

Chapter 6

Response of the Receiving System

This chapter is concerned with the response of the receiving system that accepts the signals from the antennas, amplifies and filters them, and measures the cross-correlations for the various antenna pairs. We show how the basic parameters of the system affect the output. Some of the effects were introduced in earlier chapters and are here presented in a more detailed development that leads to consideration of system design in Chaps. 7 and 8. At some point in the processing chain between the antenna and the correlator output, the form of the signals is changed from an analog voltage to a digital format, and the resulting data are thereafter processed by computer-type hardware. This does not affect the mathematical analysis of the processing and is not considered in this chapter. However, the digitization introduces a component of quantization noise, which is analyzed in Chap. 8.

6.1 Frequency Conversion, Fringe Rotation, and Complex Correlators

6.1.1 Frequency Conversion

With the exception of some systems operating below ~ 100 MHz, in most radio astronomy instruments, the frequencies of the signals received at the antennas are changed by mixing with a local oscillator (LO) signal. This feature, referred to as *frequency conversion* or (heterodyne frequency conversion), enables the major part of the signal processing to be performed at intermediate frequencies that are most appropriate for amplification, transmission, filtering, delaying, recording, and similar processes. For observations at frequencies up to roughly 50 GHz, the best sensitivity is generally obtained by using a low-noise amplifying stage before the frequency conversion.

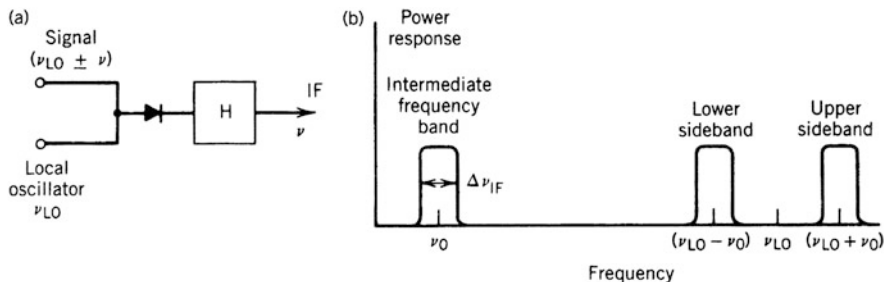


Fig. 6.1 Frequency conversion in a radio receiving system. (a) Simplified diagram of a mixer and a filter H that defines the intermediate-frequency (IF) band. The nonlinear element shown is a diode. (b) Signal spectrum showing upper and lower sidebands that are converted to the IF. Frequency ν_0 is the center of the IF band.

Frequency conversion takes place in a mixer in which the signal to be converted, plus an LO waveform, are applied to a circuit element with a nonlinear voltage–current response. This element may be a diode, as shown in Fig. 6.1a. The current i through the diode can be expressed as a power series in the applied voltage V :

$$i = a_0 + a_1 V + a_2 V^2 + a_3 V^3 + \dots \quad (6.1)$$

Now let V consist of the sum of an LO voltage $b_1 \cos(2\pi \nu_{LO} t + \theta_{LO})$ and a signal, of which one Fourier component is $b_2 \cos(2\pi \nu_s t + \phi_s)$. The second-order term in V then gives rise to a product in the mixer output of the form

$$\begin{aligned} & b_1 \cos(2\pi \nu_{LO} t + \theta_{LO}) \\ & \times b_2 \cos(2\pi \nu_s t + \phi_s) = \frac{1}{2} b_1 b_2 \cos [2\pi(\nu_s + \nu_{LO})t + \phi_s + \theta_{LO}] \\ & \quad + \frac{1}{2} b_1 b_2 \cos [2\pi(\nu_s - \nu_{LO})t + \phi_s - \theta_{LO}] . \end{aligned} \quad (6.2)$$

Thus, the current through the diode contains components at the sum and difference of ν_s and ν_{LO} . Other terms in Eq. (6.1) lead to other components, such as $3\nu_{LO} \pm \nu_s$, but the filter H shown in Fig. 6.1 passes only the wanted output components, and with proper design, unwanted combinations can be prevented from falling within the filter passband. Usually the signal voltage is much smaller than the LO voltage, so harmonics and intermodulation products (i.e., spurious signals that arise as a result of cross products of different frequency components within the input signal band) are small compared with the wanted terms containing ν_{LO} .

In most cases of frequency conversion, the signal frequency is being reduced, and the second term on the right side in Eq. (6.2) is the important one. The filter H then defines an intermediate-frequency (IF) band centered on ν_0 , as shown in Fig. 6.1b. Signals from within the bands centered on $\nu_{LO} - \nu_0$ and $\nu_{LO} + \nu_0$ are converted to the

IF band and admitted by the filter. These bands are known as the *lower* and *upper sidebands*, as shown, and if only a single sideband is wanted, the other can often be removed by a suitable filter inserted before the mixer. In some cases, both sidebands are accepted, resulting in a double-sideband response.

6.1.2 Response of a Single-Sideband System

Figure 6.2 shows a basic receiving system for two antennas, m and n , of a synthesis array. Here, we are interested in further effects of frequency conversion. The time difference τ_g between the arrival at the antennas of the signals from a radio source varies continuously as the Earth rotates and the antennas track the source across the sky. A variable instrumental delay τ_i is continuously adjusted to compensate for the

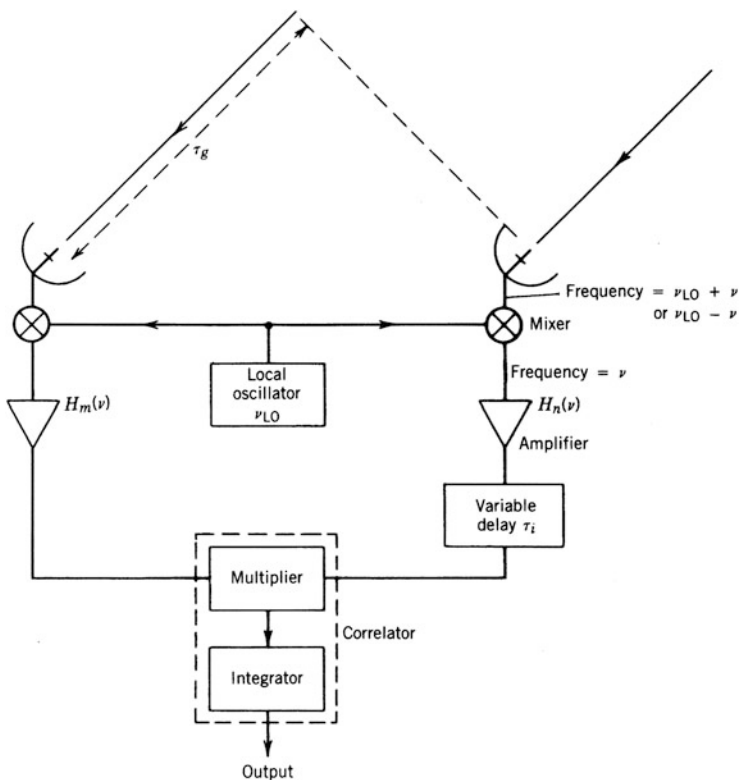


Fig. 6.2 Basic receiving system for two antennas of a synthesis array. The variable delay τ_i is continuously adjusted under computer control to compensate for the geometric delay τ_g . The frequency response functions $H_m(\nu)$ and $H_n(\nu)$ represent the overall bandpass characteristics of the amplifiers and filters in the signal channels.

geometric delay τ_g , so that the signals arrive simultaneously at the correlator. The receiving channels through which the signals pass contain amplifiers and filters, the overall amplitude (voltage) responses of which are $H_m(\nu)$ and $H_n(\nu)$ for antennas m and n . Here, ν represents a frequency at the correlator input; the corresponding frequency at the antenna is $\nu_{\text{LO}} \pm \nu$. The voltage waveforms that are processed by the receiving system result from cosmic noise and system noise; we consider the usual case in which these processes are approximately constant across the receiver passband. The spectra at the correlator inputs are thus determined mainly by the response of the receiving system. Let ϕ_m be the phase change in the signal path through antenna m resulting from τ_g and the LO phase, and let ϕ_n be the corresponding phase change in the signal for the path through antenna n , including τ_i . ϕ_m and ϕ_n , together with the instrumental phase resulting from the amplifiers and filters, represent the phases of the cosmic signal at the correlator inputs. Negative values of these parameters indicate phase lag (signal delay). The response to a source for which the visibility is $\mathcal{V}(u, \nu) = |\mathcal{V}|e^{j\phi_\nu}$ is most easily obtained by returning to Eq. (3.5) and replacing the phase difference $2\pi\mathbf{D}_\lambda \cdot \mathbf{s}_0$ by the general term $\phi_n - \phi_m$. Then the response at the correlator output resulting from a frequency band of width $d\nu$ can be written as

$$dr = \text{Re} \left\{ A_0 |\mathcal{V}| H_m(\nu) H_n^*(\nu) e^{j(\phi_n - \phi_m - \phi_\nu)} d\nu \right\}, \quad (6.3)$$

where ϕ_ν is the visibility phase. The response from the full system passband is

$$r = \text{Re} \left\{ A_0 |\mathcal{V}| \int_{-\infty}^{\infty} H_m(\nu) H_n^*(\nu) e^{j(\phi_n - \phi_m - \phi_\nu)} d\nu \right\}, \quad (6.4)$$

where we have included both positive and negative frequencies¹ in the integral and assumed that \mathcal{V} does not vary significantly over the observing bandwidth. Equation (6.4) represents the real part of the complex cross-correlation, and the way to obtain both the real and imaginary parts is explained later in this section.

6.1.3 Upper-Sideband Reception

For upper-sideband reception, a filter or amplifier at the receiver input selects frequencies in a band defined by the correlator input spectrum (frequency ν) plus ν_{LO} . In Fig. 6.2, the signal entering antenna m traverses the geometric delay τ_g at a frequency $\nu_{\text{LO}} + \nu$ and thus suffers a phase shift $2\pi(\nu_{\text{LO}} + \nu)\tau_g$. At the mixer, its

¹The negative frequencies have no physical meaning but arise as part of the mathematical representation of the frequency conversion.

phase is also decreased by the LO phase θ_m . Thus, we obtain

$$\phi_m(\nu) = -2\pi(\nu_{\text{LO}} + \nu)\tau_g - \theta_m. \quad (6.5)$$

The phase of the signal entering antenna n is decreased by the LO phase θ_n , and the signal then traverses the instrumental delay τ_i at a frequency ν , thus suffering a shift $2\pi\nu\tau_i$. The total phase shift for antenna n is

$$\phi_n(\nu) = -2\pi\nu\tau_i - \theta_n. \quad (6.6)$$

From Eqs. (6.4), (6.5), and (6.6), the correlator output is

$$r_u = \mathcal{R}e \left\{ A_0 |\mathcal{V}| e^{j[2\pi\nu_{\text{LO}}\tau_g + (\theta_m - \theta_n) - \phi_\nu]} \int_{-\infty}^{\infty} H_m(\nu) H_n^*(\nu) e^{j2\pi\nu\Delta\tau} d\nu \right\}. \quad (6.7)$$

The real part of the integral in Eq. (6.7) is one-half the Fourier transform of the (Hermitian) cross power spectrum $H_m(\nu)H_n^*(\nu)$ with respect to the delay compensation error, $\Delta\tau = \tau_g - \tau_i$, which introduces a linear phase slope across the band.² We assume that \mathcal{V} does not vary significantly over the observing bandwidth. For example, if the IF passbands are rectangular with center frequency ν_0 , width $\Delta\nu_{\text{IF}}$, and identical phase responses, then for positive frequencies,

$$|H_m(\nu)| = |H_n(\nu)| = \begin{cases} H_0, & |\nu - \nu_0| < \frac{\Delta\nu_{\text{IF}}}{2}, \\ 0, & |\nu - \nu_0| > \frac{\Delta\nu_{\text{IF}}}{2}. \end{cases} \quad (6.8)$$

Using the equality in Eq. (A3.6) of Appendix 3.1 for the Hermitian³ function $H_m H_n$, we can write

$$\begin{aligned} \int_{-\infty}^{\infty} H_m(\nu) H_n^*(\nu) e^{j2\pi\nu\Delta\tau} d\nu &= 2\mathcal{R}e \left\{ \int_{\nu_0 - (\Delta\nu_{\text{IF}}/2)}^{\nu_0 + (\Delta\nu_{\text{IF}}/2)} H_0^2 e^{j2\pi\nu\Delta\tau} d\nu \right\} \\ &= 2H_0^2 \Delta\nu_{\text{IF}} \left[\frac{\sin(\pi\Delta\nu_{\text{IF}}\Delta\tau)}{\pi\Delta\nu_{\text{IF}}\Delta\tau} \right] \cos 2\pi\nu_0\Delta\tau. \end{aligned} \quad (6.9)$$

²Here, we assume that the source is sufficiently close to the center of the field being imaged that the condition $\Delta\tau = 0$ maintains zero delay error. The effect of the variation of the delay error across a wider field of view is considered in Sect. 6.3.

