Chapter 8

Correlator I. Basics

D. Anish Roshi

8.1 Introduction

A radio interferometer measures the mutual coherence function of the electric field due to a given source brightness distribution in the sky. The antennas of the interferometer convert the electric field into voltages. The mutual coherence function is measured by cross correlating the voltages from each pair of antennas. The measured cross correlation function is also called *Visibility*. In general it is required to measure the visibility for different frequencies (spectral visibility) to get spectral information for the astronomical source. The electronic device used to measure the spectral visibility is called a *spectral correlator*. These devices are implemented using digital techniques. Digital techniques are far superior to analog techniques as far as stability and repeatability is concerned.

The first of these two chapters on correlators covers some aspects of digital signal processing used in digital correlators. Details of the hardware implementation of the GMRT spectral correlator are presented in the next lecture.

8.2 Digitization

The signals[1] at the output of the antenna/receiver system are analog voltages. Measurements using digital techniques require these voltages to be sampled and quantized.

8.2.1 Sampling

A band limited signal $s(t)$ with bandwidth $\Delta\nu$ can be uniquely represented by a time series obtained by periodically sampling $s(t)$ at a frequency f_s (the sampling frequency) which is greater than a critical frequency $2\Delta\nu$ (Shannon 1949). The signal is said to be 'Nyquist sampled' if the sampling frequency is exactly equal to the critical frequency $2\Delta\nu$.

The spectrum of signals sampled at a frequency $< 2\ \Delta\nu$ (i.e. under sampled) is distorted. Therefore the time series thus obtained is not a true representation of the band limited signal. The spectral distortion caused by under sampling is called *aliasing*.

[1]For all the analysis presented here we assume that radio astronomy signals are stationary and ergodic stochastic processes with a gaussian probability distribution. We also assume that the signals have zero mean.

8.2.2 Quantization

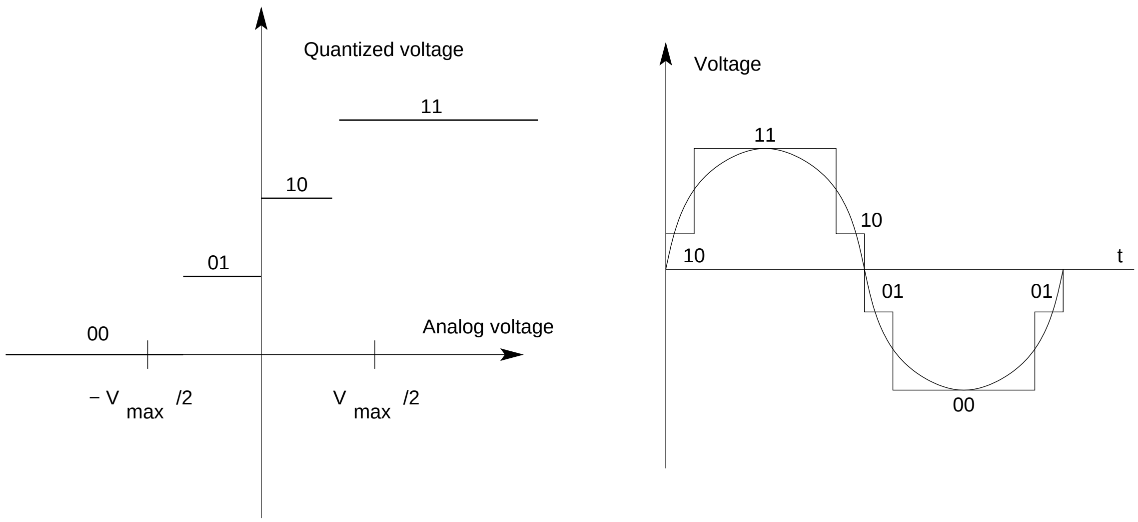

Figure 8.1: Transfer function of a two bit four level quantizer. The *binary* numbers corresponding to the quantized voltage range from 00 to 11. Quantization of a sine wave with such a quantizer is also shown.

The amplitude of the sampled signal is a continuous value. Digital systems represent values using a finite number of bits. Hence the amplitude has to be approximated and expressed with these finite number of bits. This processes is called *quantization*. The quantized values are integer multiple of a quantity q called the *quantization step*. An example of two bit (or equivalently four level) quantization is shown in Fig. 8.1. For the quantizer $q = V_{max}/2^2$, where V_{max} is the maximum voltage (peak-to-peak) that can be expressed within an error of $\pm q/2$.

