SIGNALS AND SIGNAL PROCESSING FOR ACOUSTIC MONITORING OF OCEAN PROCESSES

Theodore G. Birdsall

Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109-2122
email: birdsall@eecs.umich.edu

ABSTRACT

Systems for bistatic monitoring of ocean processes differ radically from monostatic systems searching for targets. This paper discusses basic principles, showing how these effect signal design, signal processing goals and signal processing techniques. In addition, computer advances have a strong influence in changing signal processing choices, sometimes rendering yesterday's optimums today's curiosities, and perhaps converting today's impossibilities into tomorrow's challenges.

1. INTRODUCTION

This paper and talk are based on experience with long range systems in support of propagation research or physical oceanography research. These applications have generated a steady stream of problems, and signal processing solutions. We hope you will find the problems and solutions interesting and provocative.

Actually we have only three solutions: (1) use m-sequences, (2) "divide and multiply", and (3) maximum likelihood ratio searches. These will be explained after setting the stage with some systems and physics.

1.1. Monitoring System Restrictions

The primary variable measured is the time it takes sound to travel from source to receiver. In long range work this is not a single time, but a set of times, each associated with a different path or mode of propagation. Each arrival is tagged by its traveltime, amplitude, phase and perhaps other features, and the set is one big joint random process.

Many traveltime measurement systems such as airport radars, navigation sonars and baseball speed-guns are *monostatic* systems, meaning the transmitter and the receiver are at the same place. Monostatic systems use short transmissions, followed by long silences to listen for the echo. The spreading loss at range R is basically R^4, severely limiting the range.

Using *bistatic* systems, the transmitter need not turn off while the *distant receiver* listens, so the transmission can be very long; this is one key to tens of dBs of signal processing gain. The spreading loss dependence on range R may be as weak as R^1 for total received power or for a fixed number of non-interacting modes, or be proportional to R^2 for ray propagation where the number of arrivals increases like R^1.

Measuring arrival times, and having multiple arrivals, means that the system must have adequate time-resolution to separate the arrivals. Every signal processer knows that resolution (the half-height width of the signal's autocorrelation function) is inversely proportional to the signal's bandwidth, so the system should be as wideband as practical. Since bistatic systems can have large duration signals, time measurement systems are large time-bandwidth (TW) systems. The *spread spectrum* signal processer foresees a "W" in the noise gain, but it isn't to be. The underwater acoustic background noise is dominated by large numbers of broadband sources, not intentional jammers, so the spread spectrum noise reduction relative to the per-hertz level is just "T". "W" is used to achieve resolution.

It is important for the signal processer to know that the peak-power limit of most well-designed acoustic transmitters will dominate the average power limit. (The transmitter sits in a big water bath. Even short pulses last long enough for heat to accumulate. Peak power limits are set by oil surround breakdown voltages, elastic limits of vibrating elements, and the like.) This robs the signal processor of many favorite theorems that begin with "the optimum ...".

1.2. Propagation Facts

Things happen slowly underwater. The speed of sound underwater is roughly 1.5 km/sec, which is the speed of light in vacuum divided by 2×10^5. Underwater sound will take roughly 100 minutes to travel 3 light-

 0-7803-2431-5/95 $4.00

Table 1: Atlantic & Pacific Attenuation in dB

Frequency (Hz)	1 Mm		10Mm	
	Atlantic	Pacific	Atlantic	Pacific
50	—	—	3	1.6
100	1.2	.7	12	6.5
150	2.7	1.5	27	15
250	7.1	3.9	71	39
500	25	13.7	—	—
750	46	27	—	—
1000	66	38	—	—

milliseconds, 9 Mm. (10 Mm, 10 megameters, is the distance from pole to equator.) The water itself has high inertia, and changes in currents and locations of features (such as fronts and eddies) require longer times than comparable changes in the atmosphere. These are in the signal processers favor if one is using searches in large parameter spaces, as there may be hours between measurement events during which the computer is furiously byteing away at the processing.