³The term ‘‘Hermitian’’ indicates a function in which the real part is even and the imaginary part is odd.

In the general case, we define an instrumental gain factor $G_{mn} = |G_{mn}|e^{j\phi_G}$ as follows:

$$\begin{aligned} A_0 \int_{-\infty}^{\infty} H_m(\nu)H_n^*(\nu) e^{j2\pi\nu\Delta\tau} d\nu &= G_{mn}(\Delta\tau) e^{j2\pi\nu_0\Delta\tau} \\ &= |G_{mn}(\Delta\tau)|e^{j(2\pi\nu_0\Delta\tau+\phi_G)} , \end{aligned} \quad (6.10)$$

where the $\Delta\tau$ dependence in $|G_{mn}(\Delta\tau)|$ is the sinc function in Eq. (6.9). The phase ϕ_G results from the difference in the phase responses of the amplifiers and filters in the two channels. The LO phases θ_m and θ_n are not included within the general instrumental phase term ϕ_G because they enter into the upper and lower sidebands with different signs.

Substituting Eq. (6.10) into Eq. (6.7), we obtain for upper-sideband reception

$$r_u = |\mathcal{V}||G_{mn}(\Delta\tau)| \cos [2\pi(\nu_{LO}\tau_g + \nu_0\Delta\tau) + (\theta_m - \theta_n) - \phi_v + \phi_G] . \quad (6.11)$$

The term $2\pi\nu_{LO}\tau_g$ in the cosine function results in a quasi-sinusoidal oscillation as the source moves through the fringe pattern. The phase of this oscillation depends on the delay error $\Delta\tau$, the relative phases of the LO signals, the phase responses of the signal channels, and the phase of the visibility function. The frequency of the output oscillation $\nu_{LO}d\tau_g/dt$ is often referred to as the *natural fringe frequency*. The oscillations result because the signals traverse the delays τ_g and τ_i at different frequencies, that is, at the input radio frequency for τ_g and at the intermediate frequency for τ_i , and these two frequencies differ by ν_{LO} . Thus, even if these two delays are identical, they introduce different phase shifts, and they increase or decrease progressively as the Earth rotates.

6.1.4 Lower-Sideband Reception

Consider now the situation where the frequencies accepted from the antenna are those in the lower sideband, at ν_{LO} minus the correlator input frequencies. The phases are

$$\phi_m = 2\pi(\nu_{LO} - \nu)\tau_g + \theta_m \quad (6.12)$$

and

$$\phi_n = -2\pi\nu\tau_i + \theta_n . \quad (6.13)$$

The signs of these terms and of ϕ_v differ from those in the upper-sideband case because increasing the phase of the signal at the antenna here decreases the phase at

the correlator. The expression for the correlator output is

$$r_\ell = \mathcal{R}e \left\{ A_0 |\mathcal{V}| e^{-j[2\pi\nu_{\text{LO}}\tau_g + (\theta_m - \theta_n) - \phi_v]} \int_{-\infty}^{\infty} H_m(\nu) H_n^*(\nu) e^{j2\pi\Delta\tau} d\nu \right\}. \quad (6.14)$$

Proceeding as in the upper-sideband case, we obtain

$$r_\ell = |\mathcal{V}| |G_{mn}(\Delta\tau)| \cos [2\pi(\nu_{\text{LO}}\tau_g - \nu_0\Delta\tau) + (\theta_m - \theta_n) - \phi_v - \phi_G]. \quad (6.15)$$

6.1.5 Multiple Frequency Conversions

In an operational system, the signals may undergo several frequency conversions between the antennas and the correlators. A frequency conversion in which the output is at the lower sideband (i.e., the LO frequency minus the input frequency) results in a reversal of the signal spectrum in which frequencies at the high end at the input appear at the low end at the output, and vice versa. If there is no net reversal (that is, an even number of lower-sideband conversions), Eq. (6.11) applies, except that ν_{LO} must be replaced by a corresponding combination of LO frequencies. Similarly, the oscillator phase terms θ_m and θ_n are replaced by corresponding combinations of oscillator phases.

6.1.6 Delay Tracking and Fringe Rotation

Adjustment of the compensating delay τ_i of Fig. 6.2 is usually accomplished under computer control, the required delay being a function of the antenna positions and the position of the phase center of the field under observation. This can be achieved by designating one antenna of the array as the delay reference and adjusting the instrumental delays of other antennas so that, for an incoming wavefront from the phase reference direction, the signals intercepted by the different antennas all arrive at the correlator simultaneously.

To control the frequency of the sinusoidal fringe variations in the correlator output, a continuous phase change can be inserted into one of the LO signals. Equations (6.11) and (6.15) show that the fringe frequency can be reduced to zero by causing $\theta_m - \theta_n$ to vary at a rate that maintains constant, modulo 2π , the term $[2\pi\nu_{\text{LO}}\tau_g + (\theta_m - \theta_n)]$. This requires adding a frequency $2\pi\nu_{\text{LO}}(d\tau_g/dt)$ to θ_n or subtracting it from θ_m . Note that $d\tau_g/dt$ can be evaluated from Eq. (4.9) in which w , the third component of the interferometer baseline, is equal to $c\tau_g$ measured in wavelengths; for example, for an east–west antenna spacing of 1 km, the maximum value of $d\tau_g/dt$ is 2.42×10^{-10} , so the fringe frequencies are generally small compared with the radio frequencies involved. Reduction of the output frequency reduces the quantity of data to be processed, since each correlator output must

be sampled at least twice per cycle of the output frequency (the Nyquist rate) to preserve the information, as is discussed in Sect. 8.2.1. With antenna spacings required for angular resolution of milliarcsecond order, which occur in VLBI, the natural fringe frequency, $v_{\text{LO}} d\tau_g/dt$, can exceed 10 kHz. For an array with more than one antenna pair, it is possible to reduce each output frequency to the same fraction of its natural frequency, or to zero. Reduction to zero frequency (fringe stopping) is generally the preferred practice. Some special technique, such as the use of a complex correlator, described in the following section, is then required to extract the amplitude and phase of the output.

6.1.7 Simple and Complex Correlators

A method of measuring the amplitude and phase of the correlator output signal when the fringe frequency is reduced to zero is shown in Fig. 6.3. Two correlators are used, one of which has a quadrature phase shift network in one input. For signals of finite bandwidth, this phase shift is not equivalent to a delay. The phase shift can also be effected by feeding the signal into two separate mixers and converting it with two LOs in phase quadrature. The output of the second correlator can be represented by replacing $H_m(v)$ by $H_m(v)e^{-j\pi/2}$. From Eq. (6.10), the result is to add $-\pi/2$ to ϕ_G , and thus in Eq. (6.11) and Eq. (6.15), the cosine function is replaced by \pm sine. Another way of comparing the two correlator outputs in Fig. 6.3 is to note that the real output of the complex correlator, omitting constant factors, is

$$r_{\text{real}} = \text{Re} \left\{ \mathcal{V} \int_{-\infty}^{\infty} H_m(v) H_n^*(v) dv \right\} = \text{Re} \{ \mathcal{V} \} \int_{-\infty}^{\infty} H_m(v) H_n^*(v) dv, \quad (6.16)$$

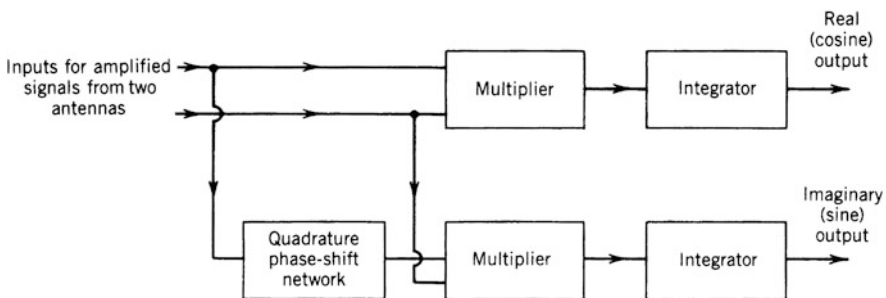


Fig. 6.3 Use of two correlators to measure the real and imaginary parts of the visibility. This system is called a complex correlator.

where the integral is real since $H_m(v)$ and $H_n^*(v)$ are Hermitian and thus $H_m(v)H_n^*(v)$ is Hermitian. The imaginary output of the correlator is proportional to

$$r_{\text{imag}} = \mathcal{Re} \left\{ \mathcal{V} \int_{-\infty}^{\infty} H_m(v)H_n^*(v)e^{-j\pi/2} dv \right\} = \mathcal{Im}\{\mathcal{V}\} \int_{-\infty}^{\infty} H_m(v)H_n^*(v) dv . \quad (6.17)$$

Thus, the two outputs respond to the real and imaginary parts of the visibility \mathcal{V} .

The combination of two correlators and the quadrature network is usually referred to as a *complex correlator*, and the two outputs as the cosine and sine, or real and imaginary, outputs. For continuum observations, the compensating delay is adjusted so that $\Delta\tau = 0$ and the fringe rotation maintains the condition $2\pi\nu_{\text{LO}}\tau_i + (\theta_m - \theta_n) = 0$. Thus, the cosine and sine outputs represent the real and imaginary parts of $G_{mm}\mathcal{V}(u, v)$. With the use of the complex correlator, the rotation of the Earth, which sweeps the fringe pattern across the source, is no longer a necessary feature in the measurement of visibility. An important feature of the complex correlator is that the noise fluctuations in the cosine and sine outputs are independent, as discussed in Sect. 6.2.2.

Spectral correlator systems, in which a number of correlators are used to measure the correlation as a function of time offset or “lag” [i.e., τ in Eq.(3.27)], are discussed in Sect. 8.8. The correlation as a function of τ measured using a correlator with a quadrature phase shift in one input is the Hilbert transform of the same quantity measured without the quadrature phase shift (Lo et al. 1984).

6.1.8 Response of a Double-Sideband System

A double-sideband (DSB) receiving system is one in which both the upper- and lower-sideband responses are accepted. From Eqs. (6.11) and (6.15), the output is

$$r_d = r_u + r_\ell = 2|\mathcal{V}||G_{mm}(\Delta\tau)| \cos(2\pi\nu_0\Delta\tau + \phi_G) \times \cos[2\pi\nu_{\text{LO}}\tau_g + (\theta_m - \theta_n) - \phi_v] . \quad (6.18)$$

There is a significant difference from the single-sideband (SSB) cases. The phase of the fringe-frequency term, which is the cosine function containing the term $2\pi\nu_{\text{LO}}\tau_g$, is no longer dependent on $\Delta\tau$ or ϕ_G , but instead these quantities appear in the term that controls the fringe amplitude:

$$|G_{mm}(\Delta\tau)| \cos(2\pi\nu_0\Delta\tau + \phi_G) . \quad (6.19)$$

If the delay τ_i is held constant, $\Delta\tau$ varies continuously, resulting in cosinusoidal modulation of the fringe oscillations through the cosine term in (6.19). Also, as

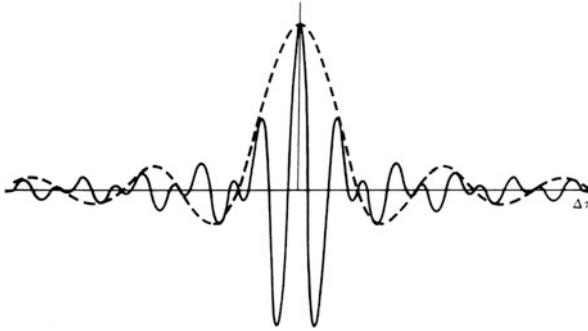


Fig. 6.4 Example of the variation of the fringe amplitude as a function of $\Delta\tau$ for a DSB system (solid line). In this case, the centers of the two sidebands are separated by three times the IF bandwidth, that is, $\nu_0 = 1.5\Delta\nu_{\text{IF}}$, and the IF response is rectangular. The broken line shows the equivalent function for an SSB system with the same IF response.

shown in Fig. 6.4, the cross-correlation (fringe amplitude) falls off more rapidly because of the cosine term in (6.19) than it does in the SSB case, in which it depends only on $G_{mn}(\Delta\tau)$. The required precision in matching the geometric and instrumental delays is correspondingly increased. The lack of dependence of the fringe phase on the phase response of the signal channel occurs because the latter has equal and opposite effects on the signals from the two sidebands.