Quantization distorts the sampled signal affecting both the amplitude and spectrum of the signal. This is evident from Fig. 8.1 for the case of a two bit four level quantized sine wave. The amplitude distortion can be expressed in terms of an error function $e(t) = s(t) - s_q(t)$, which is also called the *quantization noise*. Here $s_q(t)$ is the output of the quantizer. The variance of quantization noise under certain restricted conditions (such as uniform quantization) is $q^2/12$. The spectrum of quantization noise extends beyond the bandwidth $\Delta\nu$ of $s(t)$ (see Fig. 8.2). Sampling at the Nyquist rate ($2\Delta\nu$) therefore aliases the power of the quantization noise outside $\Delta\nu$ back into the spectral band of $s(t)$. For radio astronomy signals, the spectral density of the quantization noise within $\Delta\nu$ can be considered uniform and is $\sim q^2/12\Delta\nu$ (assuming uniform quantization). Reduction in quantization noise is hence possible by oversampling $s(t)$ (i.e. $f_s > 2\Delta\nu$) since it reduces the aliased power. For example, the signal to noise ratio of a digital measurement of the correlation function of s(t) (see Section 8.5) using a Nyquist sampling and a two bit four level quantizer is 88% of the signal to noise ratio obtained by doing analog correlation for Nyquist sampling and 94% if one were to sample at twice the Nyquist rate.

The largest value that can be expressed by a quantizer is determined by the number of bits (M) used for quantization. This value is $2^M - 1$ for binary representation. The finite number of bits puts an upper bound on the amplitude of input voltage that can be expressed within an error $\pm q/2$. Signals with amplitude above the maximum value will be 'clipped', thus producing further distortion. This distortion is minimum if the

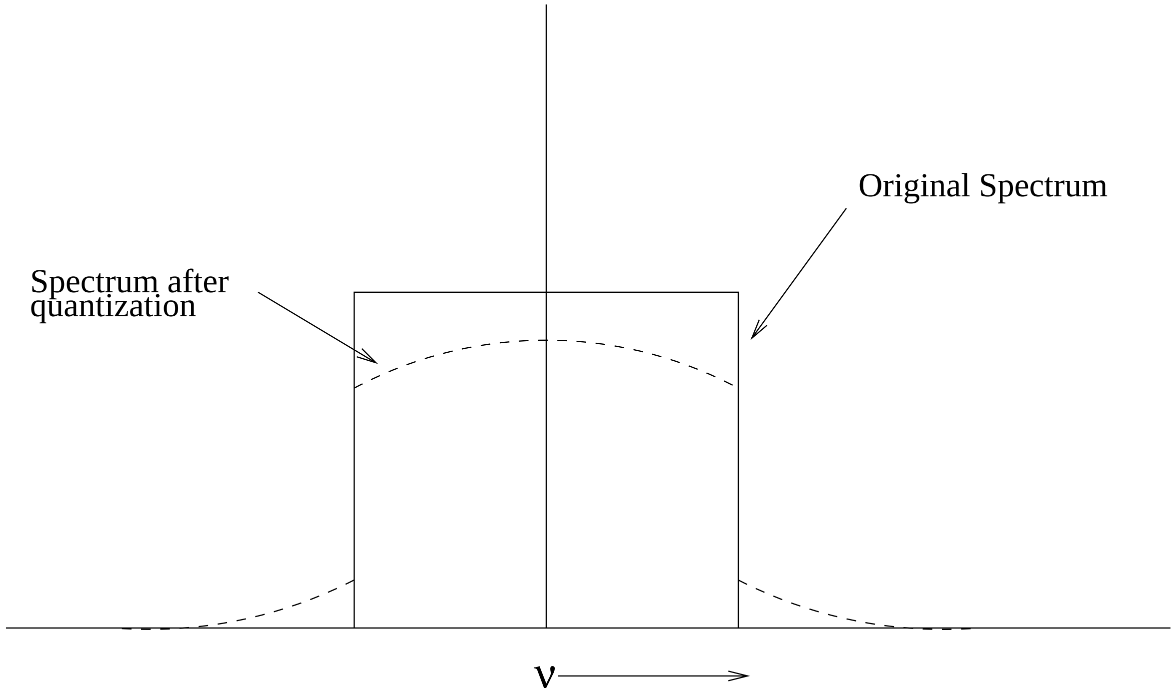

Figure 8.2: Power spectrum of band limited gaussian noise after one bit quantization. The spectrum of the original analog signal is shown with a solid line, while that of the quantized signal is shown with a dotted line.

