Signal processing and correlation techniques

PETER T. GOUGH
Acoustics Research Group,
Department of Electrical and Computer Engineering,
University of Canterbury, Christchurch, New Zealand

Preamble

Frequently we need to measure a precise time difference between two signals. This is especially true of sonar and radar systems but has applications in general signal processing, communications and biomedical engineering. Almost always we are using narrow bandwidth modulated waveforms that have been sampled so first we need to look at sampling theory.

1. Sampling Theory

The following arguments are developed for signals parameterized by time, t, and angular or radian frequency, ω. These arguments are equally applicable to signals parameterized in terms of spatial quantities such as distance, x, and wavenumber k_x. Both temporal and spatial signals are analysed extensively.

To represent a continuous time domain signal $g(t)$ digitally, it is necessary to *sample* the signal. This (impulse) sampling operation is described by

$$\begin{aligned} g_t(t) &= g(t) \cdot \sum_{m=-\infty}^{\infty} \delta(t - m\Delta_t) \\ &= \sum_{m=-\infty}^{\infty} g(m\Delta_t) \cdot \delta(t - m\Delta_t), \end{aligned} \tag{0.1}$$

where Δ_t is the sampling interval and $g_t(t)$ has a subscript 't' to indicate *temporal* sampling. It is important to note that (0.1) is still a *continuous* representation of the sampled signal. However, the sampled signal is fully characterised by the m samples given by $g(m\Delta_t)$.

The continuous Fourier transform of (0.1) is

$$\begin{aligned} G_t(\omega) &= G(\omega) \odot_\omega \frac{2\pi}{\Delta_t} \cdot \sum_{m=-\infty}^{\infty} \delta\left(\omega - m\frac{2\pi}{\Delta_t}\right) \\ &= \frac{2\pi}{\Delta_t} \cdot \sum_{m=-\infty}^{\infty} G\left(\omega - m\frac{2\pi}{\Delta_t}\right), \end{aligned} \tag{0.2}$$

where $\odot_\omega$ represents convolution in ω and the Fourier transform operation is defined in (**??**). The effect of sampling the function $g(t)$ is to generate repeated copies of its scaled continuous spectrum $2\pi/\Delta_t \cdot G(\omega)$ every $m2\pi/\Delta_t$. The repeated nature of this continuous spectrum is important when dealing with the array theory presented in Chapter **??**.

Equation (0.2) represents a *continuous* function and as such also needs to be sampled for use in digital applications. This *spectrally* sampled signal is given by

$$\begin{aligned} G_s(\omega) &= G_t(\omega) \cdot \sum_{n=-\infty}^{\infty} \delta(\omega - n\Delta_\omega) \\ &= \sum_{n=-\infty}^{\infty} G_t(n\Delta_\omega) \cdot \delta(\omega - n\Delta_\omega), \end{aligned} \tag{0.3}$$

which has the inverse Fourier transform,

$$\begin{aligned} g_s(t) &= g_t(t) \odot_t \frac{2\pi}{\Delta_\omega} \cdot \sum_{n=-\infty}^{\infty} \delta\left(t - n\frac{2\pi}{\Delta_\omega}\right) \\ &= \frac{2\pi}{\Delta_\omega} \cdot \sum_{n=-\infty}^{\infty} g_t\left(t - n\frac{2\pi}{\Delta_\omega}\right). \end{aligned} \tag{0.4}$$

The effect of frequency sampling is to repeat copies of the scaled *temporally* sampled signal $2\pi/\Delta_\omega \cdot\ g_t(t)$ every $n2\pi/\Delta_\omega$. The repeated nature of this temporally and spectrally sampled signal seems to imply that the data is corrupted by the repetition. This observation is true if m or n are allowed to take on arbitrarily high values. A digital processor can not deal with an infinite number of samples, so that the values of m and n must be finite and as such there exists a value for m and n such that the sampled data suffers minimal corruption. If the temporal signal is repeated every $2\pi/\Delta_\omega$ and is sampled every Δ_t then m can take on the values $m \in [1, M]$, where

$$M = \frac{2\pi}{\Delta_t \Delta_\omega}. \tag{0.5}$$

Similarly, if the spectrum is repeated every $2\pi/\Delta_t$ and sampled every Δ_ω then $n \in [1, N]$, where N takes on the same value as M. The discrete representation of the time-limited temporal and frequency-limited spectrally sampled signals are the essence of the discrete Fourier transform (DFT) and its efficient implementation via the fast Fourier transform (FFT). The FFT of a temporal signal containing M samples gives a *single copy* of the *repeated* spectrum, and conversely the inverse FFT of a spectral signal containing N samples gives a single copy of the repeated temporal signal.