The response of a DSB system with a complex correlator is given by Eq. (6.18) for the cosine output, and the sine output is obtained by replacing ϕ_G by $\phi_G - \pi/2$:

$$(r_d)_{\text{sine}} = 2|\mathcal{V}||G_{mn}(\Delta\tau)| \sin(2\pi\nu_0\Delta\tau + \phi_G) \times \cos[2\pi\nu_{\text{LO}}\tau_g + (\theta_m - \theta_n) - \phi_v] . \quad (6.20)$$

If the term $2\pi\nu_0\Delta\tau + \phi_G$ is adjusted to maximize either the real output [Eq. (6.18)] or the imaginary output [Eq. (6.20)], the other will be zero. Thus, for continuum observations in which the signal is of equal strength in both sidebands, the complex correlator offers no increase in sensitivity. However, it can be useful for observations in the sideband-separation mode described later.

To help visualize the difference between SSB and DSB interferometer systems, Fig. 6.5 illustrates the correlator outputs in the complex plane. The SSB case is shown in Fig. 6.5a. The output of the complex correlator is represented by the vector \mathbf{r} . If the fringes are not stopped, the vector \mathbf{r} rotates through 2π each time the geometric delay τ_g changes by one wavelength (that is, one wavelength at the LO frequency if the instrumental delay is tracking the geometric delay). The projections of the radial vector on the real and imaginary axes indicate the real and imaginary outputs of the complex correlator, which are two fringe-frequency sinusoids in phase quadrature. If the fringes are stopped, \mathbf{r} remains fixed in position angle. Figure 6.5b represents the DSB case. Vectors \mathbf{r}_u and \mathbf{r}_ℓ represent the output components from

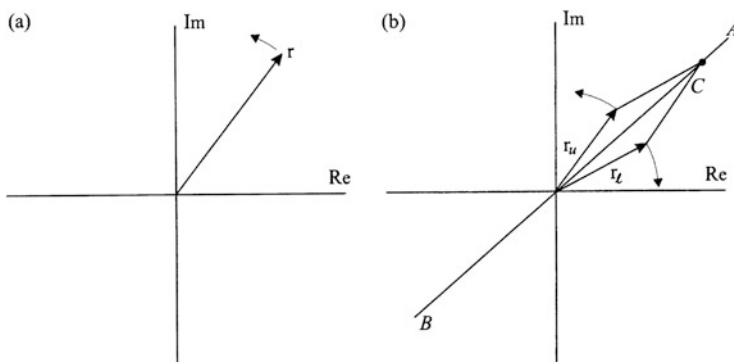


Fig. 6.5 Representation in the complex plane of the output of a correlator with (a) an SSB and (b) a DSB receiving system. The point C in (b) represents the sum of the upper- and lower-sideband outputs of the correlator.

the upper and lower sidebands. Here the variation of τ_g causes \mathbf{r}_u and \mathbf{r}_l to rotate in opposite directions. To verify this statement, note that the real parts of the correlator output are given in Eqs. (6.11) and (6.15), and the corresponding imaginary parts are obtained by replacing ϕ_G by $\phi_G - \pi/2$. Then with $(\theta_m - \theta_n) = 0$ (no fringe rotation), consider the effect of a small change in τ_g .

The contra-rotating vectors representing the two sidebands at the correlator output coincide at an angle determined by instrumental phase, which we represent by the line AB in Fig. 6.5b. Thus, the vector sum oscillates along this line, and the fringe-frequency sinusoids at the real and imaginary outputs of the correlator are in phase. Now suppose we adjust the phase term $(2\pi\nu_0\Delta\tau + \phi_G)$ in Eq. (6.18) to maximize the fringe amplitude at the real output. This action has the effect of rotating the line AB to coincide with the real axis. The imaginary output of the complex correlator then contains no signal, only noise. From Eq. (6.18), it can be seen that the visibility phase ϕ_v is represented by the phase of the vector that oscillates in amplitude along the real axis. The phase can be recovered by letting the fringes run and fitting a sinusoid to the waveform at the real output. If the fringes are stopped, it is possible to determine the amplitude and phase of the fringes by $\pi/2$ switching of the LO phase at one antenna. In Eq. (6.18), this phase switch action can be represented by $\theta_m \rightarrow (\theta_m - \pi/2)$, which results in a change of the second cosine function to a sine, thus enabling the argument in square brackets to be determined. However, in such a case, the data representing the cosine and sine components of the output are not measured simultaneously, so the effective data-averaging time is half that for the SSB, complex-correlator case. In Fig. 6.5b, a $\pi/2$ switch of the LO phase results in a rotation of \mathbf{r}_u and \mathbf{r}_l by $\pi/2$ in opposite directions, so the vector sum of the two sideband outputs remains on the line AB . Relative sensitivities of different systems are discussed in Sect. 6.2.5.

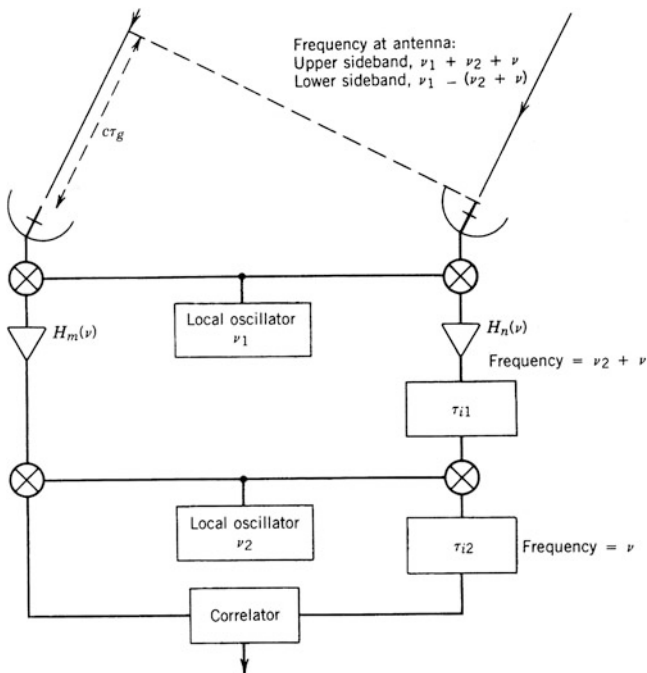


Fig. 6.6 Receiving system for two antennas that incorporates two frequency conversions, the first being DSB and the second upper-sideband. Two compensating delays, τ_{i1} and τ_{i2} , are included so that in deriving the response for a DSB system, the effect of the position of the delay relative to the first mixer can be investigated. In practice, only one compensating delay is required. The overall frequency responses H_m and H_n are specified as functions of ν , which is the corresponding frequency at the correlator input.

6.1.9 Double-Sideband System with Multiple Frequency Conversions

The response with multiple frequency conversions is more complicated for a DSB interferometer than for an SSB one and is illustrated by considering the system in Fig. 6.6. Note that for the case in which the IF signal undergoes a number of SSB frequency conversions after the first mixer, the second mixer of each antenna in Fig. 6.6 can be considered to represent several mixers in series, and ν_2 is equal to the sum of the LO frequencies with appropriate signs to take account of upper- or lower-sideband conversions. The signal phase terms are determined by considerations similar to those described in the derivation of Eqs. (6.5) and (6.6). Thus, we obtain

$$\phi_m = \mp 2\pi(\nu_1 \pm \nu_2 \pm \nu)\tau_g \mp \theta_{m1} - \theta_{m2} \tag{6.21}$$

and

$$\phi_n = -2\pi(\nu_2 + \nu)\tau_{i1} - 2\pi\nu\tau_{i2} \mp \theta_{n1} - \theta_{n2} , \quad (6.22)$$

where the upper signs correspond to upper-sideband conversion at both the first and second mixers for each antenna, and the lower signs to lower-sideband conversion at the first mixer for each antenna and upper-sideband conversion at the second. We then proceed as in the previous examples; that is, use Eqs. (6.21) and (6.22) to substitute for ϕ_m and ϕ_n in Eq. (6.4), separate out the integral of $H_m H_n^*$ with respect to frequency, ν , as in Eq. (6.7), and substitute for the integral using Eq. (6.10). The results are

$$\begin{aligned} r_u = & |\mathcal{V}||G_{mn}(\Delta\tau)| \cos[2\pi\nu_1\tau_g + 2\pi\nu_2(\tau_g - \tau_{i1}) + 2\pi\nu_0\Delta\tau \\ & + (\theta_{m1} - \theta_{n1}) + (\theta_{m2} - \theta_{n2}) - \phi_v + \phi_G] \end{aligned} \quad (6.23)$$

and

$$\begin{aligned} r_\ell = & |\mathcal{V}||G_{mn}(\Delta\tau)| \cos[2\pi\nu_1\tau_g - 2\pi\nu_2(\tau_g - \tau_{i1}) - 2\pi\nu_0\Delta\tau \\ & + (\theta_{m1} - \theta_{n1}) - (\theta_{m2} - \theta_{n2}) - \phi_v - \phi_G] . \end{aligned} \quad (6.24)$$

The DSB response is

$$\begin{aligned} r_d = & r_u + r_\ell \\ = & 2|\mathcal{V}||G_{mn}(\Delta\tau)| \cos\{2\pi[\nu_2(\tau_{i1} - \tau_g) - \nu_0\Delta\tau] - (\theta_{m2} - \theta_{n2}) - \phi_G\} \\ & \times \cos[\nu_1\tau_g + (\theta_{m1} - \theta_{n1}) - \phi_v] , \end{aligned} \quad (6.25)$$

where $\Delta\tau = \tau_g - \tau_{i1} - \tau_{i2}$. Note that the phase of the output fringe pattern, given by the second cosine term, depends only on the phase of the first LO. Thus, in the implementation of fringe rotation, the phase shift must be applied to this oscillator. The first cosine term in Eq. (6.25) affects the fringe amplitude, and two cases should be considered:

1. The delay τ_{i1} , at the IF immediately following the DSB mixer, is used as the compensating delay, and $\tau_{i2} = 0$. Then in the first cosine function in Eq. (6.25), $\tau_{i1} - \tau_g \simeq 0$, and ϕ_G should be small if the frequency responses of the two channels are similar. It is necessary only to equalize θ_{m2} and θ_{n2} to maximize the amplitude of the fringe-frequency term. This is similar to the single conversion case in Eq. (6.18).
2. The delay τ_{i2} , located after the last mixer, is used as the compensating delay, and $\tau_{i1} = 0$. (This is the case in any array in which the compensating delays are implemented digitally, which includes almost all currently operational systems.)

Then a continuously varying phase shift is required in θ_{m2} or θ_{n2} of Eq. (6.25) to keep the value of the first cosine function close to unity as τ_g varies. This phase shift does not affect the phase of the output fringe oscillations, only the amplitude [see, e.g., Wright et al. (1973)].

6.1.10 Fringe Stopping in a Double-Sideband System

Consider two antennas of an array as shown in Fig. 6.6 and the case in which the instrumental delay that compensates for τ_g is the one immediately preceding the correlator, so that $\tau_{i1} = 0$. One can think of interferometer fringes as being caused by a Doppler shift in the signal at one antenna, which results in a beat frequency when the signals are combined in the correlator. Suppose that the geometric delay, τ_g , in the signal path to antenna m (on the left side of the diagram) is increasing with time, that is, antenna m is moving away from the source relative to antenna n . Then a signal at frequency ν_{RF} at the wavefront from a source appears at frequency $\nu_{\text{RF}}(1 - d\tau_g/dt)$ when received at antenna m . If the signal is in the upper sideband, its frequency at the correlator input will be

$$\nu_{\text{RF}} \left(1 - \frac{d\tau_g}{dt} \right) - \nu_1 - \nu_2 . \quad (6.26)$$

To stop the fringes, we need to apply a corresponding decrease to the frequency of the signal from antenna n so that the signals arrive at the correlator at the same frequency. To do this, we increase the frequencies of the two LOs for antenna n by the factor $(1 + d\tau_g/dt)$. Note that this is equivalent to adding $2\pi(d\tau_g/dt)\nu_1$ to θ_{n1} and $2\pi(d\tau_g/dt)\nu_2$ to θ_{n2} , which are the rates of change of the oscillator phases required to maintain each of the two cosine functions in Eq. (6.25) at constant value. The corresponding signal from antenna n traverses the delay τ_{i2} at a frequency $\nu_{\text{RF}} - (\nu_1 + \nu_2)(1 + d\tau_g/dt)$, and since the delay is continuously adjusted to equal τ_g , the signal suffers a reduction in frequency by a factor $(1 - d\tau_g/dt)$. Thus, at the correlator input, the frequency of the antenna- n signal is

$$\left[\nu_{\text{RF}} - (\nu_1 + \nu_2) \left(1 + \frac{d\tau_g}{dt} \right) \right] \left(1 - \frac{d\tau_g}{dt} \right) , \quad (6.27)$$

which is equal to (6.26) when second-order terms in $d\tau_g/dt$ are neglected. (Recall that for, e.g., a 1-km baseline, the highest possible value of $d\tau_g/dt$ is 2.42×10^{-10} .) For the lower sideband, (6.26) and (6.27) apply if the signs of both ν_{RF} and ν_1 are reversed, and again the frequencies at the correlator input are equal. Thus, the overall effect is that the fringes are stopped for *both* sidebands.