probability of amplitude of the signal exceeding $+V_{max}/2$ and $-V_{max}/2$ is less than 10^{-5}. For a signal with a gaussian amplitude distribution this means that $V_{max} = 4.42\sigma$, σ being the standard deviation of $s(t)$.

8.2.3 Dynamic Range

As described above, the quantizer degrades the signal if its (peak-to-peak) amplitude is above an upper bound V_{max}. The minimum change in signal amplitude that can be expressed is limited by the quantization step q. Thus a given quantizer operates over a limited range of input voltage amplitude called its *dynamic range*. The Dynamic range of a quantizer is usually defined by the ratio of the power of a sinusoidal signal with peak-to-peak amplitude = V_{max} to the variance of the quantization noise. For an ideal quantizer with uniform quantization the dynamic range is $\frac{3}{2}2^{2M}$. Thus the dynamic range is larger if the number of bits used for quantization is larger.

8.3 Discrete Fourier Transform

The Fourier Transform (FT) of a signal $s(t)$ is defined as

$$S(w) = \int_{-\infty}^{+\infty} s(t)e^{-j\omega t}\mathrm{d}t \tag{8.3.1}$$

Discrete Fourier Transform (DFT) is an operation to evaluate the FT of the sampled signal $s(n)$ ($\equiv s(n\frac{1}{f_s})$) with a finite number of samples (say N). It is defined as

$$S(k) = \sum_{n=0}^{N-1} s(n)e^{-j2\pi nk/N}; \quad 0 \leq k \leq N-1 \tag{8.3.2}$$

The relationship between FT and DFT and some properties of DFT are discussed here.

Consider a time series $s(n)$, which is obtained by sampling a continuous band limited signal $s(t)$ at a rate f_s (see Fig. 8.3). The sampling function is a train of delta function $\text{III}(t)$. The length of the series is restricted to N samples by multiplying with a rectangular window function $\Pi(t)$. The modification of the signal $s(t)$ due to these operations and the corresponding changes in the spectrum are shown in Fig. 8.3. The spectral modifications can be understood from the properties of Fourier transforms. The FT of the time series can now be written as a summation (assuming N is even)

$$\begin{aligned} S(\omega) &= \int_{-\infty}^{+\infty} s(t) \sum_{n=-N/2}^{N/2-1} \delta(t - \frac{n}{f_s})e^{-j\omega t}\mathrm{d}t \\ &= \sum_{n=-N/2}^{N/2-1} s(\frac{n}{f_s})e^{-\frac{j\omega n}{f_s}} \end{aligned} \tag{8.3.3}$$

What remains is to quantize the frequency variable. For this the frequency domain is sampled such that there is no *aliasing in the time domain* (see Fig. 8.3). This is satisfied if $\Delta\omega$ = $2\pi f_s/N$. Thus Eq. 8.3.3 can be written as

$$S(k\Delta\omega) = \sum_{n=-N/2}^{N/2-1} s(\frac{n}{f_s})e^{-\frac{jk\Delta\omega n}{f_s}} \tag{8.3.4}$$

Using the relation $\Delta\omega/f_s = 2\pi/N$ and writing the variables as discrete indices we get the DFT equation. The cyclic nature of DFT (see below) allows n and k to range from 0 to $N-1$ instead of $-N/2$ to $N/2-1$.

Some properties that require attention are:

1. The spectral values computed for $N/2 \geq k \geq 3N/2-1$ are identical to those for $k = -N/2$ to $N/2-1$. In fact the computed values have a periodicity equal to $N\Delta\omega$ which makes the DFT cyclic in nature. This periodicity is a consequence of the sampling done in the time and frequency domain (see Fig. 8.3).