1.1. Real signals and Nyquist rate sampling

The signals used in synthetic aperture systems can be described in terms of *band-limited* functions. A band-limited real function, $g_r(t)$, having an amplitude function $g_0(t)$ and phase function $\phi(t)$ at a carrier radian frequency ω_0 is mathematically described by (p89 [3])

$$g_r(t) = g_0(t) \cos\left[\omega_0 t + \phi(t)\right]. \tag{0.6}$$

The amplitude function $g_0(t)$ is a slowly varying function also referred to as the envelope of the signal. In the range dimension of the synthetic aperture model, this amplitude function reflects the weighting of the transmitted pulse and in the along-track dimension, it reflects the effect of the overall radiation pattern of the transmit and receive real apertures.

In the case of purely amplitude modulated (AM) pulses, the signal bandwidth is approximately given by the inverse of the temporal duration at the 3dB point of $g_0(t)$. For the more usual case of some form of phase modulation (PM), the system bandwidth is determined by the the modulating phase. In either case, if the spectrum of the real signal is zero for radian frequencies $|\omega| \geq \omega_{\max}$, then temporal samples taken at spacings $\Delta_t \leq \pi/\omega_{\max}$ are sufficient to reconstruct the continuous signal (see Appendix A in Curlander and McDonough [2]). Sampling at the minimum allowable temporal spacing $\Delta_t = \pi/\omega_{\max}$ is termed Nyquist rate sampling. The *sampling frequency* $f_s = 1/\Delta_t$ or $\omega_s = 2\pi/\Delta_t$ represents the extent of the Fourier domain that is sampled by the system, and also represents the distance over which this sampled information repeats in the Fourier domain.

For the real valued temporal signals encountered in synthetic aperture imaging applications, the *sampled bandwidth* is typically much larger than the transmitted *signal bandwidth*. Radar systems typically transmit a wide bandwidth signal (many MHz) about a high carrier frequency (many GHz), so Nyquist rate sampling is suboptimal. There are two more efficient sampling methods that can be employed. The first technique involves multiplying (demodulating) the received signal $g_r(t)$ with

an *intermediate frequency* (IF) that acts to shift the signal bandwidth to a much lower carrier. This intermediate signal is then Nyquist rate sampled at a significantly lower rate. The second technique involves the sampling of the *complex-valued baseband signal* that is described next.

1.2. Complex-valued baseband signals

The real *band-limited* signal given by $g_r(t)$ in (0.6) can also be represented by the real part of its *complex pre-envelope* or *complex modulated representation*, $g_r(t) = \text{Real}\{\mathbf{g}_m(t)\}$, where the complex modulated signal implied by the bold face notation is given by (pp60-64 and p137 [1], pp83-91 [3],pp10-24 [5])

$$\mathbf{g}_m(t) = g_0(t) \exp\left[j\omega_0 t + j\phi(t)\right]. \tag{0.7}$$

Because only positive frequencies exist in the complex modulated form of the signal, it is also known as the one- sided form of $g_r(t)$. To generate the *complex baseband form* of the signal from the complex modulated signal requires the removal of the carrier:

$$\mathbf{g}_b(t) = \mathbf{g}_m(t) \exp(-j\omega_0 t) = g_0(t) \exp\left[j\phi(t)\right]. \tag{0.8}$$

The use of complex mathematics simplifies the modeling of the synthetic aperture system considerably. Also, as most practical systems perform operations on the complex basebanded samples, this is also the most appropriate format in which to develop the processing mathematics.

The conversion of a *real* Nyquist rate sampled signal into a complex signal is most easily performed by removing the negative frequencies in the signal. To achieved this, the signal is Fourier transformed and the negative frequency components are set to zero. This operation halves the signal power, so the one-sided spectrum is then multiplied by 2. This new spectrum is then inverse Fourier transformed to give the complex modulated form of the signal. The complex modulated signal is then demodulated to give the complex baseband signal (see p86 [3]). If this complex baseband, Nyquist-rate sampled signal is down-sampled (decimated) to the same sampling rate as quadrature sampling of the same signal, essentially identical samples are produced.

1.3. In-phase and quadrature (I-Q) sampling

In-phase and quadrature (I-Q) sampling splits the input signal into two paths. In the first path the signal is multiplied (mixed) with an in-phase version of the carrier, the result is low-pass filtered, and then sampled; this sample is assigned to the real part of a complex number. The other path is multiplied (mixed) with a quadrature (90 ° out-of-phase) version of the carrier, low-pass filtered and sampled; this sample

is assigned to the imaginary part of a complex number (see fig 2.32 on p88 of [3]). If the spectrum of the baseband complex signal has a bandwidth B_c (Hz), then *complex* samples taken at any rate $\omega_s \geq 2\pi B_c$ are sufficient to reconstruct the continuous signal.

