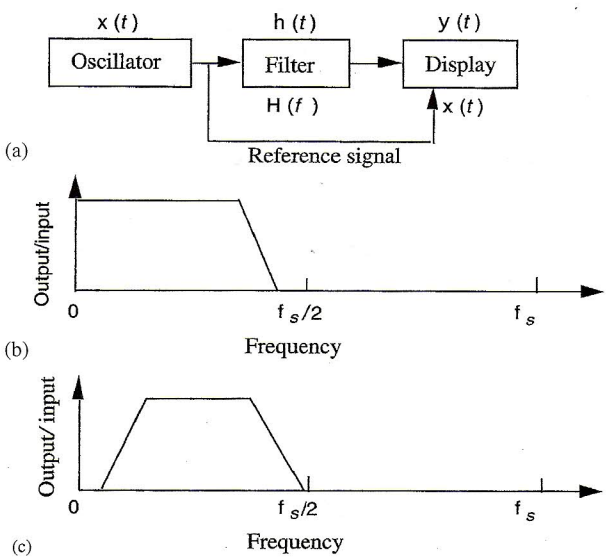

FIGURE 21.25 Filters and their responses: (a) block diagram for a typical filter response measurement; (b) response of an antialiasing, low-pass filter that is used ahead of analog-to-digital conversion at sampling frequency f_s; (c) response of a bandpass filter.

frequencies that are less than half the sampling frequency of the analog-to-digital converter and to reject higher frequencies. It thereby can prevent higher-frequency components from appearing as *alias* signals. The action of a bandpass filter is sketched in Figure 21.25(c). It is used to pass a signal and to reject unwanted signals and noise within the designed frequency range.

Time-Domain View of Bandpass Filtering. Figure 21.26(a) shows a short-duration signal, $x_1(t)$, which then passes through an appropriate bandpass filter to give the output $y_1(t)$ in Figure 21.26(b). The high- and low-pass settings of the filter were chosen to pass the signal with the signal with an acceptable amount of distortion of waveform. A longer-duration, low-frequency whale song, $x_2(t)$, is emitted during this same time so that the sum of the two signals at the input (see Figure 21.26(c)) is

$$x(t) = x_1(t) + x_2(t) \tag{21.41}$$

The output of the bandpass filter $y(t)$ is shown in Figure 21.26(d). The bandpass filter effectively removes the interfering whale song and reveals the short-duration 150-Hz ping.

Filter Operations in the Frequency Domain. Operations in the frequency domain are intuitively simple. The frequency-dependent functions, $X(f)$ and $Y(f)$, are the amplitude spectral densities of the input and output signals, and $H(f)$ is the filter-frequency response. The input–output expressions for analog signals are

$$Y(f) = H(f) \cdot X(f) \tag{21.42}$$

and

$$Y_{\text{fft}}(f) = H_{\text{fft}}(f) \cdot X_{\text{fft}}(f) \tag{21.43}$$

The output signals $Y(f)$ or $Y_{\text{fft}}(f)$ have the frequency components that are passed by the filters. Our discussion uses the amplitude spectral densities for brevity.

Figure 21.27 represents a spectrum analyzer that was constructed of many bandpass filters. The complex signal is an input to each of the filters. The spectral output of the jth filter is

$$Y_j(f) = H_j(f) \cdot X_j(f) \tag{21.44}$$

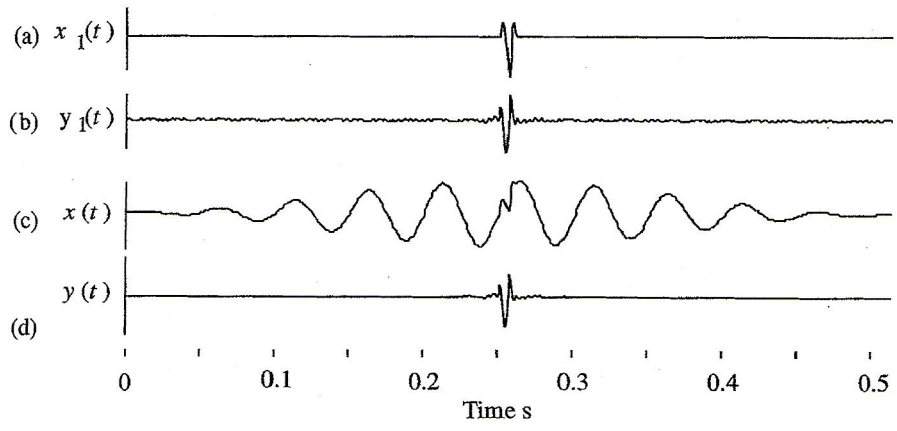

FIGURE 21.26 Filter operation shown in the time-domain: (a) signal input is a 150-Hz ping having a duration of 0.01sec; (b) signal out of a 50 to 150 Hz bandpass filter; (c) input 150-Hz ping and a 20-Hz whale song; (d) filtered signal output using the 50 to 150 Hz bandpass filter.