Typical acoustic propagation ranges depend on the application. A classical sonar range is 2-4 km, as the sound refracts downward away from surface or near-surface objects due to the temperature decrease with depth. A higher power sonar may bounce power off the bottom and achieve ranges of 50-200 km. Ocean acoustic tomography of the 1980's monitored medium scale eddies with diameters of the order of 100 km and worked ranges of 200km-2Mm using the refractive SOFAR channel. Current work in acoustic thermometry of ocean climate plans to use ranges of 3-10 Mm.

Once one requires several watts acoustic, every 10 dB increase in power will raise the price by a factor of 10, raise the weight by a factor of 10 and hence the cost of installation from a tossing ship, and so on. Table 1 shows the attenuation loss at 1 and 10 Mm, for the mid-Atlantic and the mid-Pacific. This attenuation drives the choice of operating frequency down, and raises the value of signal processing gains. The financial pressure to go to low frequencies is apparent.

Normally bandwidth and center frequency scale together. Frequencies have to go low to beat attenuation, bandwidths have to stay up to achieve resolution, so the ratio $Q = f_c/W$ has to decrease; if the hardware engineers can't meet that, the signal processer must (at least, must try). An example is in section 4.

Signal processers should be warned that in the time domain, received amplitudes are much more variable than received phases, even when these are measured on isolated arrivals. This is not a noise effect or gross multipath effect. The data is scarce but the fact is important; do not bet on stable or knowable amplitudes.

2. SIGNAL DESIGN

A signal processer's heaven is to be able to design and implement the signal, reception and processing of a repeatable measurement when the competing designs are 10 dB to 30 dB deficient in signal to noise performance, and the propagation is linear and almost time invariant. It's even better when the frequencies are so low that the signal can be digitized and stored, then read out through a D/A converter and filter, amplified and sent to the transmitter.

The resolution specification is that the signal should have a large TW with an autocorrelation function with acceptable resolution and low ripple; don't go overboard on pushing the ripple too far down, as ripples over 30 dB down may be below the noise level. Power spectra such as the main lobes of sinc-squared or cosine-squared are popular. For these two the *operating band* is twice the *3 dB bandwidth*; don't let the hardware designer confuse the bandwidth with the band over which the hardware must operate.

Part of the power specification is that the signal should drive the transmitter close to peak power at all times, since the hardware limitation is the peak power, but the processed output quality will depend on the total signal energy, i.e., on the average signal power.

2.1. Periodic Signals

I champion the use of transmissions that send many periods of the same waveform. Let me set up the notation and then explain the reasoning. The signal period T_{per} should be about twice the anticipated arrival time spread T_{sprd} including any dispersive effects. This avoids wraparound, covers for estimation errors, and allows the user to identify the arrivals' *start* and *stop* easily. The number of periods analyzed sets the processing gain by setting the duration of the reception to be analyzed, $T_{anal} = PT_{per}$. The actual transmission time is typically $T_{send} \geq T_{anal} + 2\, T_{sprd}$. At the receiver, the user carefully estimates the first arrival time + T_{sprd} and begins recording T_{anal} seconds. The objective is to record and analyze P periods when the signal and ocean are in *steady state*, when all paths are present in the reception even though their amplitudes may be varying. (Remember that the user is interested in what the reception tells about the ocean, not in the details of the reception.)

Periodicity introduces flexibility in processing. In many situations one knows that there is sufficient phase stability so all P periods can be summed to yield a one-period *snapshot* of the reception. If the user suspects that the ocean may be changing significantly, and there is sufficient signal power in the reception, sets of several periods may be summed to increase the S/N, and subsequent processing will yield results that can be tracked, or coherent summations with subsequent incoherent combination may yield the desired results.

Such periodic processing means that the relevant spectral math is Fourier Series, which is a close match to the actual sampling and DFT processing. (This is known as "designing your situation so the math fits the actuality.") This brings us to the actual single period waveform and the methods used to achieve a continuous steady-power wideband waveform that can be *inverted* to a resolvable event, a pulse. We touch on digital codes and on unidirectional FM slides.