As stated above, it is not practical to Nyquist rate sample radar signals so either an intermediate frequency is employed, or the signal is I-Q sampled. Sonar systems operate at significantly lower frequencies than radar so they have a choice of either Nyquist rate or quadrature sampling. However, the I-Q method allows the use of a lower sampling frequency in most practical situations.

1.4. A comment on spectral notation

A baseband signal spectrum is a "window" of the signal's continuous spectrum. This window has a width ω_s and is centered about a carrier ω_0. Convention denotes the baseband frequencies $\omega \in [-\omega_s/2, \omega_s/2]$, however the spectral coverage actually correspond to frequencies from $\omega_0 \pm \omega_s/2$. When dealing with Nyquist- rate sampled signals, the frequencies are also denoted $\omega \in [-\omega_s/2, \omega_s/2]$ (where ω_s is generally much higher than the baseband case), however, this time the frequencies correspond to the correct spectral components. To avoid notational confusion, this text denotes the baseband frequencies as $\omega_b \in [-\omega_s/2, \omega_s/2]$ and the actual frequencies as ω, where $\omega = \omega_b + \omega_0$ when dealing with baseband signals.

2. Convolution and correlation

Many time or range resolving techniques rely on complex cross-correlation and this is a good place to summarise the various correlation operations and their convolutional equivalents. To start, assume we have two arbitrary complex time functions $\mathbf{g}(t)$ and $\mathbf{h}(t)$ where the Fourier transform relationship is summarised as

$$\begin{aligned} \mathbf{g}(t) &\Leftrightarrow G(\omega) = \int_t \mathbf{g}(t) \exp(-j\omega t)\, dt \\ \mathbf{h}(t) &\Leftrightarrow H(\omega) = \int_t \mathbf{h}(t) \exp(-j\omega t)\, dt \end{aligned}$$

then a few equivalences

$$\begin{aligned} \mathbf{h}(-t) &\Leftrightarrow H(-\omega) \\ \mathbf{h}^*(t) &\Leftrightarrow H^*(-\omega) \\ \mathbf{h}^*(-t) &\Leftrightarrow H^*(\omega) \end{aligned}$$

So if we now start with convolution

$$\mathbf{y}(t) = \mathbf{g}(t) \odot \mathbf{h}(t) = \int \mathbf{g}(\tau) \cdot \mathbf{h}(t-\tau)\, d\tau, \tag{0.9}$$

Fourier theory gives us its frequency domain equivalent

$$Y(\omega) = G(\omega) \cdot H(\omega)$$

Now for correlation. If

$$\begin{aligned} \mathbf{y}(t) = \mathbf{g}(t) \star \mathbf{h}(t) &= \int \mathbf{g}(\tau) \cdot \mathbf{h}^*(\tau - t)\, d\tau \\ &= \mathbf{g}(t) \odot \mathbf{h}^*(-t), \\ \text{then} \qquad Y(\omega) &= G(\omega) \cdot H^*(\omega). \end{aligned} \tag{0.10}$$

However if

$$\begin{aligned} \mathbf{y}(t) = \mathbf{g}(t) \star \mathbf{h}^*(-t) &= \int \mathbf{g}(\tau) \cdot \mathbf{h}(t-\tau)\, d\tau, \\ &= \mathbf{g}(t) \odot \mathbf{h}(t) \\ \text{then} \qquad Y(\omega) &= G(\omega) \cdot H(\omega) \end{aligned} \tag{0.11}$$

which is exactly the same as straightforward convolution, but if

$$\begin{aligned} \mathbf{y}(t) = \mathbf{g}(t) \star \mathbf{h}(-t) &= \int \mathbf{g}(\tau) \cdot \mathbf{h}^*(t-\tau)\, d\tau, \\ &= \mathbf{g}(t) \odot \mathbf{h}^*(t), \\ \text{then} \qquad Y(\omega) &= G(\omega) \cdot H^*(-\omega). \end{aligned} \tag{0.12}$$

So finally if

$$\begin{aligned} \mathbf{y}(t) = \mathbf{g}(t) \star \mathbf{h}^*(t) &= \int \mathbf{g}(\tau) \cdot \mathbf{h}(\tau - t)\, d\tau, \\ &= \mathbf{g}(t) \odot \mathbf{h}(-t) \\ \text{then} \qquad Y(\omega) &= G(\omega) \cdot H(-\omega) \end{aligned} \tag{0.13}$$

and that completes the matched set for the complex correlation operations that form the basis of many algorithms especially those used to estimate fine time delays.