6.1.11 *Relative Advantages of Double- and Single-Sideband Systems*

The principal reason for using DSB reception in interferometry is that in certain cases, the lowest receiver noise temperatures are obtained by using input stages that are inherently DSB devices. As frequency increases above ~ 100 GHz, it becomes increasingly difficult to make low-noise amplifiers, and receiving systems often use a mixer of the superconductor–insulator–superconductor (SIS) type [see, e.g., Tucker and Feldman (1985)] as the input stage followed by a low-noise IF amplifier. Both the mixer and the IF amplifier are cryogenically cooled to obtain superconductivity in the mixer and to minimize the amplifier noise. If a filter is placed between the antenna and the mixer to cut out one sideband, the received signal power is halved, but there is no reduction in the receiver noise generated in the mixer and IF stages. Thus, the signal-to-noise ratio (SNR) in the IF stages is reduced, and in this case, the best continuum sensitivity may be obtained if both sidebands are retained. As a historical note, DSB systems were used at centimeter wavelengths during the 1960s and early 1970s [see, e.g., Read (1961)], sometimes with a degenerate type of parametric amplifier as the low-noise input stage. These amplifiers were inherently DSB devices, and their use in interferometry is discussed by Vander Vorst and Colvin (1966).

DSB systems have a number of disadvantages. Increased accuracy of delay setting is required, frequency and phase adjustment on more than one LO is likely to be required, interpretation of spectral line data is complicated if there are lines in both sidebands, and the width of the interference-free spectrum required is doubled. Also, the smearing effect of a finite bandwidth, to be discussed in Sect. 6.3, is increased. These problems have stimulated the development of schemes by which the responses for upper and lower sidebands can be separated.

6.1.12 *Sideband Separation*

To illustrate the method by which the responses for the two sidebands can be separated at the correlator output of a DSB receiving system, we examine the sum of the upper- and lower-sideband responses from Eqs. (6.11) and (6.15). This is

$$r_d = r_u + r_\ell = |\mathcal{V}||G_{mn}(\Delta\tau)| \left\{ \cos [2\pi(\nu_{\text{LO}}\tau_g + \nu_0\Delta\tau) + \theta_{mn} - \phi_v + \phi_G] + \cos [2\pi(\nu_{\text{LO}}\tau_g - \nu_0\Delta\tau) + \theta_{mn} - \phi_v - \phi_G] \right\}, \quad (6.28)$$

where $\theta_{mn} = \theta_m - \theta_n$. Equation (6.28) represents the real output of a complex correlator. We rewrite Eq. (6.28) as

$$r_d = |\mathcal{V}||G_{mn}|(\cos \Psi_u + \cos \Psi_\ell), \quad (6.29)$$

where Ψ_u and Ψ_ℓ represent the corresponding expressions in square brackets in Eq. (6.28). The responses considered above represent the normal output of the interferometer, which we call condition 1. The expression for the imaginary output of the correlator is obtained by replacing ϕ_G by $\phi_G - \pi/2$. Consider a second condition in which a $\pi/2$ phase shift is introduced into the first LO signal of antenna m , so that θ_{mn} becomes $\theta_{mn} - \pi/2$. The correlator outputs for the two conditions are obtained from Eqs. (6.28) and (6.29):

$$\underline{\text{condition 1}} \quad (6.30)$$

$$r_1 = |\mathcal{V}||G_{mn}|(\cos \Psi_u + \cos \Psi_\ell)$$

$$r_2 = |\mathcal{V}||G_{mn}|(\sin \Psi_u - \sin \Psi_\ell)$$

$$\underline{\text{condition 2}} \quad (\theta_{mn} \rightarrow \theta_{mn} - \pi/2) \quad (6.31)$$

$$r_3 = |\mathcal{V}||G_{mn}|(\sin \Psi_u + \sin \Psi_\ell)$$

$$r_4 = |\mathcal{V}||G_{mn}|(-\cos \Psi_u + \cos \Psi_\ell)$$

where r_1 and r_3 represent the real outputs of the correlator and r_2 and r_4 the imaginary outputs. Thus, the upper-sideband response, expressed in complex form, is

$$|\mathcal{V}||G_{mn}|(\cos \Psi_u + j \sin \Psi_u) = \frac{1}{2} [(r_1 - r_4) + j(r_2 + r_3)] . \quad (6.32)$$

Similarly, the lower-sideband response is

$$|\mathcal{V}||G_{mn}|(\cos \Psi_\ell + j \sin \Psi_\ell) = \frac{1}{2} [(r_1 + r_4) - j(r_2 - r_3)] . \quad (6.33)$$

If the $\pi/2$ phase shift is periodically switched into and out of the LO signal, the upper- and lower-sideband responses can be obtained as indicated by Eqs. (6.32) and (6.33).

A similar implementation of sideband separation that makes use of fringe frequencies is attributable to B. G. Clark. This method is based on the fact that a small frequency shift in the first LO adds the same frequency shift to the fringes at the correlator for both sidebands, but a similar shift in a later LO adds to the fringe frequency for one sideband but subtracts from it for the other. Consider two antennas of an array in which the fringes have been stopped as in the discussion associated with expressions (6.26) and (6.27). Now suppose that we increase the frequency of the first LO at antenna n by a frequency $\delta\nu$ and decrease the frequency of the second LO by the same amount. The fringe frequency for the upper-sideband signal will be unchanged; that is, the fringes will remain stopped. For the lower sideband, the signal frequencies after the second mixer will be decreased by $2\delta\nu$. The lower-sideband output will consist of fringes at frequency $2\delta\nu(1 - d\tau_g/dt) \approx 2\delta\nu$ and

will be averaged to a small residual if $(2\delta\nu)^{-1}$ is small compared with the integration period at the correlator output, or if an integral number of fringe cycles fall within such an integration period. If the frequency of the second LO is increased by $\delta\nu$ instead of decreased, the lower sideband will be stopped and the upper one averaged out. To apply this scheme to an array of n_a antennas, the offset must be different for each antenna, and this can be achieved by using an offset $n\delta\nu$ for antenna n , where n runs from 0 to $n_a - 1$. An advantage of this sideband-separating scheme is that it can be implemented using the variable LOs required for fringe stopping, and no other special hardware is needed. Unlike the $\pi/2$ phase-switching scheme, one sideband is lost in this method. However, as mentioned above, sideband separation schemes of this type separate only the correlated component of the signal and not the noise. To separate the noise, the SIS mixers at the receiver inputs can be mounted in a sideband-separating circuit of the type described in Appendix 7.1. In such cases, the isolation of the sidebands achieved in the mixer circuit may be only ~ 15 dB, which is sufficient to remove most of the noise contributed by an unwanted sideband, but not sufficient to remove strong spectral lines. The Clark technique described above is nicely suited to increasing the suppression of an unwanted sideband that has already suffered limited rejection at the mixer.

Fringe-frequency effects can also be used for sideband separation in VLBI observations. In VLBI systems, the fringe rotation is usually applied during playback. Fringe rotation then has the effect of reducing the fringe frequency for one sideband and increasing it for the other. If the fringe rotation is set to stop the fringes in one sideband, then since the baselines are so long, fringes resulting from the other sideband will often have a sufficiently high frequency that they will be reduced to a negligible level by the time averaging at the correlator output. The data are played back to the correlator twice, once for each sideband, with appropriate fringe rotation.

6.2 Response to the Noise

The ultimate sensitivity of a receiving system is determined principally by the system noise. We now consider the response to the noise and the resulting threshold of sensitivity, beginning with the effect at the correlator output and the resulting uncertainty in the real and imaginary parts of the visibility, \mathcal{V} . This leads to calculation of the rms noise level in a synthesized image in terms of the peak response to a source of given flux density. Finally, we consider the effect of noise in terms of the rms fluctuations in the amplitude and phase of \mathcal{V} .

6.2.1 Signal and Noise Processing in the Correlator

Consider an observation in which the field to be imaged contains only a point source located at the phase reference position. Let $V_m(t)$ and $V_n(t)$ be the waveforms at the

correlator input from the signal channels of antennas m and n . The output is

$$r = \langle V_m(t)V_n(t) \rangle, \quad (6.34)$$

where all three functions are real, and the expectation denoted by the angular brackets is approximated in practice by a finite time average. To determine the relative power levels of the signal and noise components of r , we determine their power spectra by first calculating the autocorrelation functions. The autocorrelation of the signal product in Eq. (6.34) is

$$\rho_r(\tau) = \langle V_m(t)V_n(t)V_m(t-\tau)V_n(t-\tau) \rangle. \quad (6.35)$$

This expression can be evaluated using the following fourth-order moment relation⁴:

$$\langle z_1z_2z_3z_4 \rangle = \langle z_1z_2 \rangle \langle z_3z_4 \rangle + \langle z_1z_3 \rangle \langle z_2z_4 \rangle + \langle z_1z_4 \rangle \langle z_2z_3 \rangle, \quad (6.36)$$

where z_1, z_2, z_3 , and z_4 are joint Gaussian random variables with zero mean. Thus,

$$\begin{aligned} \rho_r(\tau) &= \langle V_m(t)V_n(t) \rangle \langle V_m(t-\tau)V_n(t-\tau) \rangle \\ &\quad + \langle V_m(t)V_m(t-\tau) \rangle \langle V_n(t)V_n(t-\tau) \rangle \\ &\quad + \langle V_m(t)V_n(t-\tau) \rangle \langle V_m(t-\tau)V_n(t) \rangle \\ &= \rho_{mm}^2(0) + \rho_m(\tau)\rho_n(\tau) + \rho_{mn}(\tau)\rho_{mn}(-\tau), \end{aligned} \quad (6.37)$$

where ρ_m and ρ_n are the unnormalized autocorrelation functions of the two signals V_m and V_n , respectively, and ρ_{mn} is their cross-correlation function. Each V term is the sum of a signal component s and a noise component n , and to examine how these components contribute to the correlator output, we substitute them in Eq. (6.37). Products of uncorrelated terms, that is, products of signal and noise voltages, or noise voltages from different antennas, have an expectation of zero, and omitting them, we obtain

$$\begin{aligned} \rho_r(\tau) &= \langle s_m(t)s_n(t) \rangle \langle s_m(t-\tau)s_n(t-\tau) \rangle \\ &\quad + \langle s_m(t)s_m(t-\tau) + n_m(t)n_m(t-\tau) \rangle \langle s_n(t)s_n(t-\tau) + n_n(t)n_n(t-\tau) \rangle \\ &\quad + \langle s_m(t)s_n(t-\tau) \rangle \langle s_m(t-\tau)s_n(t) \rangle, \end{aligned} \quad (6.38)$$

where the three lines on the right side correspond to the three terms on the last line of Eq. (6.37). To determine the effect of the frequency response of the receiving

⁴This relation is a special case of a more general expression for the expectation of the product of N such variables, which is zero if N is odd and a sum of pair products if N is even. A form of Eq. (6.36) can be found in Lawson and Uhlenbeck (1950), Middleton (1960), and Wozencraft and Jacobs (1965).

system on the various terms of $\rho(\tau)$, we need to convert them to power spectra. By the Wiener–Khinchin relation, we should therefore examine the Fourier transforms of each term on the right sides of Eqs. (6.37) and (6.38).

The first term from Eq. (6.37), $\rho_{mn}^2(0)$, is a constant, and its Fourier transform is a delta function at the origin in the frequency domain, multiplied by $\rho_{mn}^2(0)$. From Eq. (6.38), we see that $\rho_{mn}^2(0)$ involves only the signal terms, which it is convenient to express as antenna temperatures. By the integral theorem of Fourier transforms, $\rho_{mn}(0)$ is the infinite integral of the Fourier transform of $\rho_{mn}(\tau)$, and thus the Fourier transform of $\rho_{mn}^2(0)$ is

$$k^2 T_{Am} T_{An} \left[\int_{-\infty}^{\infty} H_m(\nu) H_n^*(\nu) d\nu \right]^2 \Delta(\nu), \quad (6.39)$$

where k is Boltzmann's constant, T_{Am} and T_{An} are the components of antenna temperature resulting from the source, $H_m(\nu)$ and $H_n(\nu)$ are the frequency responses of the signal channels, and $\Delta(\nu)$ is the bandwidth.