A digital code consists of some length L of digits. Let $T_1 = T_{per}/L$. The k^{th} digit has complex value d_k. This is converted to a complex valued waveform based on a pulse waveform $p(t)$ whose resolution is of the order of T_1, but whose duration may be longer that T_1. The waveform is

$$s(t) = \sum_{k=0}^{L-1} d_k p(t - kT_1) \qquad 0 < t < T_{per}$$

The spectrum of this signal is

$$S(f) = A \sum_{k=0}^{L-1} d_k \, e^{-j2\pi f k T_1} \, P(f) = AD(f)P(f)$$

If the code is an m-sequence using two phases, $D(f)$ has the same magnitude at all the eigenfrequencies $f_m = m/T_{per}$ with the possible exception of $f = 0$; if furthermore the phases are $\pm tan^{-1}(\sqrt{L})$, the $D(f_m)$ spectral magnitude is a constant.

A possible FM signal design [1] is to use a monotone instantaneous frequency achieving a power spectrum $|X(f)P(f)|^2$ approximating the desired power spectrum $|P(f)|^2$. Here the magnitude of $X(f_m)$ is reasonably constant and contains the error in approximating the power spectrum, while the phase of $X(f)$ is related to the instantaneous frequency. If the instantaneous frequency is ideal, the magnitude of $X(f_m)$ is a constant. So for FM we may write $S(f) = AX(f)P(f)$.

Spectral factorization opens a signal processing door [2], using the ***divide and multiply*** thought. If $S(f) = AX(f)P(f)$, then the spectrum of the signal part of the received snapshot is the multipath sum

$$R(f) = \sum_{p} c_k \, e^{-j2\pi f \tau_k} \, S(f)$$

The Factor Inverse Filtering step is the division

$$R(f)X^{-1}(f) = A \sum_{p} c_k \, e^{-j2\pi f \tau_k} \, P(f)$$

which results in a picture in the time domain as if a large pulse were transmitted

$$IFFT(RX^{-1}) = A \sum_{p} c_k \, p(t - \tau_k)$$

For better S/N, "match" to p(t) by using the final multiplication; the FIMF result is

$$R(f)X^{-1}(f)P^*(f) = A \sum_{p} c_k \, e^{-j2\pi f \tau_k} \, |P(f)|^2$$

FIF has better resolution but poorer S/N than the FIMF result. The beauty of this is that the user may view both the FIF and FIMF outputs and does not have to choose "which is better".

This processing discussion was spectral. For 3 decades it was done with fast time-domain algorithms. The pinnacle was in the use of the Hadamard transform for m-sequences; these have ultimate speed advantages of 30-100 or so over PFA-FFT spectral processing because they use ***real, integer, demultiplexing*** algorithms instead of ***complex, floating-point*** algorithms. The Hadamard's disadvantage was that few signal processers understood its use. Its time has passed. PCs are now so big, so fast and so inexpensive that the well-understood spectral processing is usually adequate.

3. DOPPLER & DISPERSION

Searching for either the rate of a linear doppler or for the amount of dispersion that has affected arriving modes depends on routines that search over a sufficient mesh of possible parameter values. (For simplicity assume that either doppler or dispersion occur, but not both.) They are similar in that the searches attempt to cause local peaks to appear which can be identified as meaningful arrivals. They differ in two aspects: the dispersion processing must occur after the FIF stage that makes the arrivals appear pulse-like, and the amount of dispersion may be different for each arrival, while the doppler search must occur jointly with the FIF processing, and usually the amount of doppler is (nearly) the same for all arrivals.

Doppler means that the traveltime of each arrival is changing during the observation, due to source or receiver motion or acceleration of ocean currents. If an isolated arrival is

$$r_a(t) = As(t - \tau_a(t)) = As(t - \tau_a(0) - \tau_a'(0)\, t - ...)$$

and this is rearranged as

$$r_a(t) = As([1 - \tau_a'(0)]\,t - \tau_a(0) - ...)$$

it is obvious that Doppler is a scaling or compression effect, not a frequency shift (unless the signal is a monochromatic tone). In simple situations the derivative becomes the familiar "v/c" or "$v\ \cos(\theta)/c$". In digital processing the search involves interpolating the sampled data to achieve a new sampling rate so that the sample values match what they would have been if there were no doppler. The resampled data is then FIF processed, and if the rate is nearly right the arrival peaks appear. The dopplered signal first zeros when $\tau_a'(0) = \pm 1/f_c T_{anal}$, so a practical step size in this search variable is $1/4 f_c T_{anal}$.