3. Correlation techniques and limits of accuracy

Many time delay estimation techniques use cross-correlation of two sequential echoes to get fine time delay differences. Consequently it is appropriate to review the techniques of cross-correlation — and the limits

of accuracy. We can assume the transmitted signals and the two noise-corrupted, real echo signals have a double-sided, narrow-band, spectrum $G_r(\omega)$ with the positive and negative frequency response pefectly symetrical about the carrier frequency ω_0. In that case,

$$\begin{aligned} g_{r0}(t) &= \mathcal{F}^{-1}\{G_r(\omega)\} \\ g_{r1}(t) &= \mathcal{F}^{-1}\{G_r(\omega)\exp(-j\omega t_1)\} + n_1(t) \qquad (0.14) \\ g_{r2}(t) &= \mathcal{F}^{-1}\{G_r(\omega)\exp(-j\omega t_2)\} + n_2(t) \qquad (0.15) \\ &\approx g_{r1}(t-\Delta t) \end{aligned}$$

where we assume the three real signals g_{r0}, $g_{r1}(t)$ and $g_{r2}(t)$ are pulse-compressed, CW waveforms. Signals 1 and 2 are the received echos, delayed in time by t_1 and t_2 and corrupted by AWGN $n_1(t)$ and $n_2(t)$. We can treat transmitted signal $g_{r0}(t)$ and the two echoes $g_{r1}(t)$ and $g_{r2}(t)$ as having a slowly varying envelope $g_0(t)$ superimposed on a carrier as described here:

$$\begin{aligned} g_{r0}(t) &= g_0(t)\cos(\omega_0 t) \\ g_{r1}(t) &= g_0(t-t_1)\cos(\omega_0(t-t_1)) + n_1(t) \qquad (0.16) \\ g_{r2}(t) &= g_0(t-t_2)\cos(\omega_0(t-t_2)) + n_2(t) \qquad (0.17) \end{aligned}$$

and where $g_0(t)$ is symetrical about $t = 0$. Now the time difference that we want to estimate is $\Delta t = t_2 - t_1$. The cross-correlation between the two real echoes is also purely real and given by

$$\begin{aligned} rr_{21}(t) &= g_{r2}(t) \star g_{r1}(t) \\ &= \frac{\int g_{r2}(\tau)\, g_{r1}(\tau - t)\, d\tau}{\sqrt{\left(\int g_{r0}(\tau)\, g_{r0}(\tau)\, d\tau \cdot \int g_{r1}(\tau)\, g_{r1}(\tau)\, d\tau\right)}} \qquad (0.18) \end{aligned}$$

To save having to trail around scaling constants to ensure that the auto-corrrelation functions are always less than one, let us normalise the equations so that

$$\sqrt{\left(\int g_{r0}(\tau)\, g_{r0}(\tau)\, d\tau \cdot \int g_{r1}(\tau)\, g_{r1}(\tau)\, d\tau\right)} = 1$$

With this requirement, the autocorrelation function is now defined as

$$\begin{aligned} rr_{11}(t) &= g_{r1}(t) \star g_{r1}(t) \\ &= \int g_{r1}(\tau)\, g_{r1}(\tau - t)\, d\tau \qquad (0.19) \end{aligned}$$

and the cross-correlation function as

$$\begin{aligned} rr_{21}(t) &= g_{r2}(t) \star g_{r1}(t) \\ &= \int g_{r2}(\tau)\, g_{r1}(\tau - t)\, d\tau \end{aligned} \tag{0.20}$$

The direct application of (0.19) and (0.20) are the basis of time domain auto- and cross-correlation.

The crosscorrelation function is related to the spectrum by

$$\begin{aligned} rr_{21}(t) &= \mathcal{F}^{-1}\{G_r(\omega) \exp(-j\omega\Delta t) \cdot G_r^*(\omega)\} + n_1(t) \star n_2(t) \\ &= \mathcal{F}^{-1}\{|G_r(\omega)|^2 \exp(-j\omega\Delta t)\} + n_1(t) \star n_2(t) \\ &\approx rr_{11}(t - \Delta t) \end{aligned} \tag{0.21}$$

Thus the cross-correlation of $g_{r2}(t)$ and $g_{r1}(t)$ is almost the same as the autocorrelation of $g_{r1}(t)$ delayed by Δt. The"almost" is needed to allow for the noise cross-correlation between the two signals $n_1(t) \star n_2(t)$ that is slightly different from the noise autocorrelation $n_1(t) \star n_1(t)$. The uncorrelated AWGN can be clearly seen in the two traces of fig **??**(a). This figure shows the results of a simulation of the echoes received from