The Fourier transform of the second term of Eq. (6.37), $\rho_m(\tau)\rho_n(\tau)$, is the convolution of the transforms of ρ_m and ρ_n , that is

$$k^2 (T_{Sm} + T_{Am})(T_{Sn} + T_{An}) \int_{-\infty}^{\infty} H_m(\nu) H_m^*(\nu) H_n(\nu' - \nu) H_n^*(\nu' - \nu) d\nu, \quad (6.40)$$

where T_{Sm} and T_{Sn} are the system temperatures. Note that the magnitude of this term is proportional to the product of the total noise temperatures.

The Fourier transform of the third term of Eq. (6.37), $\rho_{mn}(\tau)\rho_{mn}(-\tau)$, is the convolution of the transforms of $\rho_{mn}(\tau)$ and $\rho_{mn}(-\tau)$, and the latter is the complex conjugate of the former, since ρ_{mn} is real. Thus, the Fourier transform of $\rho_{mn}(\tau)\rho_{mn}(-\tau)$ is

$$k^2 T_{Am} T_{An} \int_{-\infty}^{\infty} H_m(\nu) H_n^*(\nu) H_m^*(\nu' - \nu) H_n(\nu' - \nu) d\nu. \quad (6.41)$$

In expression (6.39), as in Eq. (6.37), only the antenna temperatures appear, because the receiver noise for different antennas makes no contribution to the cross-correlation.

Expression (6.39) represents the signal power in the correlator output, and (6.40) and (6.41) represent the noise. The effect of the time averaging at the correlator output can be modeled in terms of a filter that passes frequencies from 0 to $\Delta\nu_{LF}$. The output bandwidth $\Delta\nu_{LF}$ is less than the correlator input bandwidth by several or many orders of magnitude. Therefore, the spectral density of the output noise can be assumed to be equal to its value at zero frequency, that is, for $\nu' = 0$ in (6.40) and (6.41). From these considerations, and because $H_m(\nu)$ and $H_n(\nu)$ are Hermitian, the ratio of the signal voltage to the rms noise voltage after averaging at

the correlator output is

$$\mathcal{R}_{\text{sn}} = \frac{\sqrt{T_{Am}T_{An}} \int_{-\infty}^{\infty} H_m(\nu)H_n^*(\nu) d\nu}{\sqrt{(T_{Am} + T_{Sm})(T_{An} + T_{Sn}) + T_{Am}T_{An}} \sqrt{2\Delta\nu_{\text{LF}} \int_{-\infty}^{\infty} |H_m(\nu)|^2|H_n(\nu)|^2 d\nu}}, \quad (6.42)$$

where $2\Delta\nu_{\text{LF}}$ is the equivalent bandwidth after averaging, with negative frequencies included. It is unusual for \mathcal{R}_{sn} , the estimate of the SNR at the output of a simple correlator, to be required to an accuracy better than a few percent. Indeed, it is usually difficult to specify T_S to any greater accuracy since the effects of ground radiation and atmospheric absorption on T_S vary as the antennas track. Thus, it is usually satisfactory to approximate $H_m(\nu)$ and $H_n(\nu)$ by identical rectangular functions of width $\Delta\nu_{\text{IF}}$. Also, in sensitivity calculations, one is concerned most often with sources near the threshold of detectability, for which $T_A \ll T_S$. With these simplifications, Eq. (6.42) becomes

$$\mathcal{R}_{\text{sn}} = \sqrt{\frac{T_{Am}T_{An}}{T_{Sm}T_{Sn}}} \sqrt{\frac{\Delta\nu_{\text{IF}}}{\Delta\nu_{\text{LF}}}}. \quad (6.43)$$

Figure 6.7 shows the signal and noise spectra for the rectangular bandpass approximation. Note that the input spectra $|H_m(\nu)|^2$ and $|H_n(\nu)|^2$ contain both positive and negative frequencies and are symmetric about the origin in ν . Thus, the output noise spectrum can be described as proportional to either the convolution or the cross-correlation function of $|H_m(\nu)|^2$ and $|H_n(\nu)|^2$.

The output bandwidth is related to the data averaging time τ_a since the averaging can be described as convolution in the time domain with a rectangular function of unit area and width τ_a . The power response of the averaging circuit as a function of frequency is the square of the Fourier transform of the rectangular function, that is, $\sin^2(\pi\tau_a\nu)/(\pi\tau_a\nu)^2$. The equivalent bandwidth, including both positive and negative frequencies, is

$$2\Delta\nu_{\text{LF}} = \int_{-\infty}^{\infty} \frac{\sin^2(\pi\tau_a\nu)}{(\pi\tau_a\nu)^2} d\nu = \frac{1}{\tau_a}. \quad (6.44)$$

Then from Eq. (6.43), we obtain

$$\mathcal{R}_{\text{sn}} = \sqrt{\left(\frac{T_{Am}T_{An}}{T_{Sm}T_{Sn}}\right)} 2\Delta\nu_{\text{IF}}\tau_a. \quad (6.45)$$

Note that $2\Delta\nu_{\text{IF}}\tau_a$ is the number of independent samples of the signal in time τ_a .

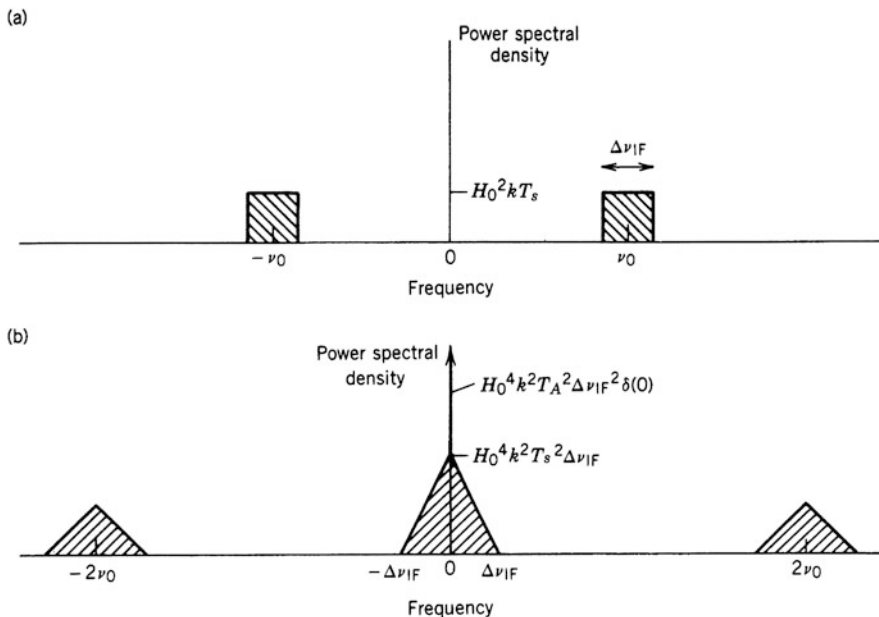


Fig. 6.7 Spectra of (a) the input and (b) the output waveforms of a correlator. The input passbands are rectangular of width $\Delta\nu_{IF}$. Shown in (b) is the complete spectrum of signals generated in the multiplication process, including noise bands at twice the input frequency. Only frequencies very close to zero are passed by the averaging circuit at the correlator output. These include the wanted signal, the spectrum of which has the form of a delta function and is represented by the arrow. It is assumed that $T_A \ll T_S$.

If the source is unpolarized, each antenna responds to half the total flux density S , and the received power density is

$$kT_A = \frac{1}{2}AS, \tag{6.46}$$

where A is the effective collecting area of the antenna. For identical antennas and system temperatures, we obtain, from Eqs. (6.45) and (6.46),

$$\mathcal{R}_{sn} = \frac{AS}{kT_S} \sqrt{\frac{\Delta\nu_{IF}\tau_a}{2}}. \tag{6.47}$$

Similar derivations of this result can be found in Blum (1959), Colvin (1961), and Tiuri (1964). Usually the result in Eq. (6.47), in which we have assumed $T_A \ll T_S$, is the one needed. At the other extreme, which may be encountered in observations of very strong, unresolved sources for which $T_A \gg T_S$, we have $\mathcal{R}_{sn} = \sqrt{\Delta\nu_{IF}\tau_a}$. The SNR is then determined by the fluctuations in signal level and is independent of the areas of the antennas. Anantharamaiah et al. (1989) give a discussion of noise levels in the observation of very strong sources.

From Fig. 6.7, we can see how the factor $\sqrt{\Delta\nu_{\text{IF}}\tau_a}$ in Eq. (6.47), which enables very high sensitivity to be achieved in radio astronomy, arises. The noise within the correlator results from beats between components in the two input bands and thus extends in frequency up to $\Delta\nu_{\text{IF}}$. The triangular noise spectrum in Fig. 6.7 is simply proportional to the number of beats per unit frequency interval. However, only the very small fraction of this noise that falls within the output bandwidth is retained after the averaging. Note that the signal bandwidth $\Delta\nu_{\text{IF}}$ that is important here is the bandwidth at the correlator input. In a DSB system, this is only one-half of the total input bandwidth at the antenna.

One other factor that affects the SNR should be introduced at this point. If the signals are quantized and digitized before entering the correlators, a quantization efficiency η_Q related to the quantization must be included, and Eq. (6.47) becomes

$$\mathcal{R}_{\text{sn}} = \frac{AS\eta_Q}{kT_S} \sqrt{\frac{\Delta\nu_{\text{IF}}\tau_a}{2}}, \quad (6.48)$$

or in terms of antenna temperature,

$$\mathcal{R}_{\text{sn}} = \frac{T_A\eta_Q}{T_S} \sqrt{2\Delta\nu_{\text{IF}}\tau_a}. \quad (6.49)$$

Values of η_Q vary between 0.637 and 1.0 and are discussed in Chap. 8 (see Table 8.1). In VLBI observing, other losses affect the SNR, as discussed in Sect. 9.7.

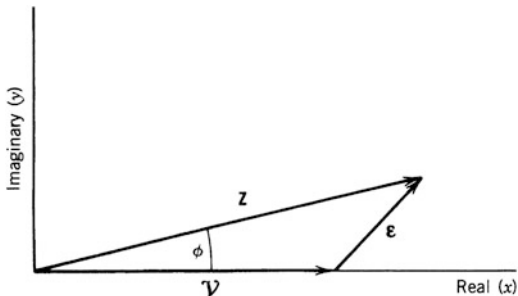
6.2.2 Noise in the Measurement of Complex Visibility

To understand precisely what \mathcal{R}_{sn} represents, note that in deriving Eqs. (6.48) and (6.49), no delay was introduced between the signal components at the correlator, and the phase responses of the signal channels were assumed to be identical. Thus, the source is in the central fringe of the interferometer pattern, and the response is the peak fringe amplitude, which represents the modulus of the visibility. To express the rms noise level at the correlator output in terms of the flux density σ of an unresolved source for which the peak fringe amplitude produces an equal output, we put $\mathcal{R}_{\text{sn}} = 1$ in Eq. (6.48) and replace S by σ :

$$\sigma = \frac{\sqrt{2}kT_S}{A\eta_Q} \sqrt{\Delta\nu_{\text{IF}}\tau_a}, \quad (6.50)$$

where σ is in units of $\text{W m}^{-2} \text{Hz}^{-1}$. Consider the case of an instrument with a complex correlator in which the output oscillations are slowed to zero frequency as described earlier. The noise fluctuations in the real and imaginary outputs are uncorrelated, as we now show. Suppose the antennas are pointed at blank sky so

Fig. 6.8 Complex quantity \mathbf{Z} , which is the sum of the modulus of the true complex visibility \mathcal{V} and the noise $\boldsymbol{\varepsilon}$. The noise has real and imaginary components of rms amplitude σ , and ϕ is the phase deviation resulting from the noise.



that the only inputs to the correlators in Fig. 6.3 are the noise waveforms n_m , n_n , and n_m^H , where the last is the Hilbert transform of n_m produced by the quadrature phase shift. The expectation of the product of the real and imaginary outputs is $\langle n_m n_n n_m^H n_n \rangle$, which can be shown to be zero by using Eq. (6.36) and noting that the expectations $\langle n_m n_n \rangle$, $\langle n_m n_m^H \rangle$, and $\langle n_m^H n_n \rangle$ must all be zero. Thus, the noise from the real and imaginary outputs is uncorrelated.⁵

The signal and noise components in the measurement of the complex visibility are shown in Fig. 6.8 as vectors in the complex plane. Here \mathcal{V} represents the visibility as it would be measured in the absence of noise, which is assumed to be along the x , or real, axis; and \mathbf{Z} represents the sum of the visibility and noise, $\mathcal{V} + \boldsymbol{\varepsilon}$. We consider \mathbf{Z} and $\boldsymbol{\varepsilon}$ to be vectors whose components correspond to the real and imaginary parts of the corresponding quantities. The components of $\boldsymbol{\varepsilon}$ are independent Gaussian random variables with zero mean and variance σ^2 . Hence, the noise in both components of \mathbf{Z} has an rms amplitude σ , and

$$\langle |\mathbf{Z}|^2 \rangle = |\mathcal{V}|^2 + 2\sigma^2 . \tag{6.51}$$

The factor of two arises because of the contributions of the real and imaginary parts of $\boldsymbol{\varepsilon}$. If the measurement is made using only a single-multiplier correlator, one can periodically introduce a quadrature phase shift at one input, thus obtaining real and

⁵The noise in the correlator outputs is composed of an ensemble of components of frequency $|\nu_m - \nu_n|$, where ν_m and ν_n are frequency components of the correlator inputs n_m and n_n . Components of the imaginary output are shifted in phase by $\pm\pi/2$ relative to the corresponding components of the real output. Note that for any pair of input components, the sign of the shift in the imaginary output takes opposite values depending on whether $\nu_m > \nu_n$ or $\nu_m < \nu_n$. As a result, the noise waveforms at the correlator outputs are not a Hilbert transform pair, and one cannot be derived from the other.

imaginary outputs, each for half of the observing time. Then the data are half of those that would be obtained with a complex correlator, and the noise in the visibility measurement is greater by $\sqrt{2}$.