4. TRANSMITTER Q LOWERING

In a recent project, to obtain a transmitted resolution of 64 ms. a transmitter purchase specified a transmitter bandwidth of 20 hertz. This was changed by the contract writer into an "operating band" of 20 hertz, meaning it had to be able to deliver the specified output power while transmitting a single steady tone at any frequency in that 20 hertz band. The delivered product had a 3-dB bandwidth of 7 hertz, an apparent disaster. It did have some dB to sell in its power margin.

Two signal processing cures were designed. The easier one used a periodic am/fm slide designed to drive the transmitter as hard as permitted at each frequency, with instantaneous frequency trajectory chosen to produce a desired power spectrum [1].

The second method is based on the divide & multiply principle. The manufacturer provided an equivalent circuit with points of critical voltage and mechanical stress identified; some of the parameters had uncertain depth dependencies, and most of the elements were frequency dependent as was the terminating radiation impedance. In situ measurements using the power amplifier, miles of cable and the actual transducer are made to obtain the actual transfer function $H_{actual}(f)$ at low excitation level. The user's signal processer makes an apt selection of what the desired transfer function should be, $H_{desired}(f)$, and what the corresponding desired drive signal would be; designate the latter by its spectrum $S_{desired}(f)$.

In concept, apply the desired signal to $H_{desired}(f) \div H_{actual}(f)$ and call the output $S_{actual}(f)$. Then apply $AS_{actual}(f)$ to the model of the actual system; $H_{actual}(f)$ cancels and the output is $A\ S_{desired}(f)\ H_{desired}(f)$. Raise the level of the input until the model touches a critical limitation; this determines the maximum output power level one may use safely. Now apply $A\ s_{actual}(t)$ to the actual hardware and check the results by using the actual signal processing that will be used by the receivers.

There are some subtle aspects that constitute the real work. The signal processer's *desired* choices require understanding and experience. Desire too much, and you may get too little output power. Division by zero is forbidden; if the actual transmitter has spectral zeros or near-zeros it indicates a power-sink at such frequencies, a vampire capable of sucking the life out of any signal amplifier attempting to feed it at these frequencies. One chooses $H_{desired}(f)$ to avoid powering at such frequencies, and also avoids them in all measurements. If pulse-like digits are desired, the drive signals may have spectral zeros, so will the measurements; this sets up a 0/0 situation in measuring $H_{actual}(f)$. The measurements are made at high S/N, so noise is not so critical, but quantization errors may be painful in 0/0.

Kurt Metzger uses a combination of high sampling rate, with median filtering followed by linear filtering to obtain reasonable input and output spectra near enough to spectral zeros in $S_{desired}(f)$ that the computer can make reasonable evaluations of $H_{actual}(f)$ without human supervision. Finally, there are problems if the measurements are desired near reflecting boundaries; one must first work at mid-depths to adjust the parameter values of a free-field model, and then use the model alone near boundaries. In all such operations it is important to keep the math linked to reality; one must do spectral operations at exactly the eigenfrequencies of the measurement and operational systems.

5. ACKNOWLEDGEMENTS

This work is supported by the Advanced Research Projects Agency (ARPA) on the project Acoustic Thermometry of Ocean Climate (ATOC).

6. REFERENCES

[1] S. Wang, M.L. Grabb, and T.G. Birdsall, "Design of periodic signals using FM Sweeps and amplitude modulation for ocean acoustic traveltime measurements," *IEEE J.Oceanic Engr*, Oct 1994

[2] T.G. Birdsall and K. Metzger, Jr., "Factor inverse matched filtering," *Journ. Acoustical Soc Am*, vol.79, no.1, pp.91-99, Jan 1986