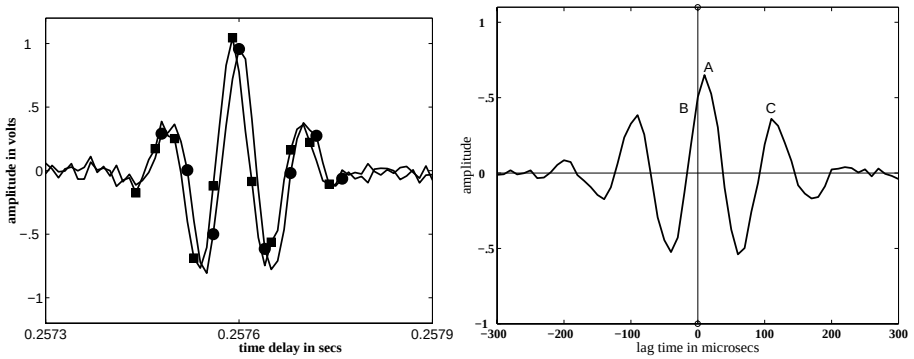


Figure 0.1. (a) Two sequential pulses with a differential delay between them plus AWGN, (b) The cross-correlation of the two sequential pulses where "A" denotes the peak of the cross- correlation, "B" the amplitude at zero lag and "C" the next highest positive peak value.

a small target at a range of 158m with a SNR of about 25. In the plots, the squares indicate the echo returned from ping 1 and the circles that from ping 2 and where the hydrophone has been displaced 7.5mm away from the targets between the two pings. Thus the second ping has a lag (i.e., is differentially delayed by) of 10 μs as is clearly shown in fig. 0.1(a). Clearly in the plot, "A" is accurate and unambiguous and the point "C"

indicates the next highest positive peak. With the SNR and bandwidths used, "C" is clearly smaller than "A" but with decreasing bandwidth, the envelope stretches and/or with increasing noise, the possibility of selecting "C" rather than "A" increases. The peak value of $rr_{21}(t)$ (call it μ and it is the amplitude of the point represented by A in fig. 0.1(b)) must be less than one to account for uncorrelated noise as well as baseline decorrrelation due to any spatial separation and this peak occurs at a time lag of Δt.

But a full correlation calculation is not always required. A simple approximation, but only useful for very small delays, high SNR and high-Q systems, requires the computation of only

$$rr_{21}(0) = \int g_{r2}(\tau) \cdot g_{r1}(\tau)\, d\tau \tag{0.22}$$

since for a pulsed CW with envelope $g_0(t)$, the cross- correlation of the two reflected signals is

$$\begin{aligned} rr_{21}(t) &\approx rr_{11}(t - \Delta t) \\ &= \mu[g_0(t - \Delta t) \star g_0(t - \Delta t)] \cos(\omega_0(t - \Delta t)). \end{aligned} \tag{0.23}$$

When $t = 0$ and assuming a symetrical envelope (i.e., $g_0(t) = g_0(-t)$)

$$rr_{21}(0) \approx rr_{11}(-\Delta t) = \mu[g_0(\Delta t) \star g_0(\Delta t)] \cos(\omega_0 \Delta t)$$

and this is shown as point B in fig. 0.1. So a quick and dirty estimate of Δt can be found without having to compute a full cross-correlation and is

$$\widehat{\Delta t} = (1/\omega_0) \cos^{-1}\{rr_{21}(0)/[g_0(\widehat{\Delta t}) \star g_0(\widehat{\Delta t})]/\mu\}, \tag{0.24}$$

however as $\widehat{\Delta t}$ appears on both sides of (0.24), this needs a recursive solution. The only provisos are that there should be sufficient pulse overlap at zero lag to have some meaning and that μ is known. Hence the need for small delays, high SNR and high-Q (i.e., long pulse) operation.

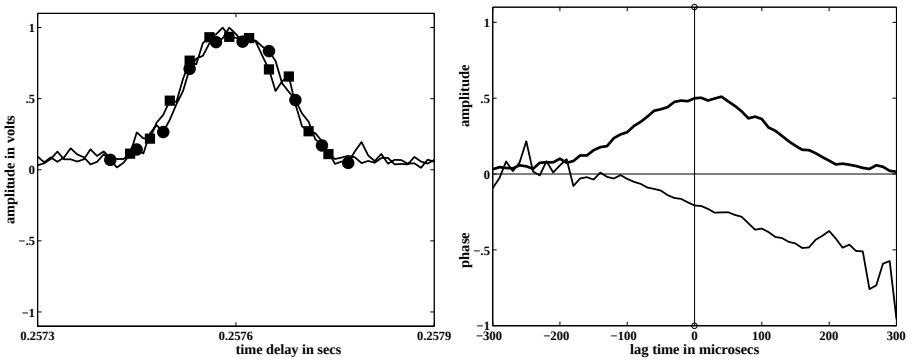
3.1. Complex-baseband Cross-correlation

Few sonars record the raw data sampled at the Nyquist rate. To save on data storage and the subsequent computing effort, the incoming signals are almost always demodulated to complex baseband on the fly and decimated to just twice the baseband. Staying with our convention,