6.2.3 Signal-to-Noise Ratio in a Synthesized Image

Having determined the noise-induced error in the visibility, the next step is to consider the SNR in an image. Consider an array with n_p antenna pairs, and suppose that the visibility data are averaged for time τ_a and that the whole observation covers a time interval τ_0 . The total number of independent data points in the (u, v) plane is therefore

$$n_d = n_p \frac{\tau_0}{\tau_a} . \quad (6.52)$$

In imaging an unresolved source at the field center for which the visibility data combine in phase, we should thus expect the SNR in the image to be greater than that in Eqs. (6.48) and (6.49) by a factor $\sqrt{n_p \tau_0 / \tau_a}$. This simple consideration gives the correct result for the case in which the data are combined with equal weights. We now derive the result for the more general case of arbitrarily weighted data.

The ensemble of measured data can be represented by

$$\sum_{i=1}^{n_d} \left[{}^2\delta(u - u_i, v - v_i)(\mathcal{V}_i + \varepsilon_i) + {}^2\delta(u + u_i, v + v_i)(\mathcal{V}_i^* + \varepsilon_i^*) \right] , \quad (6.53)$$

where ${}^2\delta$ is the two-dimensional delta function and ε_i is the complex noise contribution to the i th measurement. Each such data point appears at two (u, v) locations, reflected through the origin of the (u, v) plane. Before taking the Fourier transform of the data in Eq. (6.53), each data point is assigned a weight w_i (the choice of weighting factors is discussed in Sect. 10.2.2). To simplify the calculation, we assume that the source is unresolved and located at the phase reference point of the image and therefore produces a constant real visibility \mathcal{V} equal to its flux density S . The intensity at the center of the image is then

$$I_0 = \frac{\sum_{i=1}^{n_d} w_i (\mathcal{V} + \varepsilon_{\mathcal{R}i})}{\sum w_i} , \quad (6.54)$$

where $\varepsilon_{\mathcal{R}i}$ is the real part of the noise, ε_i . Note that the imaginary part of ε_i vanishes at the origin of the image when the conjugate components are summed. For neighboring points in the image, the same rms level of noise is distributed between the real and imaginary parts of ε . The expectation of I_0 is

$$\langle I_0 \rangle = \mathcal{V} = S, \quad (6.55)$$

since $\langle \varepsilon_{\mathcal{R}i} \rangle = 0$. The variance of the estimate of the intensity, σ_m^2 , is

$$\sigma_m^2 = \langle I_0^2 \rangle - \langle I_0 \rangle^2 = \frac{\sum w_i^2 \langle \varepsilon_{\mathcal{R}i}^2 \rangle}{\left(\sum w_i \right)^2}. \quad (6.56)$$

Equation (6.56) is derived directly from Eq. (6.54) using the fact that the noise terms from different (u, v) locations are uncorrelated, that is, $\langle \varepsilon_{\mathcal{R}i} \varepsilon_{\mathcal{R}j} \rangle = 0$, for $i \neq j$. We define the mean weighting factor w_{mean} and rms weighting factor w_{rms} by the equations

$$w_{\text{mean}} = \frac{1}{n_d} \sum w_i \quad (6.57)$$

and

$$w_{\text{rms}}^2 = \frac{1}{n_d} \sum w_i^2. \quad (6.58)$$

The rms noise contribution [see Eq. (6.51)] is the same for each (u, v) point and is equal to $\langle \varepsilon_{\mathcal{R}i}^2 \rangle = \sigma^2$, where σ is given by Eq. (6.50). Thus, the SNR can be calculated from Eqs. (6.55), (6.56), (6.57), and (6.58) as

$$\frac{\langle I_0 \rangle}{\sigma_m} = \frac{S \sqrt{n_d}}{\sigma} \frac{w_{\text{mean}}}{w_{\text{rms}}}. \quad (6.59)$$

For an array with complex correlators, we have, from Eq. (6.50),

$$\frac{\langle I_0 \rangle}{\sigma_m} = \frac{AS\eta_Q \sqrt{n_d \Delta v_{\text{IF}} \tau_a}}{\sqrt{2kT_S}} \frac{w_{\text{mean}}}{w_{\text{rms}}}. \quad (6.60)$$

If combinations of all pairs of antennas are used, $n_p = \frac{1}{2}n_a(n_a - 1)$, where n_a is the number of antennas. Since, from Eq. (6.52), $n_d = n_p \tau_0 / \tau_a$, we obtain

$$\frac{\langle I_0 \rangle}{\sigma_m} = \frac{AS\eta_Q \sqrt{n_a(n_a - 1) \Delta v_{\text{IF}} \tau_0}}{2kT_S} \frac{w_{\text{mean}}}{w_{\text{rms}}}. \quad (6.61)$$

To express the rms noise level in terms of flux density, we put $I_0/\sigma_m = 1$ in Eq. (6.61). S then represents the flux density of a point source for which the peak response is equal to the rms noise level. If we represent this particular value of S by S_{rms} , then

$$S_{\text{rms}} = \frac{2kT_S}{A\eta_Q \sqrt{n_a(n_a - 1)} \Delta\nu_{\text{IF}} \tau_0} \frac{w_{\text{rms}}}{w_{\text{mean}}}. \quad (6.62)$$

If all the weighting factors w_i are equal, $w_{\text{mean}}/w_{\text{rms}} = 1$, and this situation is referred to as the use of *natural weighting*. In such a case, the SNR given by Eq. (6.61) is equal to the corresponding sensitivity for a total-power receiver combined with an antenna of aperture $\sqrt{n_a(n_a - 1)}A$, which approaches $n_a A$ as n_a becomes large. For an analysis of the sensitivity of single-antenna systems, see Appendix 1.1.

We have considered the *point-source sensitivity* in Eq. (6.62). In the case of a source that is wider than the synthesized beam, it is useful to know the *brightness sensitivity*. The flux density (in $\text{W m}^{-2} \text{Hz}^{-1}$) received from a broad source of mean intensity I ($\text{W m}^{-2} \text{Hz}^{-1} \text{sr}^{-1}$) across the synthesized beam is $I\Omega$, where Ω is the effective solid angle of the synthesized beam. Thus, the intensity level that is equal to the rms noise is S_{rms}/Ω . Note that the brightness sensitivity decreases as the synthesized beam becomes smaller, so compact arrays are best for detecting broad, faint sources. However, to measure the intensity of a uniform background, a measurement of the total power received in an antenna is required because a correlation interferometer does not respond to such a background.

The ratio $w_{\text{mean}}/w_{\text{rms}}$ is less than unity except when the weighting is uniform. Although the SNR depends on the choice of weighting, in practice, this dependence is not highly critical. The use of natural weighting maximizes the sensitivity for detection of a point source in a largely blank field but can also substantially broaden the synthesized beam. The advantage in sensitivity is usually small. For example, if the density of data points is inversely proportional to the distance from the (u, v) origin, as is the case for an east–west array with uniform increments in antenna spacing, the weighting factors required to obtain effective uniform density of data result in $w_{\text{mean}}/w_{\text{rms}} = 2\sqrt{2}/3 = 0.94$. In this case, the natural weighting results in an undesirable beam profile in which the response remains positive for large angular distances from the beam axis and dies away only slowly.

Methods of Fourier transformation of visibility data are reviewed in Chap. 10, and the results derived in Eqs. (6.61) and (6.62) can be applied to these by using the appropriate values of w_{mean} and w_{rms} . Convolution of the visibility data in the (u, v) plane to obtain values at points on a rectangular grid is a widely used process. In general, the data at adjacent grid points are then not independent, and a tapering of the signal and noise is introduced into the image. Aliasing can also cause the SNR to vary across the image. (These effects are explained in Fig. 10.5 and the associated discussion.) In such cases, the results derived here apply near the origin of the image, where the effects of tapering and aliasing are unimportant. The rms noise level over

the image can be obtained by the application of Parseval's theorem to the noise in the visibility data (see Appendix 2.1).

In practice, a number of factors that affect the SNR are difficult to determine precisely. For example, T_S varies somewhat with antenna elevation. There are also a number of effects that can reduce the response to a source without reducing the noise, but these are important only for sources not near the (l, m) origin of an image. These include the smearing resulting from the receiving bandwidth and from visibility averaging, discussed later in this chapter, and the effect of non-coplanar baselines, discussed in Sect. 11.7.

Note also that in many instruments, two oppositely polarized signals (with crossed linear or opposite circular polarizations) are received and processed using separate IF amplifiers and correlators. For unpolarized sources, the overall SNR is then $\sqrt{2}$ greater than the values derived above, which include only one signal from each antenna.

6.2.4 Noise in Visibility Amplitude and Phase

In synthesis imaging, we are usually concerned with data in the form of the real and imaginary parts of the complex visibility \mathcal{V} , but sometimes it is necessary to work with the amplitude and phase. The sum of the visibility and noise is represented by $\mathbf{Z} = Z e^{j\phi}$, where we choose the real axis so that the phase ϕ is measured with respect to the phase of \mathcal{V} , as in Fig. 6.8. Then for $T_A \ll T_S$ (the antenna temperature resulting from the source is much less than the system temperature), the probability distributions of the resulting amplitude and phase are

$$p(Z) = \frac{Z}{\sigma^2} \exp\left(-\frac{Z^2 + |\mathcal{V}|^2}{2\sigma^2}\right) I_0\left(\frac{Z|\mathcal{V}|}{\sigma^2}\right), \quad Z > 0 \quad (6.63a)$$

$$p(\phi) = \frac{1}{2\pi} \exp\left(-\frac{|\mathcal{V}|^2}{2\sigma^2}\right) \left\{ 1 + \sqrt{\frac{\pi}{2}} \frac{|\mathcal{V}| \cos \phi}{\sigma} \exp\left(\frac{|\mathcal{V}|^2 \cos^2 \phi}{2\sigma^2}\right) \times \left[1 + \operatorname{erf}\left(\frac{|\mathcal{V}| \cos \phi}{\sqrt{2}\sigma}\right) \right] \right\}, \quad (6.63b)$$

and erf is the error function (Abramowitz and Stegun 1968).

$$\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2/2} dt, \quad (6.63c)$$

where I_0 is the modified Bessel function of zero order and σ is as given by Eq. (6.50). The amplitude distribution is identical to that for a sine wave in noise, and the derivation is given by Rice (1944, 1954), Vinokur (1965), and Papoulis (1965), of which the last two also derive the result for the phase. $p(Z)$ is sometimes referred to as the Rice distribution, and for $\mathcal{V} = 0$, it reduces to the Rayleigh distribution. Curves of $p(Z)$ and $p(\phi)$ are given in Fig. 6.9. Comparison of the curves

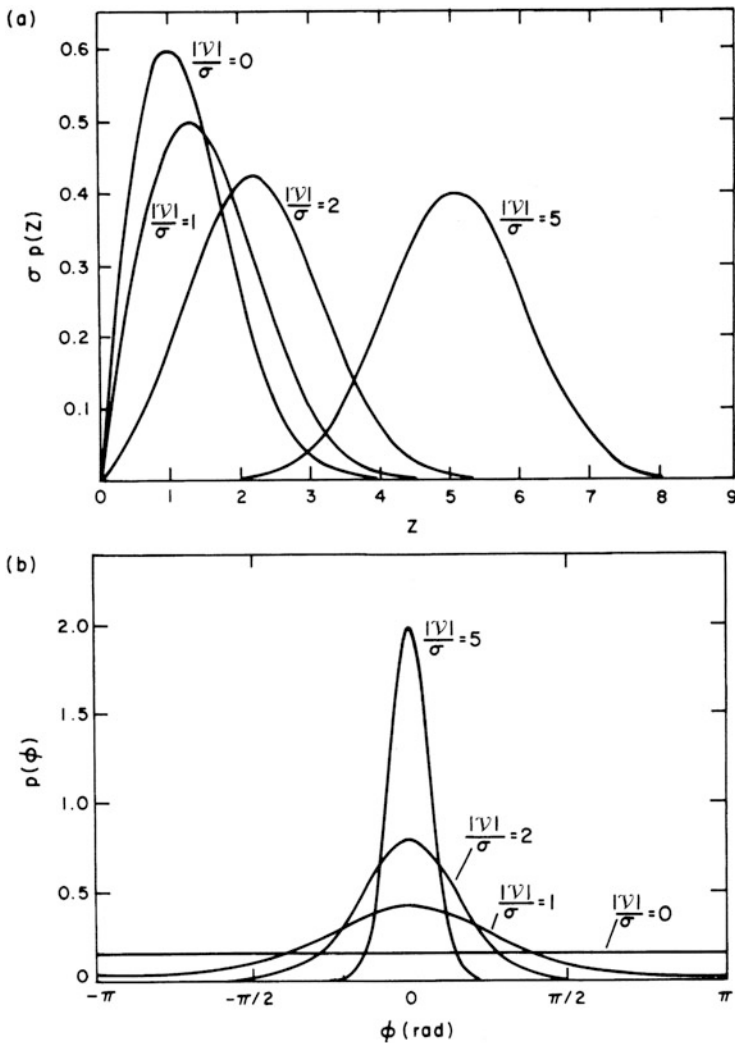


Fig. 6.9 Probability distributions of (a) the amplitude, and (b) the phase, of the measured complex visibility as functions of the SNR. $|V|$ is the modulus of the signal component. Reprinted from Moran (1976), © 1976, with permission from Elsevier.

for $|\mathcal{V}|/\sigma = 0$ and 1 indicates that the presence of a weak signal is more easily detected by examining the visibility phase than by examining the amplitude.