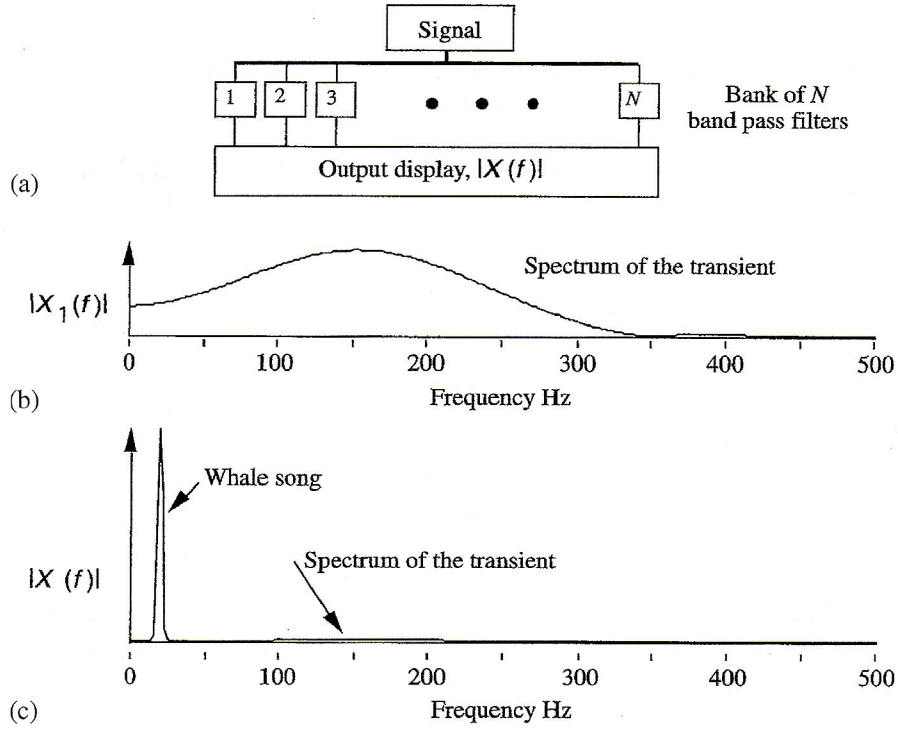

FIGURE 21.27 Spectral analysis of the signals of Figure 21.26 using a digital spectral analysis. The bandwidths of the equivalent bandpass filters are 2 Hz: (a) an analog spectrum analyzer that uses a bank of bandpass filters; (b) the digitally calculated spectrum of the 150-Hz ping in Figure 21.26(a); (c) the digitally calculated spectrum of the ping and whale song in Figure 21.26(c). The spectral amplitude factor of the ping is 0.025 that of the whale song. The ping does not show the detail of Figure 21.26(c) because of the change of scale.

Parseval's theorem gives the equivalence between the integral squares of signals in the time- and frequency-domains:

$$\int_{-\infty}^{\infty} y^2(t)\mathrm{d}t = \int_{-\infty}^{\infty} |Y(f)|^2 \mathrm{d}f = 2\int_{0}^{\infty} |Y(f)|^2 \mathrm{d}f \tag{21.45}$$

where the time integral of $y^2(t)$ is finite, and the absolute squares of $|Y(f)|$ and $|Y(-f)|$ are equal. For signals that start at 0 time, the limits of the doubly infinite integral become 0 to ∞. Using the filter input and output, Equation (21.44), and Parseval's theorem [7], the integral square output of the jth filter is

$$\int_{0}^{\infty} y_j^2(t)\mathrm{d}t = 2\int_{0}^{\infty} |Y_j(f)|^2 \mathrm{d}f \tag{21.46}$$

The substitution of Equation (21.44) into Equation (21.46) gives the filter output:

$$\int_{0}^{\infty} y_j^2(t)\mathrm{d}t = 2\int_{0}^{\infty} |X_j(f)\cdot H_j(f)|^2 \mathrm{d}f \tag{21.47}$$

For an approximation, let $H_j(f)$ be a *boxcar* filter defined by

$$H_j(f) = \begin{cases} 1 & \text{for} \quad f_j - 0.5\Delta f \leqslant f \leqslant f_j + 0.5\Delta f \\ 0 & \text{otherwise} \end{cases} \tag{21.48}$$

If $X(f)$ is approximately constant in the pass band (Equation (21.48)), then the time integral square (Equation (21.47)) is approximately:

$$\int_0^\infty y_j^2(t)\mathrm{d}t \approx 2|X(f_j)|^2 \Delta f \tag{21.49}$$

This is the raw output of the boxcar spectrum analyzer. The spectral density in a 1-Hz bandwidth is obtained by dividing both sides of Equation (21.49) by Δf:

$$\frac{1}{\Delta f}\int_0^\infty y_j^2(t)\mathrm{d}t \approx 2|X(f_j)|^2 \tag{21.50}$$

The quantity in Equation (21.50) is sometimes called an *energy spectral density*, $E_{xx}(f_j)$, when the signal amplitude is a voltage, because the *integral square* over time is proportional to the energy of the electrical signal during that time:

$$E_{xx}(f_j) = \frac{1}{\Delta f}\int_0^\infty y_j^2(t)\mathrm{d}t \approx 2|X(f_j)|^2 \tag{21.51}$$

The subscript $_{xx}$ indicates that $x(t)$ is the function being analyzed.