2. The sampling interval of the frequency variable $\Delta\omega$ (= $2\pi f_s/N$) is inversely proportional to the total number of samples used in the DFT. This is discussed further in Section 8.3.1.

There are several algorithms developed to reduce the number of operations in the DFT computation, which are called Fast Fourier Transform (FFT) algorithms. These algorithms reduce the time required for the computation of the DFT from $O(N^2)$ to $O(N\log(N))$. The FFT implementation used in the GMRT correlator uses Radix 4 and Radix 2 algorithms.

In the digital implementation of FFTs the quantization of the coefficients $e^{-j2\pi nk/N}$ degrades the signal to noise ratio the of spectrum. This degradation is in addition to the quantization noise introduced by the quantizer. Thus the dynamic range reduces further due to coefficient quantization. Coefficient quantization can also produce systematics in the computed spectrum. This effect also depends on the statistics of the input signal, and in general can be reduced only by using a larger number of bits for coefficient representation.

8.3.1 Filtering and Windowing

The Fourier transform of a signal $s(t)$ is a decomposition into frequency or spectral components. The DFT also performs a spectral decomposition but with a finite *spectral resolution*. The spectrum of a signal $s(t)$ obtained using a DFT operation is the convolution of the true spectrum of the signal $S(f)$ convolved by the FT $W(f)$ of the window function, and sampled at discrete frequencies. Thus a DFT is equivalent to a filter bank with filters spaced at $\Delta\omega$ in frequency. The response of each filter is the Fourier transform of the *window function* used to restrict the number of samples to N. For example, in the above analysis (see Section 8.3) the response of each 'filter' is the *sinc* function, (which is the FT of the rectangular window $\Pi(t)$). The spectral resolution (defined as the full width at half maximum (FWHM) of the filter response) of the sinc function is $\frac{1.21\Delta\omega}{2\pi}$. Different window functions $w(n)$ give different 'filter' responses, i.e. for

$$S(k) = \sum_{n=0}^{N-1} w(n)s(n)e^{-j2\pi nk/N} \tag{8.3.5}$$

the Hanning window

$$\begin{aligned} w(n) &= 0.5(1+\cos(2\pi n/N)) \text{ for } -N/2 \le n \le N/2-1 \\ &= 0 \text{ elsewhere} \end{aligned} \tag{8.3.6}$$

has a spectral resolution $\frac{2\Delta\omega}{2\pi}$. Side lobe reduction and resolution are the two principal considerations in choosing a given window function (or equivalently a given filter response). The rectangular window (i.e. sinc response function) has high resolution but a peak sidelobe of 22% while the Hanning window has poorer resolution but peak sidelobe level of only 2.6%.

8.4 Digital Delay

In interferometry the geometric delay suffered by a signal (see Chapter 4) has to be compensated before correlation is done. In an analog system this can be achieved by adding or removing cables from the signal path. An equivalent method in digital processing is to take sampled data that are offset in time. Mathematically, $s(n-m)$ is the sample delayed by $m \times 1/f_s$ with respect to $s(n)$ (where f_s is the sampling frequency). In such an implementation of delay it is obvious that the delay can be corrected only to the nearest integral multiple of $1/f_s$.

A delay less than $1/f_s$ (called *fractional delay*) can also be achieved digitally. A delay τ introduced in the path of a narrow band signal with angular frequency ω produces a phase $\phi = \omega\tau$. Thus, for a broad band signal, the delay introduces a phase gradient across the spectrum. The slope of the phase gradient is equal to the delay or $\tau = \frac{\mathrm{d}\phi}{\mathrm{d}\omega}$. This means that introducing a phase gradient in the FT of $s(t)$ is equivalent to introducing a delay is $s(t)$. Small enough phase gradients can be applied to realize a delay $< 1/f_s$. In the GMRT correlator, residual delays $\tau < 1/f_s$ is compensated using this method. This correction is called the Fractional Sampling Time Correction or FSTC.