we can write the transmitted signal and the two echo waveforms as

$$\begin{aligned} g_{b0}(t) &= \mathcal{F}^{-1}\{G(\omega_b)\} \\ g_{b1}(t) &= \mathcal{F}^{-1}\{G(\omega_b)\exp(-j(\omega_0+\omega_b)t_1)\} + \mathbf{n}_1(t) \qquad (0.25) \\ g_{b2}(t) &= \mathcal{F}^{-1}\{G(\omega_b)\exp(-j(\omega_0+\omega_b)t_2)\} + \mathbf{n}_2(t) \qquad (0.26) \\ &\approx g_{b1}(t-\Delta t) \end{aligned}$$

where the spectrum $G(\omega_b)$ is now the complex baseband equivalent of $G(\omega)$ and the noise values $\mathbf{n}_{1,2}(t)$ are complex baseband equivalents of the real noise values $n_{1,2}(t)$ used in the real-only cross correlation. Thus



(a) (b)

Figure 0.2. (a) The unfiltered, complex-baseband envelopes of two sequential pulses with a differential delay between them as well as AWGN, (b) The cross-correlation where the upper trace is the amplitude and the lower trace the phase as a proportion of π. The peak value of the ampliude occurs at $10\mu s$ with a phase of 0.2π .

the baseband cross-correlation function is now complex and given by

$$\begin{aligned} \mathbf{cc}_{21}(t) &= \mathcal{F}^{-1}\{G(\omega_b)\exp(-j(\omega_0+\omega_b)\Delta t)\cdot G^*(\omega_b)\} + \mathbf{n}_1(t)\star\mathbf{n}_2(t) \\ &= \mathcal{F}^{-1}\{|G(\omega_b)|^2\exp(-j\omega_b\Delta t)\exp(-j\omega_0\Delta t)\} + \mathbf{n}_1(t)\star\mathbf{n}_2(t) \\ &\approx \mathbf{cc}_{11}(t-\Delta t)\cdot\exp(-j\omega_0\Delta t). \qquad (0.27) \end{aligned}$$

Again the cross-correlation of the two echoes is "almost" the same as the auto-correlation of $g_{b1}(t)$ delayed by Δt but this time with an extra carrier shift of $\exp(-j\omega_0\Delta t)$ to account for the complex demodulation process. By finding the peak amplitude of the complex cross- correlation of $|\mathbf{cc}_{21}(t)|$, shown in fig. 0.2(b), the *coarse time delay* $\widehat{\Delta t}$ can be estimated without any ambiguity. Since the peak is broad, the accuracy of the estimate of the coarse time delay $\widehat{\Delta t}$ from the amplitude is not as good as that estimated from the real-only signal in fig. 0.1(b) since we have ignored the carrier or phase effects.

A more accurate, but ambiguous, measurement of the time delay requires the phase of $\mathbf{cc}_{21}(t = \Delta t)$ (i.e., when $\mathbf{cc}_{11}(t - \Delta t)$ and $\mathbf{cc}_{21}(t)$ are a maximum). At this time, the phase is no more than $(\omega_0 \Delta t)$. Thus it is possible to estimate a *fine time delay* δt with the following computation

$$\widehat{\delta t} = 1/\omega_0 \arg[\mathbf{cc}_{21}(\Delta t)] \tag{0.28}$$

and this is commonly known as carrier-only correlation. Even better, if (but only if) we can ensure a symmetrical baseband so that $|G(\omega_b)| = |G(-\omega_b)|$, the phase is the same where ever $|\mathbf{cc}_{21}(t)|$ has a significant value. However, we still have to account for any 2π ambiguity. The favoured technique is to get the coarse time delay estimate from the peak of the amplitude in fig. 0.2(b), then unwrap the fine delay time estimate obtained from (0.28) until their difference is mimimised. this can be summarised as

$$\widehat{\widehat{\Delta t}} = \widehat{\delta t} + mT \tag{0.29}$$

where T is the period and m an integer selected to minimise $|\widehat{\widehat{\Delta t}} - \widehat{\Delta t}|$. This is known as the full or complete cross-correlation estimate of the time delay between the two complex baseband signals.