The $x(t)$ and $y(t)$ have units of Pa, the units of $E_{xx}(f_j)$ are Pa^2 sec/Hz. Although the quantity $E_{xx}(f)$ is often called an *energy spectral density*, it is actually proportional to the energy spectral density of the wave. As shown in the discussion of intensity [1], true expressions for energy spectral density (joules/m^2 Hz) require that Pa^2 sec/Hz be divided by $\rho_A c$, where ρ_A is the ambient density of the medium and c is the sound speed or velocity. Most spectrum analyzers are digital and use computers to do the spectral analysis. The digital spectrum analyzers often can have the equivalent of more than 1000 very narrow band-pass filters. Examples of spectrum analysis are shown in Figure 21.27(b) and (c). The spectrum of the 150-Hz ping (see Figure 21.27(b)) and the spectrum of the short-duration ping and longer-duration whale song are shown in Figure 21.27(c).

Gated Signals

The spectrum of a signal depends on its time-domain waveform. Consider some pings and their spectra. These comparisons display the relation of periodicity and duration, in the time-domain, to the peak frequency and bandwidth of the spectrum. For these examples, the ping has a slow turn-on and turn-off. The signal $x(t)$ is

$$x(t) = \begin{cases} 0.5\left[1 - \cos\dfrac{2\pi t}{t_p}\right]\sin 2\pi f_c t & \text{for} \quad 0 < t < t_p \\ 0 & \text{otherwise} \end{cases} \tag{21.52}$$

where t_p is the total nonzero ping duration and f_c is the carrier frequency. The amplitude factor in the brackets gives a spectrum with very small side lobes. This signal is similar to the sound-pressure signal radiated by many sonar transducers and some marine animals. The envelope of the sine wave is tapered from zero to a maximum and then back to zero.

Dependence of Spectrum on Ping Carrier Periodicity. Figure 21.16 and Figure 21.17 show the dependence of the peak of the spectrum on the carrier frequency of the pings. The durations of the pings were chosen to be long so that the widths of spectral peaks are narrow. The 50-Hz signal has a spectral peak at approximately 50 Hz. The other signals have their spectral peaks at 100 and 150 Hz. The period of the signal can be measured to estimate the peak or central frequency of the spectrum. This is the first rule of thumb.

Dependence of Spectrum on Ping Duration. Figure 21.28 shows the dependence of the widths of the spectral peaks on the durations of the pings. The effective durations of the signals t_d are a little less than the t_p in Equation (21.52) because the turn-ons and turn-offs are very gradual. The same frequency, 100 Hz, was used for all examples. To define the bandwidth Δf, we use the half-power width given by the two frequencies where the amplitude is 0.707 of the peak amplitude. The spectra shown in Figure 21.28(b) are the module of the absolute amplitudes. The widths of the spectra decrease as the signal duration increases and are approximately the reciprocals of the durations of the signals. The comparisons are in Table 21.1. These comparisons give a second rule of thumb:

$$\Delta f \cdot t_d \geqslant 1 \tag{21.53}$$

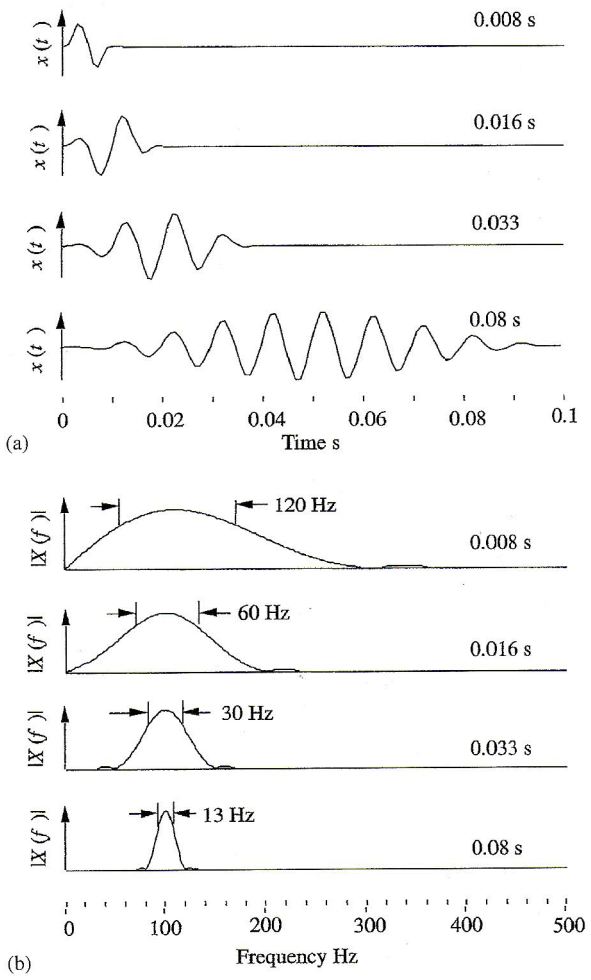

FIGURE 21.28 Signals having the same carrier frequency and different durations: (a) signals in the time-domain; (b) module of spectral amplitudes in the frequency-domain. The bandwidths were measured at the half-power points (i.e., at 0.707 of the peak amplitudes).

TABLE 21.1 Comparison between the Signal Duration and Spectrum Width

t_p, sec	t_d, sec	$\frac{1}{t_d}$, Hz	Δf, Hz
0.01	0.0085	118.0	120
0.02	0.017	58.8	60
0.04	0.034	29.4	30
0.10	0.08	12.5	13

We use the $\geqslant$ sign for $\Delta f \cdot t_d$ because many signals have durations greater than Δf^{-1}. The time t_d gives the minimum duration of a signal for a sonar system to have a bandwidth Δf.

Power Spectra of Random Signals

Sound pressures that have random characteristics are often called noise, whether they are cleverly created as such or are the result of random and uncontrolled processes in the ocean. The spectral analysis of both is the same. As inputs to a spectrum analyzer, they are signals, and their spectral descriptions are to be determined. For a short name, these are called *random signals.*

Signals Having Random Characteristics. In their simplest form, signals that have random characteristics are the result of some process that is not predictable. In honest games, the toss of a coin and the roll of a die give sequences of random events. In the Earth, processes that range from the occurrence and location of earthquakes to rainfall at sea are generators of random signals. For simulations and laboratory tests, we use computers and function generators to make sequences or sets of random numbers. Many of the algorithms generate sequences that repeat, and these algorithms are known as *pseudorandom number generators.* The numerical recipe books give random number-generating algorithms. Programming languages usually include a function call such as *rng* (.) in the library of functions.

Spectral Density and Correlation Methods. The correlation or covariance method of analyzing random signals is discussed in detail by Blackman and Tukey in their monograph *The Measurement of Power Spectra* [1958]. The random signal is the sequence of numbers $x(n)$, and the sequence has $N + k_{max}$ numbers. The covariance of the random signal is the summation:

$$c_{xx}(k) = \begin{cases} \frac{1}{N}\sum_{n=0}^{N-1} x(n)\cdot x(n+k); \\ 0 \qquad\qquad \text{otherwise} \end{cases} \tag{21.54}$$

The covariance $c_{xx}(k)$ is symmetric, and $c_{xx}(k) = c_{xx}(-k)$. The Fourier transformation of $c_{xx}(k)$ is, using Equation (21.38):

$$C_{xx,\text{fft}}(m) = \sum_{k=0}^{N-1} c_{xx}(k) \cdot e^{-2j\frac{\pi k m}{N}} \tag{21.55}$$

The substitution of Equation (21.54) in Equation (21.55) gives:

$$C_{\text{fft}}(m) = \frac{1}{N}\sum_{k=0}^{N-1}\sum_{n=0}^{N-1} x(n) \cdot x(n+k) \cdot e^{-2j\frac{\pi m k}{N}} \tag{21.56}$$

Change variables by letting $j = n + k$, and Equation (21.56) becomes:

$$C_{\text{fft}}(m) = \frac{1}{N}\sum_{n=0}^{N-1} x(n) \cdot e^{2j\frac{\pi m n}{N}} \sum_{n=0}^{N-1} x(k) \cdot e^{-2j\frac{\pi m k}{N}} \tag{21.57}$$

The first summation is the complex conjugate $X^*_{\text{fft}}(m)$ and the second summation is $X_{\text{fft}}(m)$. The spectrum is

$$C_{xx,\text{fft}}(m) = \frac{X^*_{\text{fft}}(m) \cdot X_{\text{fft}}(m)}{N} \tag{21.58}$$

and, using Equation 6.2.27 in [1], the spectral density is

$$C_{xx}(m) = C_{xx,\text{fft}}(m) \cdot t_0 \tag{21.59}$$

where t_0 is the time between samples. Frequency domain expressions for the autocovariance are

$$c_{xx}(\tau) = \int_{-\infty}^{\infty} C_{xx}(f) \cdot \mathrm{e}^{2\mathrm{j}\pi f \tau} \mathrm{d}f \tag{21.60}$$

and

$$C_{xx}(f) = \int_{-\infty}^{\infty} c_{xx}(\tau) \cdot \mathrm{e}^{-2\mathrm{j}\pi f \tau} \mathrm{d}\tau \tag{21.61}$$

where $C_{xx}(f)$ has positive and negative frequencies, and $C_{xx}(f) = C_{xx}(-f)$, or

$$c_{xx}(\tau) = 2\int_{0}^{\infty} C_{xx}(f) \cos(2\pi f \tau) \mathrm{d}f \tag{21.62}$$