8.5 Discrete Correlation and the Power Spectral Density

The cross correlation of two signals $s_1(t)$ and $s_2(t)$ is given by

$$R_c(\tau) = \; < s_1(t)s_2(t+\tau) > \tag{8.5.7}$$

where τ is the time delay between the the two signals. In the above equation the angle bracket indicates averaging in time. For measuring $R_c(\tau)$ in practice an estimator is defined as

$$R(m) = \frac{1}{N}\sum_{n=0}^{N-1} s_1(n)s_2(n+m) \;\; 0 \le m \le M \tag{8.5.8}$$

where m denotes the number of samples by which $s_2(n)$ is delayed, M is the maximum delay ($M \ll N$). By definition $R(m)$ is a random variable. The expectation value of $R(m)$ converges to $R_c(\tau = \frac{m}{f_s})$ when $N \to \infty$. The autocorrelation of the time series $s_1(n)$ is also obtained using a similar equation as Eq. 8.5.8 by replacing $s_2(n+m)$ by $s_1(n+m)$.

The correlation function estimated from the quantized samples in general deviates from the measurements taken with infinite amplitude precision. The deviation depends on the true correlation value of the signals. The relationship between the two measurement can be expressed as

$$\hat{R}_c(m/f_s) = \mathrm{F}(\hat{R}(m)) \tag{8.5.9}$$

where $\hat{R}_c(m/f_s)$ and $\hat{R}(m)$ are the normalized correlation functions (normalized with zero lag correlation in the case of autocorrelation and with square root of zero lag autocorrelations of the signal $s_1(t)$ and $s_2(t)$ in the case of cross correlation) and F is a correction function. It can be shown that the correction function is monotonic (Van Vleck & Middelton 1966, Cooper 1970, Hagan & Farley 1973, Kogan 1998). For example, the functional dependence for a one-bit quantization (the 'Van Vleck Correction') is

$$\hat{R}_c(m/f_s) = \sin(\frac{\pi}{2}\hat{R}(m)) \tag{8.5.10}$$

Note that the correction function is non-linear and hence this correction should be applied before any further operation on the correlation function. If the number of bits used for quantization is large then over a large range of correlation values the correction function is approximately linear.

The power spectral density (PSD) of a stationary stochastic process is defined to be the FT of its auto-correlation function (the Wiener-Khinchin theorem). That is if $R_c(\tau) = \;< s(t)s(t-\tau) >$ then the PSD, $S_c(f)$ is

$$S_c(f) = \int_{-\infty}^{\infty} R_c(\tau)e^{-j2\pi f\tau}d\tau \tag{8.5.11}$$

From the properties of Fourier transforms we have

$$R_c(0) = \;< s(t)s(t) > \;=\; \int_{-\infty}^{\infty} S_c(f)df \tag{8.5.12}$$

i.e. the function $S_c(f)$ is a decomposition of the variance (i.e. 'power') of $s(t)$ into different frequency components.

For sampled signals, the PSD is estimated by the Fourier transform of the discrete auto-correlation function. In case the signal is also quantized before the correlation, then one has to apply a Van Vleck correction *prior* to taking the DFT. Exactly as before, this estimate of the PSD is related to the true PSD via convolution with the window function.

One could also imagine trying to determine the PSD of a function $s(t)$ in the following way. Take the DFTs of the sampled signal $s(n)$ for several periods of length N and average them together and use this as an estimate of the PSD. It can be shown that this process is exactly equivalent to taking the DFT of the discrete auto-correlation function.

The cross power spectrum of the two signals is defined as the FT of the cross correlation function and the estimator is defined in a similar manner to that of the auto-correlation case.

8.6 Further Reading

1. Cooper, B.F.C. 1970, Aust. J. Phys. 23, 521
2. Hagen, J.B., Farley, D.T. 1973, Radio Science, 8, 775
3. Kogan, L. 1998, Radio Science, 33, N5, p 1289
4. Thompson, R.A., Moran, J.M., Swenson, Jr. G.W., “Interferometry and Synthesis in Radio Astronomy”, Chapter 8, John Wiley & Sons, 1986.
5. Oppenheim, A.V. & Schafer, R.W., “Digital Signal Processing”, Prentice-Hall, Englewood Cliffs, New Jersey, 1975.
6. Rabiner L.R. & Gold B, “Theory and Application of DSP”
7. Shannon, C.E. 1949, Proc. IRE, 37, 10
8. Thompson, A.R. & D’Addario, L.R. in “Synthesis Imaging in Radio Astronomy”, R.A. Perley, F.R. Schwab, & A.H. Bridle, eds., ASP Conf. Series, vol. 6.
9. Van Vleck, J. H., Middelton, D. 1966, Proc IEEE, 54, 2