Unfortunately we almost never have a perfectly symmetrical baseband in a practical sonar system. This perfect symetry would only be true if the reflecting target was exactly on boresight and the transducers are incredibly well-calibrated so that the $|G(\omega_0 - \omega_b)| = |G(\omega_0 + \omega_b)|$ which in any reasonable situation would be highly unlikely and especially so for broadband systems. The consequence of the non-symmetrical baseband, where the negative frequencies would usually be slightly larger than the positive frequencies, is a linear phase shift on top of the phase offset we wish to measure. Consequently it is always necessary to measure the phase of $\mathbf{cc}_{21}(t = \Delta t)$ where the amplitude $|\mathbf{cc}_{21}(t)|$ has its maximum value. This linear phase shift through the complex cross- correlation measurement can also be clearly seen in real data e.g., Pinto,*et al*,2003 [4]

However before we discuss the accuracy of the time delay measurement, we need to include some form of noise supression. The signal either at modulated or at complex baseband is a band- limited function where the noise is white noise so it makes sense to filter the signals before cross-correlation. The filter that gives the most increase in SNR is the "matched filter"; in this case $G(\omega_b)$; the equivalent baseband spectrum of the transmitted signal $g_{r0}(t)$. In (0.16), (0.17), (0.25) and (0.26), there is the assumption that the power spectrum of the received echoes is the same as the transmitted signal which for real physical transducers is not

the case. Thus it is best to describe the transmitted signal and the two delayed echoes as

$$\begin{aligned} g_{b0}(t) &= \mathcal{F}^{-1}\{G(\omega_b)\} \\ g_{b1}(t) &= \mathcal{F}^{-1}\{\widehat{G}(\omega_b)\exp(-j(\omega_0+\omega_b)t_1)\} + \mathbf{n}_1(t) \qquad (0.30) \\ g_{b2}(t) &= \mathcal{F}^{-1}\{\widehat{G}(\omega_b)\exp(-j(\omega_0+\omega_b)t_2)\} + \mathbf{n}_2(t) \qquad (0.31) \\ &\approx g_{b1}(t-\Delta t) \end{aligned}$$

where $|\widehat{G}(\omega_b)| \approx |G(\omega_b)|$ but without the requirement that it must be symetrical about DC, i.e., $|\widehat{G}(\omega_b)| \neq |\widehat{G}(-\omega_b)|$. With that proviso, the matched filter response

$$\begin{aligned} g_{b3}(t) &= g_{b1}(t) \star g_{b0}(t) \\ &= \mathcal{F}^{-1}\{\widehat{G}(\omega_b)\, G^*(\omega_b)\exp(-j(\omega_0+\omega_b)t_1)\} + \mathbf{n}_1(t) \star g_{b0}(t) \\ g_{b4}(t) &= g_{b2}(t) \star g_{b0}(t) \\ &= \mathcal{F}^{-1}\{\widehat{G}(\omega_b)\, G^*(\omega_b)\exp(-j(\omega_0+\omega_b)t_1)\} + \mathbf{n}_2(t) \star g_{b0}(t) \end{aligned}$$

and given that the cross-correlation of signals 3 and 4 is

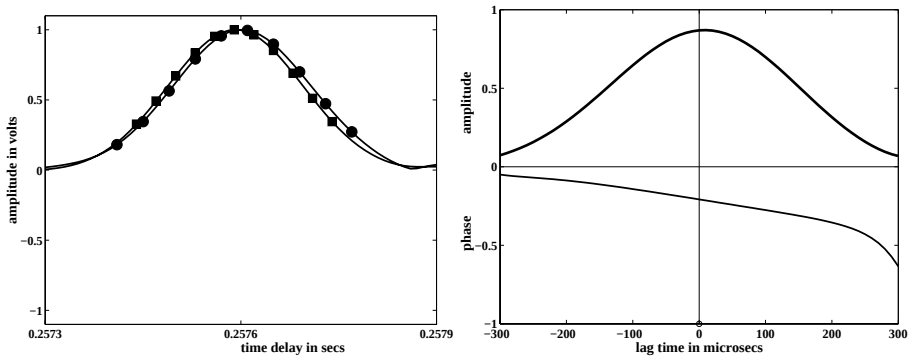
$$\mathbf{cc}_{34}(t) = g_{b3}(t) \star g_{b4}(t), \qquad (0.32)$$

we can estimate the coarse time delay $\widehat{\Delta t}$ from $|\mathbf{cc}_{34}(t)|$, the fine time delay $\widehat{\delta t}$ from the phase of $\mathbf{cc}_{34}(t=\Delta t)$ and then using (0.29) to determine if any unwrapping is required. Now we have our estimate of the time delay $\widehat{\widehat{\Delta t}}$ to the highest possible accuracy since the SNR has been maximised.