This pair of transformations, Equation (21.61) and Equation (21.62), is known as the Wiener–Khinchine theorem. The *power spectral density* of $x(t)$ is the sum of the positive and negative frequency components:

$$\Pi_{xx}(f) = 2C_{xx}(f) \tag{21.63}$$

Random Signal Simulations: Intensity Spectral Density

In the simulation of a random signal, the random function generator gives a sequence of random numbers: $x(0), x(1)$, and so on. Figure 21.29(a) shows a sequence where the numbers have been connected by interpolation lines. The result of bandpass filtering the input random signal gives a new random signal, the output $y(n)$ in Figure 21.29(b). The operations of bandpass filtering, squaring the signal, and summing or integrating the squared signal are indicated in Figure 21.30. The effective number of independent trials is

$$N_{\text{it}} = t_{\text{d}} \cdot \Delta f_m \tag{21.64}$$

where t_{d} is the duration of the signal and Δf_m is the filter bandwidth:

$$\Pi_{xx}(f_m) = \frac{1}{N\Delta f_m} \sum_{n=0}^{N-1} y_m^2(n) \tag{21.65}$$

where N is the number of samples.

The filtered signal is $y_m(n)$, where the subscript is added to indicate the filtering by the mth filter. The signal is squared, summed, and averaged over t_{d} to give the power. Since the mean square output (e.g., volt^2) is

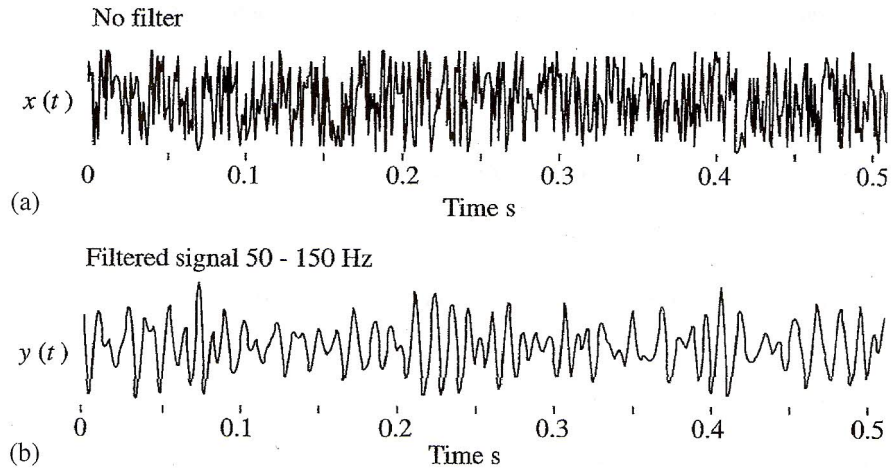

FIGURE 21.29 (a) Random signal created by using a random number generator (the sampling interval was 1 msec). (b) Result of bandpass filtering the signal in (a) through a 50- to 150-Hz bandpass filter and the duration of the signal is t_d.

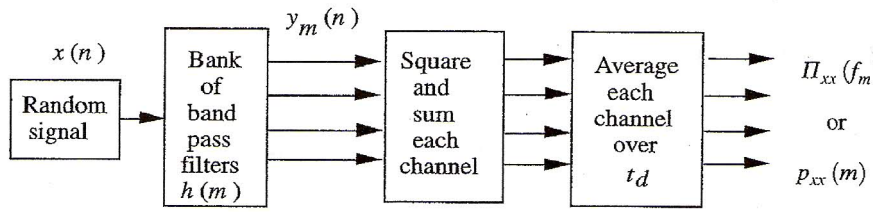

FIGURE 21.30 Block diagram of spectral analysis of a random signal in a computer or a single dedicated instrument; the set of filters $h(m)$ have center frequencies f_m and filter widths Δf_m; the duration of the random signal is t_d.

proportional to the filter bandwidth and the duration of the signal, it is customary to define the power spectral density by Equation (21.65). The first step in deriving an equivalent integral expression for continuous functions of time uses the multiplication and division by t_0:

$$\Pi_{xx}(f_m) = \frac{1}{Nt_0\Delta f_m}\sum_{n=0}^{N-1} y_m^2(n)t_0 \tag{21.66}$$