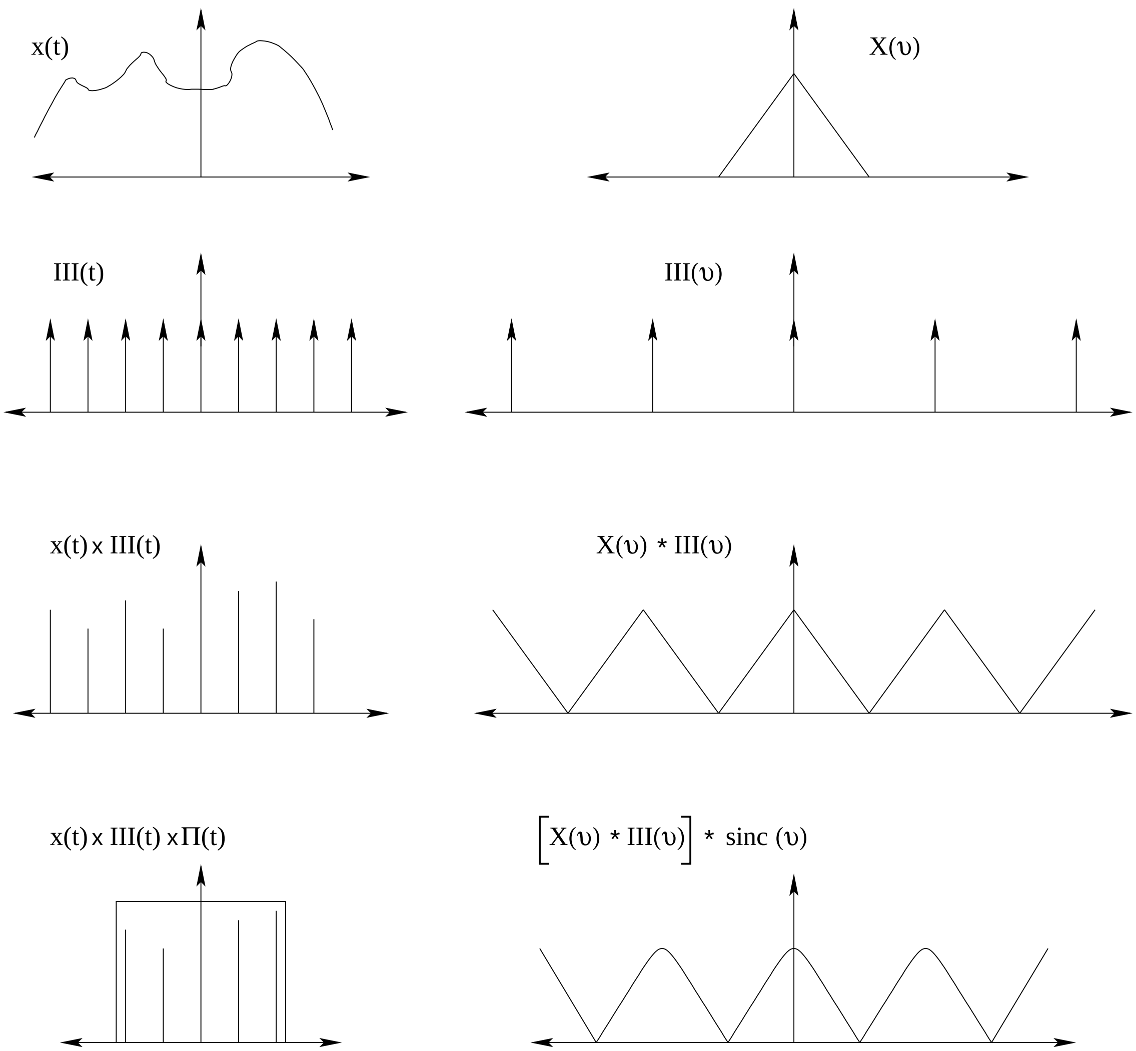

Figure 8.3: The relation between the continuous Fourier transform and the discrete Fourier transform. The panels on the left show the time domain signal and those on the right show the corresponding spectra.