It is interesting to compare $\mathbf{cc}_{12}(t)$, fig.0.2(b), and $\mathbf{cc}_{34}(t)$, fig. 0.3(b). The effect of the matched filtering is to broaden the amplitude response, increase the degree of correlation and reduce the noise. It also decreases the amount of linear slope across the phase plot. This latter effect is the result of the centroids of $G_{b3}(\omega_b)$ and $G_{b4}(\omega_b)$ being closer to DC than $G_{b1}(\omega_b)$ and $G_{b2}(\omega_b)$.

But just how accurately can we measure the time delay (i.e., the lag) between to similar echoes? Fortunately here we can resort to the Cramer Rao Lower Bound (CRLB) calculations by Pinto *et al*, 2003 [4] who derive the following bound for the complex cross-correlation of the baseband echoes

$$\mathrm{CRLB}_{\mathrm{cc}} = \frac{1}{\sqrt{B_N \tau_p}} \frac{1}{\sqrt{1+\frac{B_N^2}{12\omega_0}}} \frac{1}{\omega_0} \sqrt{\frac{1}{SNR} + \frac{1}{2\,SNR^2}} \qquad (0.33)$$



(a) (b)

Figure 0.3. (a) The matched-filtered, complex-baseband envelopes of two sequential pulses with a differential delay between them as well as AWGN, (b) The cross-correlation where the upper trace is the ampitude and the lower trace the phase as a proportion of π. The peak value of the amplitude occurs at $10\mu s$ with a phase of 0.2π .

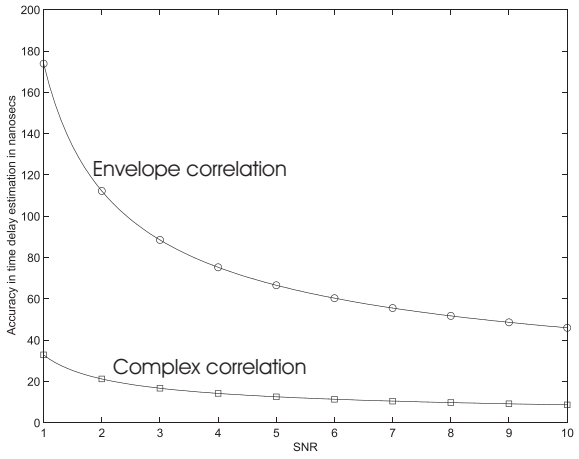


Figure 0.4. The Cramer Rao Lower Bound for full and envelope- only cross-correlation

where B_N is the noise bandwidth, τ_p the pulse (ping) repetition period and ω_0 the carrier frequency. This should be compared to the CRLB for the envelope-only cross correlation given by

$$\mathrm{CRLB}_{\mathrm{env}} = \frac{1}{\sqrt{B_N \tau_p}} \frac{2\sqrt{3}}{B_N} \sqrt{\frac{1}{SNR} + \frac{1}{2\,SNR^2}} \tag{0.34}$$

To see how these two bounds compare, fig 0.4 shows how the bounds decrease with increasing SNR when we have a noise bandwith of 20kHz, a ping repetition period of 0.3s and a carrier frequency of 30kHz.

4. Summary

Measuring relative time delays accurately is a common requirement in many signal processing and filtering operations. This is mostly done using correlation techniques. With sampled signals this can be done at the Nyquist sample rate in which case the signals are all positive real or more commonly at complex baseband in order to minimise the computational requirements.

The actual mathematical oeration can be carried out in the time domain or more efficiently in the frequency domain using the convolution/multiplication equivalents of the Fourier transform.

References

[1] C. E. Cook and M. Bernfeld. *Radar signals: an introduction to theory and application*. Academic Press Inc., 1967.

[2] J. C. Curlander and R. N. McDonough. *Synthetic Aperture Radar: systems and signal processing*. John Wiley & Sons, Inc., 1991.

[3] S. Haykin. *Communication Systems*. John Wiley and Sons, Inc, 3rd edition, 1994.

[4] M. A. Pinto, A. Bellettini, R. Hollett, and S. Chapman. Very long baseline interferometry for real and synthetic aperture sonar. *IEEE Journal of Oceanic Engineering*, 2003. In submission. Note: previously titled 'Bathymetric imaging with wideband interferometric synthetic aperture sonar'.

[5] A. W. Rihaczek. *Principles of high resolution radar*. McGraw Hill, Inc, 1969.