Let Nt_0 become t_d, the duration of the signal, and t_0 become dt. The summation becomes the integral:

$$\Pi_{xx}(f_m) = \frac{1}{t_d\Delta f_m}\int_0^{t_d} y_m^2(n)\mathrm{d}t \tag{21.67}$$

If $x(n)$ has the units of volts, the so-called power spectral density has units of (volts)^2/Hz. True power spectral density would require division by a load resistance in an electrical circuit to give watts/Hz. Since a hydrophone output in volts is proportional to the acoustic pressure, $x(n)$ has the units of Pa, the spectral density has units of (Pa)^2/Hz, and the true *intensity spectral density* requires division by $\rho_A c$, to give

$(\text{Pa})^2/\rho_A c\text{Hz} = \text{watts/m}^2\text{Hz}$. Acoustic spectra are often reported in dB relative to one $(\mu\text{Pa})^2/\text{Hz}$, so that the intensity spectrum level (ISL):

$$\text{ISL} = 10 \log_{10} \left\{ \frac{\Pi_{xx}(f_m)}{\dfrac{(\mu\text{Pa})^2}{\text{Hz}}} \right\} \tag{21.68}$$

The spectrum levels depend on the reference sound pressure, which is sometimes unclear. It is better to use *Pascal* units such as $(\text{Pa})^2/\text{Hz}$ or $\text{watts/m}^2\text{Hz}$.

Spectral Smoothing. Consider the following example of spectrum analysis. A random signal is constructed of 512 magnitudes at separation $t_0 = 0.001$ sec and duration 0.512 sec. Figure 21.31 shows the results of processing the signal by the equivalents of very narrow, wide, and very wide bandpass filters. The output of the narrow 2-Hz filter (Figure 21.31(a)) is extremely rough. The number of independent samples given by Equation (21.54) in the 2-Hz bandpass filter is 1. Figure 21.31(b) shows the result of using a wider filter, $\Delta f = 64$ Hz. Here the number of independent samples is 32. The spectrum is much smoother and has less detail. An increase in the filter width to $\Delta f = 128$ Hz and the number of independent samples to 64 is shown in Figure 21.31(c). Another random signal would have a different spectrum. These examples show the basic trade-off between resolution and reduction of roughness or variance of the estimate of the spectral density. The importance of smoothing power spectra and the trade-off between the reduction of frequency resolution and the reduction of fluctuations is given in detail by Blackman and Tukey in their monograph *The Measurement of Power Spectra* [1958].

Traditional Measures of Sound Spectra. The measurement of underwater sounds has inherited the instrumentation and the vocabulary that were developed for measurements of sounds heard by humans in air.

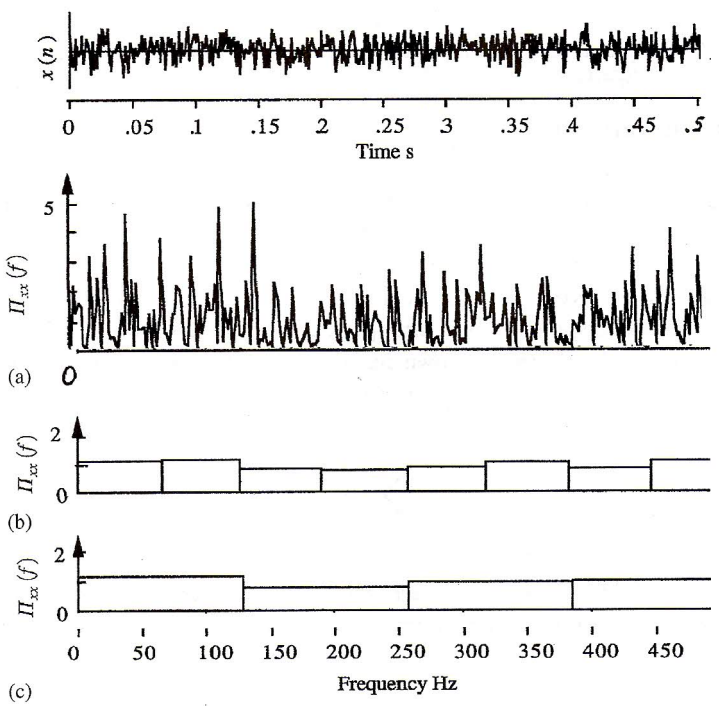

FIGURE 21.31 Smoothing of power spectra by filtering. The top trace is a random signal $x(n)$ or $x(t)$. Filter bandwidths are (a) $\Delta f = 2$ Hz, (b) $\Delta f = 64$ Hz, and (c) $\Delta f = 128$ Hz.