Approximation for $p(Z)$ and $p(\phi)$ for the cases in which $|\mathcal{V}|/\sigma \ll 1$ and $|\mathcal{V}|/\sigma \gg 1$ are given in Sect. 9.3. Expressions for the moments of Z and ϕ and their rms deviations are also given in that section. The rms phase deviation σ_ϕ is a particularly useful quantity, especially for astrometric and diagnostic work. The expression for σ_ϕ , valid for the case in which $|\mathcal{V}|/\sigma \gg 1$, is $\sigma_\phi \simeq \sigma/|\mathcal{V}|$ [Eq. (9.67)]. This result also follows intuitively from an examination of Fig. 6.8. By substituting Eq. (6.50) into the expression for σ_ϕ , setting $|\mathcal{V}|$ equal to the flux density S of the source, which is appropriate if the source is unresolved, and using Eq. (6.46) to relate the flux density and antenna temperature, we obtain

$$\sigma_\phi = \frac{T_S}{\eta_Q T_A \sqrt{2\Delta\nu_{\text{IF}}\tau_a}}. \quad (6.64)$$

This equation is valid for the conditions $T_S/\sqrt{2\Delta\nu_{\text{IF}}\tau_a} \ll T_A \ll T_S$, which are the conditions most frequently encountered, and is useful for determining whether the noise in the phase measurements of an interferometer is due exclusively to receiver noise. Excess phase noise can be contributed by the atmosphere, by system instabilities, and, in the case of VLBI, by the frequency standards.

6.2.5 Relative Sensitivities of Different Interferometer Systems

Next we compare the sensitivity of several different interferometer systems, using as a measure of sensitivity the modulus of the signal divided by the rms noise, that is, \mathcal{V}/ε in terms of the quantities at the correlator output in Fig. 6.8. Parameters such as averaging times and IF bandwidths are the same for all cases considered. To compare DSB and SSB cases, it is convenient to introduce a factor

$$\alpha = \frac{\text{double-sideband system temperature of double-sideband system}}{\text{system temperature of single-sideband system}}. \quad (6.65)$$

Recall that the system temperature of a receiver can be defined as the noise temperature of a thermal source at the input of a hypothetical noise-free (but otherwise identical) receiver that would produce the same noise level at the receiver output. [Equation (1.4) can be used for the equivalent noise temperature if the Rayleigh–Jeans approximation does not apply.] For a DSB receiver, the system temperature is described as DSB or SSB depending on whether the thermal noise source emits noise in both sidebands or only one. With these definitions, the SSB noise temperature is twice the DSB noise temperature.

For an SSB system, the rms noise from one output of a correlator (either the real or imaginary output in the case of a complex correlator) is σ after averaging for a time τ_a , as given by Eq. (6.50). The corresponding noise power is σ^2 . For a DSB

system, the rms output noise at a correlator output is $2\alpha\sigma$. In all cases, the signal results from an unresolved source. For an SSB system, we take the signal voltage from the correlator output to be \mathcal{V} , as in Fig. 6.8. For a DSB system with the input signal in one sideband only, the signal at the correlator output is \mathcal{V} , and for a DSB system with input in both sidebands, the correlator output is $2\mathcal{V}$.

Values of the relative sensitivity for various systems are discussed below and summarized in Table 6.1. Similar results are given by Rogers (1976).

1. *SSB system with complex correlator.* The output signal is \mathcal{V} , and the rms noise from each correlator output is σ . As shown by Fig. 6.9 and Eq. (6.51), the ratio of the signal amplitude to rms noise is $\mathcal{V}/(\sqrt{2}\sigma)$. We shall take this as the standard with respect to which the relative sensitivities of other systems are defined.
2. *SSB system and simple correlator with fringe fitting.* To measure both the real and imaginary parts of the complex visibility, the fringes are not stopped but appear as a sinusoid of amplitude \mathcal{V} at the fringe frequency ν_f . The signal is accompanied by noise of rms amplitude σ . The amplitude and phase are measured by “fringe fitting,” that is, performing a least-mean-squares fit of a sinusoid to the correlator output. This procedure involves multiplying the correlator output waveform by $\cos(2\pi\nu_f t)$ and $\sin(2\pi\nu_f t)$ and integrating over the period τ_a . The results

Table 6.1 Relative signal-to-noise ratios for several types of systems

System type	Relative SNR
1. SSB with complex correlator	1
2. SSB with simple correlator	$\frac{1}{\sqrt{2}}$
3. SSB, simple correlator, fringe stopping, $\pi/2$ phase switching	$\frac{1}{\sqrt{2}}$
4. DSB, simple correlator, ^a fringe fitting, continuum signal	$\frac{1}{\sqrt{2\alpha}}$
5. DSB, simple correlator, fringe stopping, $\pi/2$ phase switching, continuum signal	$\frac{1}{\sqrt{2\alpha}}$
6a. DSB, fringe stopping, sideband-separation [Eqs. (6.30) to (6.33)], signal in one sideband only	$\frac{1}{2\alpha}$
6b. As (6a), but for continuum signal and visibilities in both sidebands combined	$\frac{1}{\sqrt{2\alpha}}$
7a. VLBI, DSB, complex correlator, one sideband removed by averaging of fast fringes	$\frac{1}{2\alpha}$
7b. As (7a), but for continuum signal, correlated separately for each sideband and results combined	$\frac{1}{\sqrt{2\alpha}}$
8. SSB, digital spectral correlator with simple correlator elements and correlation measured as a function of time offsets (see Sect. 8.8)	1

^aFor DSB with complex correlator, see text pertaining to Fig. 6.5

represent the real and imaginary parts, respectively, of the cross-correlation. We calculate the effects of fringe fitting on the signal and noise separately and assume, with no loss of generality, that the fringes are in phase with the cosine component in the fringe fitting, in which case the sine component of the signal is zero. The correlator output has a bandwidth $\Delta\nu_c$ which is sufficient to pass the fringe-frequency waveform, and it is sampled at time intervals $\tau_s = 1/(2\nu_c)$ and digitized. Within the period τ_a , there are $N = 2\Delta\nu_c\tau_a$ samples. Thus, for the cosine component of the signal, the amplitude is

$$\frac{1}{N} \sum_{i=1}^N \mathcal{V} \cos^2(2\pi i\nu_f\tau_s) = \frac{\mathcal{V}}{2} + \frac{\mathcal{V}}{2N} \sum_{i=1}^N \cos(4\pi i\nu_f\tau_s). \quad (6.66)$$

The second term on the right side represents the end effects and is approximately zero if there are an integral number of half-cycles of the fringe frequency within the period τ_a . It also becomes relatively small as $\nu_f\tau_a$ increases, and we assume here that there are enough fringe cycles (say, ten or more) within time τ_a that end effects can be neglected. To determine the effect of fringe fitting on the noise, we represent the sampled noise by $n(i\tau_s)$, multiply by the cosine function, and determine the variance (mean squared value). Averaged over time τ_a , the result is

$$\begin{aligned} & \frac{1}{N} \left[\sum_{i=1}^N n(i\tau_s) \cos(2\pi i\nu_f\tau_s) \right]^2 \\ &= \frac{1}{N} \sum_{i=1}^N \sum_{k=1}^N n(i\tau_s) \cos(2\pi i\nu_f\tau_s) n(k\tau_s) \cos(2\pi k\nu_f\tau_s). \end{aligned} \quad (6.67)$$

We need to determine the expectation value of this expression, denoted by angle brackets. Only terms for which $i = k$ contribute to the expectation. Thus, the noise variance becomes

$$\left\langle \frac{1}{2N} \sum_{i=1}^N n^2(i\tau_s) [1 + \cos(4\pi i\nu_f\tau_s)] \right\rangle = \frac{\sigma^2}{2}. \quad (6.68)$$

This result shows that half of the noise power, σ^2 , that is available at the correlator output appears in the cosine component of the fringe fitting. Similarly, the other half appears in the sine component. The combined rms noise of the two components is σ , and the SNR after fringe fitting is $\mathcal{V}/(2\sigma)$. The relative sensitivity is $1/\sqrt{2}$.

3. *SSB system with simple correlator and $\pi/2$ phase switching of LO.* In this case, the fringes have been stopped, and to determine the complex visibility, a phase change of $\pi/2$ is periodically inserted into one oscillator [e.g., $\theta_n \rightarrow \theta_n + \pi/2$ in Eq. (6.11) or (6.15)] so that the correlator is effectively time-shared between the real and imaginary parts of the cross-correlation function, which are averaged

separately. The visibility phase can thereby be determined. The signal in the two phase conditions is $\mathcal{V} \cos(\phi_v)$ and $\mathcal{V} \sin(\phi_v)$, and the rms noise associated with each of these terms is $\sqrt{2}\sigma$ (the $\sqrt{2}$ factor enters because the noise in each output is averaged over time $\tau/2$ only). Thus, the modulus of the signal is \mathcal{V} and the rms noise from the two components is 2σ . The SNR is $\mathcal{V}/(2\sigma)$ and the relative sensitivity is $1/\sqrt{2}$.

4. *DSB system with simple correlator and fringe fitting.* We consider the case of a continuum source with signal in both sidebands and assume that the instrumental delay is adjusted so that the signal appears entirely in the (real) output of a simple correlator, as a fringe-frequency cosine wave of amplitude \mathcal{V} . In terms of Eq. (6.18), the factor $\cos(2\pi\nu_0\Delta\tau_a + \phi_G)$ is unity. Then for the DSB system, the signal amplitude is $2\mathcal{V}$, and the rms noise is $2\alpha\sigma$. The fringe-fitting procedure follows that of case 2, but in this case, the signal amplitude is greater by a factor of two and is equal to \mathcal{V} . The rms noise is greater by a factor of 2α . Thus, the SNR is $\mathcal{V}/(2\alpha\sigma)$, and the relative sensitivity is $1/(\sqrt{2}\alpha)$.
5. *DSB system with simple correlator and $\pi/2$ phase switching of LO.* Here, the fringes have been stopped, and to determine the visibility phase, it is necessary to perform $\pi/2$ phase switching as in case 3 above. (For a DSB system, the phase switching must be on the first LO.) The amplitude of the signal is $2\mathcal{V}$ because the system is DSB, and the rms noise level from the correlator output is increased to $2\sqrt{2}\alpha\sigma$ because the averaging time for each component is reduced to $\tau_a/2$ by the time sharing of the correlator between the two phase conditions. This rms level is associated with both the cosine and sine components of the signal, so the SNR is $\mathcal{V}/(2\alpha\sigma)$. The relative sensitivity is $1/(\sqrt{2}\alpha)$.
6. *One sideband of a DSB system with $\pi/2$ phase switching of the LO and sideband separation after correlation.* A complex correlator is used, and the procedure corresponding to Eqs. (6.30) to (6.33) is followed. We consider the upper sideband and ignore lower-sideband signal terms. The components r_1, r_2, r_3, r_4 have amplitudes \mathcal{V} multiplied by the cosine or sine of Ψ_u . Thus, from Eqs. (6.30) and (6.31), ignoring lower-sideband terms, the right side of Eq. (6.32) becomes $\frac{1}{2}(2\mathcal{V} \cos \Psi_u + j2\mathcal{V} \sin \Psi_u)$, the modulus of which is \mathcal{V} . The rms noise associated with each term $r_1, r_2, r_3,$ and r_4 is $2\sqrt{2}\alpha\sigma$ since the system is DSB, and because of the LO switching, the effective averaging time is $\tau_a/2$. Thus, the rms noise associated with the right side of Eq. (6.32) is $2\sqrt{2}\alpha\sigma$, as in case 5. The SNR is $\mathcal{V}/(2\sqrt{2}\alpha\sigma)$, and the relative sensitivity is $1/(2\alpha)$. This applies to a signal in one sideband such as a spectral line. For a continuum source, the cross-correlation can be measured for each of the two sidebands, and if the results are then averaged, the relative sensitivity becomes $1/(\sqrt{2}\alpha)$. The terms r_2 and r_4 are eliminated in averaging the right sides of Eqs. (6.32) and (6.33), and the result is the same as for a simple correlator with LO phase switching described under case 5 above.
7. *VLBI observations with a DSB system and complex correlator.* In VLBI observations, a DSB system is sometimes used, and fringe rotation is inserted after playback of the recorded signal, as mentioned in Sect. 6.1. For one sideband,

the fringes are stopped, but for the other, they are lost in the averaging at the correlator output because the fringe frequencies are high. Thus, for one playback, we have the signal of an SSB system and the noise of a DSB system in each of the real and imaginary outputs, that is, an SNR of $\mathcal{V}/(2\sqrt{2}\alpha\sigma)$ and a relative sensitivity of $1/(2\alpha)$ for each individual sideband.