The principal areas of interest to humans have been acoustic pressure threshold for hearing; acoustic threshold of damage to hearing; threshold for speech communication in the presence of noise; and community response to annoying sounds. The vast amount of data required to evaluate human responses, and then to communicate the recommendations to laymen, forced psychoacousticians and noise-control engineers to adopt simple instrumentation and a simple vocabulary that would provide simple numbers for complex problems. Originally this was appropriate to the analog instrumentation. But even now digital measurements are reported according to former constraints. For example, the *octave band*, which is named for the eight notes of musical notation that corresponds to the 2:1 ratio of the top of the frequency band to the bottom, remains common in noise-control work. For finer analysis, one-third octave band instruments are used; they have an upper-to-lower-band frequency ratio of $2^{0.33}$, so that three bands span one octave.

The use in water of instruments and references that were designed for air has caused great confusion. The air reference for acoustic pressure level in dB was logically set at the threshold of hearing (approximately 20 μPa at 1000 Hz) for the average adult human. This is certainly not appropriate for underwater measurements, where the chosen reference is 1 μPa or 1 Pa. Furthermore, plane-wave intensity of CW is calculated from Equation 2.5.16 in Reference 1, where Intensity $= P_{\text{rms}}^2/\rho_A c$ (where P_{rms}^2 is the mean squared pressure; ρ_A is the water density; and c is the speed of sound in water). Therefore, the dB reference for sound intensity in water is clearly different from that in air because the specific acoustic impedance $\rho_A c$ is about 420 kg/m^{2}sec for air compared with 1.5×10^6 for water. This ratio corresponds to about 36 dB, if one insists on using the decibel as a reference.

The potential for confusion in describing the effects of sound on marine animals is aggravated when physical scientists use the decibel notation in talking to biological scientists. Confusion will be minimized if psychoacoustical characteristics of marine mammals—such as thresholds of pain, hearing communication perception, and so forth—are described by the use of SI units (i.e., pascals; acoustic pressure at a receiver), watts/m^2 (acoustic intensity for CW at a receiver) and joules/m^2 (impulse energy/area at a receiver). Likewise, only SI units should be used for sources—that is, watts (power output of a continuous source) and joules (energy output of a transient impulse source). The directivity of the source should always be part of its specification. All of these quantities are functions of sound frequency and can be expressed as *spectral* densities (i.e., per 1-Hz frequency band).

Matched Filters and Autocorrelation

The coded signal and its matched filter and associated concepts have become very important in the applications to sound transmission in the ocean. The simple elegance of the original paper [8] is well worth a trip to the library. The generality of their concepts was far ahead of the then-existing signal processing methods. Digital signal processing facilitates the design of many types of filters for processing sonar signals. Each definition of an optimum condition also defines a class of optimum filter. We limit our discussion to the simplest of the optimum filters, the *matched filter* [8]. An example is shown in Figure 21.32.

An example of a coded signal $x(n)$ is shown in Figure 21.32(b). Recalling the convolution summation, Equation 6.2.29 in Reference 1, the convolution of $h_M(n)$ and $x(n)$ is

$$y_M(j) = \sum_{m=0}^{m_1} x(m) \cdot h_M(j-m) \tag{21.69}$$

where the subscript $_M$ means the matched filter. The matched filter uses the criterion that the square of the peak output value $y_M(0)$ is a maximum. To maximize the square of $y_M(0)$, we use Cauchy's inequality [9]:

$$y_M^2(0) \leqslant \sum_{m=0}^{m_1} h_M^2(m) \times \sum_{m=0}^{m_1} x^2(m) \tag{21.70}$$

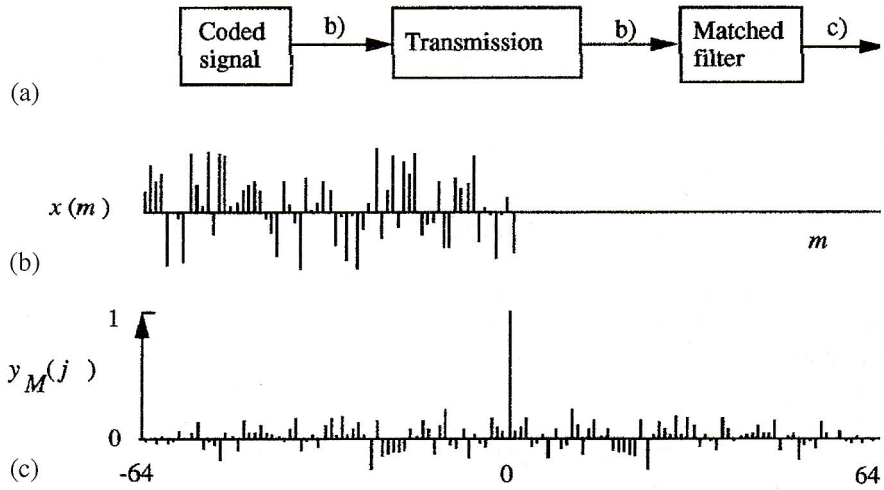

FIGURE 21.32 Coded signal, matched filter, and output: (a) simple matched filter system; (b) coded signal (here, the coded signal is a short sequence from a random-number generator); (c) output of the matched filter.

The filter produces a maximum output (the equal sign in Equation (21.70)) when $x(m)$ is time-reversed to $x(m) = x(-m)$:

$$h_{\mathrm{M}}(m) = Ax(-m) \quad \text{or} \quad h_{\mathrm{M}}(-m) = Ax(m) \tag{21.71}$$

where A is a constant of proportionality. In our examples we use $A = \dfrac{1}{m_1 + 1}$. The filter defined by Equation (21.71) is called the *matched filter*. In computations, it is convenient to shift the indices by m_1 and to shift the time by $m_1 t_0$:

$$h_{\mathrm{M}}(m_1 - m) = h_{\mathrm{M}}(-m) \tag{21.72}$$

to give a causal filter. Except for constants and normalization, $y_{\mathrm{M}}(j)$ has the same form as the autocorrelation (see Equation (21.54)):

$$y_{\mathrm{M}}(j) = A \sum_{m=0}^{m_1} x(m) \cdot x(m + j) \tag{21.73}$$

The output of the matched filter is shown in Figure 21.32(c). Here, the output of the matched filter is the autocorrelation or covariance of $x(n)$. The signal-to-noise amplitude ratio gain is proportional to the square root of the number of independent samples of the coded signal. While we do not prove it, ignoring noise in our derivation is equivalent to assuming that the noise is an uncorrelated sequence of random numbers having a mean value of zero. This kind of noise is often called *white noise*.

In 1962, Parvulescu obtained a classified patent for the use of the *matched equivalent signal*, for measuring the reproducibility of signal transmissions over large ranges in the ocean. In this first use of the matched filter technique in the ocean, the multipath received signals were regarded as a coded signal. An analog tape recorder was employed, with the tape direction reversed to convert the multipath arrivals into a matched filter [10,11].

Summary

Development and employment of acoustical techniques allow us to image underwater features, communicate information via the oceanic wave-guide or measure oceanic properties. Representative applications of these techniques can be summarized in the following form:

- *Image underwater features*: detection, classification and localization of objects in the water column and in the sediments using monostatic or bistatic sonars; obstacle avoidance using forward-looking sonars; navigation using echo sounders or sidescan sonars to recognize sea-floor topographic reference features.
- *Communicate information via the oceanic wave-guide*: acoustic transmission and reception of voice or data signals in the oceanic wave-guide; navigation and docking guided by acoustic transponders; release of moored instrumentation packages using acoustically activated mechanisms.
- *Measure oceanic properties*: measurement of ocean volume and boundaries using either direct or indirect acoustical methods; acoustical monitoring of the marine environment for regulatory compliance; acoustical surveying of organic and inorganic marine resources.

Nomenclature

Symbol	Quantity	Symbol	Quantity
ASE	Amplified spontaneous emission	IFFT	Inverse finite Fourier transformation
ASW	Anti submarine warfare		
CW	Continuous waves	ISL	Intensity spectrum level
DFT	Discrete Fourier transformation	SI	Sound intensity
FFT	Finite Fourier transformation	TVG	Time-varying gain

References

1. H. Medwin and C. Clay, *Fundamentals of Acoustical Oceanography*, London: Academic Press, 1998.
2. H. Medwin, "Speed of sound in water: a simple equation for realistic parameters," *J. Acoust. Soc. Am.*, 58, 1318–1319, 1975.
3. V. Knudsen, R. Alford, and J. Emliing, "Underwater ambient noise," *J. Marine Res.*, 7, 410–429, 1948.
4. J. Traynor and J. Ehrenberg, "Fish and standard-sphere, target-strength measurements obtained with a dual-beam and split-beam echo-sounding system," *Rapp. P.-V. Reun. Cons. Int. Explor. Mer.*, 189, 325–335, 1990.
5. R. Craig and S. Forbes, "Design of a sonar for fish counting," *Fisk. Dir. Ser. Havunders.*, 15, 210–219, 1969.
6. P. Etter, *Underwater Acoustic Modeling and Simulation*, 3rd ed., London: Spon Press, 2003.
7. L. Ziomek, *Acoustic Field Theory and Space–Time Signal Processing*, Boca Raton, FL: CRC Press, 1995.
8. J. Van Vleck and D. Middleton, "A theoretical comparison of the visual, aural, and meter reception of pulsed signals in the presence of noise," *J. Appl. Phys.*, 17, 940–971, 1946.
9. M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standard Applied Mathematical Series 55, Washington, D.C.: Government Printing Office, 1964.
10. A. Parvulescu, "Matched-signal (MESS) processing by the ocean," *J. Acoust. Soc. Am.*, 98, 943–960, 1995.
11. I. Tolstoy and C. Clay, *Ocean Acoustics*, New York: McGraw-Hill, 1966. Reprinted by the Acoustical Society of America, Woodbury, NY, 1987.