8. *Measurement of cross-correlation as a function of time offset.* Digital spectral correlators that measure cross-correlation as a function of time delay are described in Sect. 8.8. In a lag-type correlator, the cross-correlation is measured as a function of time offset, implemented by introducing instrumental delays. The Fourier transform of the cross-correlation as a function of relative time delay between the signals is the cross-correlation as a function of frequency, as required in spectral line measurements. As mentioned in Sect. 6.1.7, it is necessary to use only simple correlators for this measurement. The range of time offsets of the two signals covers both positive and negative values, and the resulting measurements of cross-correlation contain both even and odd components. Fourier transformation then provides both the real and imaginary components of the cross-correlation as a function of frequency. The full sensitivity is obtained so long as the range of time offsets is comparable to the reciprocal signal bandwidth or greater; see Table 9.7. Note that in Table 6.1, we have not included the quantization loss discussed in Sect. 8.3.3. A demonstration of the sensitivity using a simple correlator when the measurements are made as a function of time delay is given by Mickelson and Swenson (1991).

Of the cases included in Table 6.1, the SSB with complex correlator is the one generally used where possible, because of the sensitivity and avoidance of the complications of DSB operation. Cases 2 and 3 in the table are included mainly for completeness of the discussion. As mentioned earlier, for frequencies of several hundred gigahertz, the most sensitive type of receiver input stage may be an SIS mixer. This has an inherently double-sided response, and if necessary, a sideband can be removed by filtering or using a sideband-separating arrangement (Appendix 7.1). For DSB operation, the most important cases in Table 6.1 are 6a and 6b. The case in which the unwanted sideband is only partially rejected is discussed in Appendix 6.1.

6.2.6 System Temperature Parameter α

As already noted, DSB systems are mainly used at millimeter and submillimeter wavelengths, at which the receiver input stage is commonly a cooled SIS mixer. Such a system can be converted to SSB operation by filtering out the unwanted sideband and terminating the corresponding input in a cold load. If the atmospheric losses are high and the receiver temperature is low, most of the system noise will come from the antenna, and terminating one sideband in a cold load will approximately halve the level of noise within the receiver. The system temperature

of the SSB system will then be approximately equal to the *double*-sideband system temperature of the DSB system, and the value of α [defined in Eq. (6.65)] tends toward 1. On the other hand, if atmospheric and antenna losses are low and most of the system noise comes from the mixer and IF stages, then terminating one sideband input in a cold load rather than the cold sky makes little difference to the noise level in the receiver. The system temperature of the SSB system will be close to the *single*-sideband system temperature of the DSB system, which is twice the DSB value. The value of α then tends toward $1/2$. To recapitulate: If the atmospheric noise dominates the receiver noise, then α tends toward 1, but if the receiver noise dominates, then α tends toward $1/2$. Note, however, that α is not confined to the range $1/2 < \alpha < 1$. For example, if noise from the antenna is low but the termination of the image sideband in the SSB system is uncooled and injects a high noise level, then α can be $< 1/2$. If the front end is tuned close to an atmospheric absorption line in such a way that the additional sideband of the DSB system falls in a frequency range of enhanced atmospheric noise, then α can be > 1 .

6.3 Effect of Bandwidth

As seen in the preceding section, the sensitivity of a receiving system to a broadband cosmic signal increases with the system bandwidth. Here we are concerned with the effect of bandwidth on the angular range over which fringes are detected, and on the fringe amplitude. These effects result from the variation of fringe frequency, in cycles per radian on the sky, with the received radio frequency. If the monochromatic response is integrated over the bandwidth, the fringes are reinforced for directions close to that for which the time delays from the source to the correlator inputs are equal, but for other directions, the fringes vary in phase across the bandwidth. This effect, when measured in a plane containing the interferometer baseline, causes the fringe amplitude to decrease with angle in a manner similar to that caused by the antenna beams (Swenson and Mathur 1969) and is sometimes referred to as the *delay beam*. It can be used to confine the response of an interferometer to a limited area of the sky and thereby reduce the possibility of source confusion, which can occur when the fringe patterns of two or more sources are recorded simultaneously. Examples of such usage can be found in some early interferometers built for operation at frequencies below 100 MHz (Goldstein 1959; Douglas et al. 1973).

6.3.1 Imaging in the Continuum Mode

The effect of bandwidth on the fringe amplitude was discussed in Sect. 2.2. Equation (2.3) gives an expression for the fringes observed for a point source with an east–west baseline of length D and a rectangular signal passband of width $\Delta\nu$.

The fringe amplitude is proportional to a factor

$$R'_b = \frac{\sin(\pi D l \Delta v / c)}{\pi D l \Delta v / c} . \quad (6.69)$$

Consider an array for which D is typical of the longest baselines. The synthesized beamwidth of the array, θ_b , is approximately equal to $\lambda_0 / D = c / \nu_0 D$, where ν_0 is the observing frequency and λ_0 the corresponding wavelength. (Note that in this section, ν_0 is the center frequency of the RF input band, not an IF band.) Thus, Eq. (6.69) becomes

$$R'_b \simeq \frac{\sin(\pi \Delta v l / \nu_0 \theta_b)}{\pi \Delta v l / \nu_0 \theta_b} . \quad (6.70)$$

The parameter $\Delta v l / \nu_0 \theta_b$ is equal to the fractional bandwidth multiplied by the angular distance of the source from the (l, m) origin measured in beamwidths. If this parameter is equal to unity, $R'_b = 0$ and the measured visibility is reduced to zero. To keep R'_b close to unity, we require $\Delta v l / \nu_0 \theta_b \ll 1$. Thus, to avoid underestimation of the visibility at long baselines, there is a limit on the angular size of the image that is inversely proportional to the fractional bandwidth.

We now examine the same effect in more detail by considering the distortion in the synthesized image. First, recall that the response of an array can be written as

$$\mathcal{V}(u, v) W(u, v) \longleftrightarrow I(l, m) * * b_0(l, m) , \quad (6.71)$$

where \longleftrightarrow represents Fourier transformation, and the double asterisk represents convolution in two dimensions. The fringe visibility is multiplied by $W(u, v)$, the spatial sensitivity function of the array for a particular observation. The Fourier transform of the left side of Eq. (6.71) gives the intensity distribution $I(l, m)$ convolved with the synthesized beam function $b_0(l, m)$. For simplicity, we have omitted the primary antenna beam and minor effects related to use of the discrete Fourier transform. The synthesized beam is defined here as the Fourier transform of $W(u, v)$.

In operation in the continuum mode, the visibility data measured with bandwidth Δv are treated as though they were measured with a single-frequency receiving system tuned to the center frequency, ν_0 . Thus, for all frequencies within the bandwidth, the assigned values of u and v are those appropriate to frequency ν_0 . At another frequency ν within the passband, the true spatial frequency coordinates u_ν and v_ν are related to the assigned values u and v by

$$(u, v) = \left(\frac{u_\nu \nu_0}{\nu}, \frac{v_\nu \nu_0}{\nu} \right) . \quad (6.72)$$

The contribution to the measured visibility from a narrow band of frequencies centered on ν is

$$\mathcal{V}_\nu(u, v) = \mathcal{V}_\nu\left(\frac{u\nu\nu_0}{\nu}, \frac{v\nu\nu_0}{\nu}\right) \longleftrightarrow \left(\frac{\nu}{\nu_0}\right)^2 I\left(\frac{l\nu}{\nu_0}, \frac{m\nu}{\nu_0}\right), \quad (6.73)$$

where we have used the similarity theorem of Fourier transforms.⁶ Thus, the contribution to the measured intensity is the true intensity distribution scaled in (l, m) by a factor ν/ν_0 and in intensity by $(\nu/\nu_0)^2$. The derived intensity distribution is convolved with $b_0(l, m)$, the synthesized beam corresponding to frequency ν_0 . We have assumed that the beam does not vary significantly with frequency and have used the same spacial sensitivity function $W(u, v)$ to represent the whole frequency passband. The overall response is obtained by integrating over the passband with appropriate weighting and is

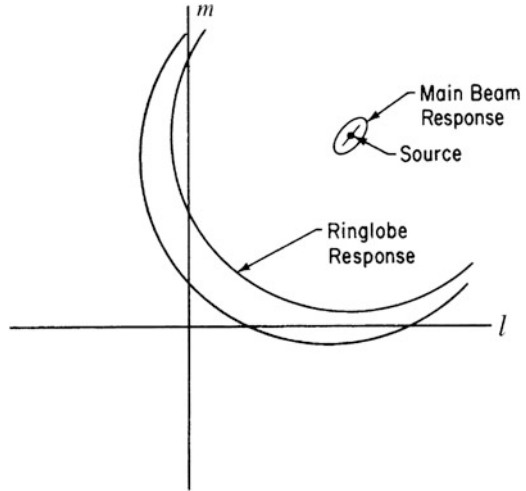
$$I_b(l, m) = \left[\frac{\int_0^\infty \left(\frac{\nu}{\nu_0}\right)^2 |H_{\text{RF}}(\nu)|^2 I\left(\frac{l\nu}{\nu_0}, \frac{m\nu}{\nu_0}\right) d\nu}{\int_0^\infty |H_{\text{RF}}(\nu)|^2 d\nu} \right] ** b_0(l, m). \quad (6.74)$$

Note that the integrals must be taken over the whole radio-frequency passband, denoted by the subscript RF, which includes both sidebands in the case of a DSB system. We assume that the passband function $H_{\text{RF}}(\nu)$ is identical for all antennas. The values of l and m in the intensity function in Eq. (6.74) are multiplied by the factor ν/ν_0 , which varies as we integrate over the passband, being equal to unity at the band center. Thus, one can envisage the integrals in the square brackets in Eq. (6.74) as resulting in a process of averaging a large number of images, each with a different scale factor. The scale factors are equal to ν/ν_0 , and the range of values of ν is determined by the observing passband. The images are aligned at the origin, and thus the effect of the integration over frequency is to produce a radial smearing of the intensity distribution before it is convolved with the beam. The response to a point source at position (l, m) is radially elongated by a factor equal to $\sqrt{l^2 + m^2} \Delta\nu/\nu_0$. For distances from the origin at which the elongation is large compared with the synthesized beamwidth, features on the sky become attenuated by the smearing, so there is an effective limitation of the useful field of view. The measured intensity is the smeared distribution convolved with the synthesized beam.

Details of the behavior of the derived intensity distribution can be deduced from Eq. (6.74). For example, suppose that the beam contains a circularly symmetrical sidelobe at a large angular distance from the beam axis and that in an image, the response to a distant source causes the sidelobe to fall near the origin. Is the sidelobe broadened near the origin? Since the distant source is elongated, the sidelobe will be smeared in a direction parallel to that of a line joining the source and the origin, as

⁶If $f(x)$ has a Fourier transform $F(s)$, then $f(ax)$ has a Fourier transform $|a|^{-1}F(s/a)$ (Bracewell 2000).

Fig. 6.10 Radial smearing resulting from the bandwidth effect for a point source at (l_1, m_1) . The effects on the responses of the main beam and a ringlobe (i.e., a sidelobe of the form in Fig. 5.15) are shown.



shown in Fig. 6.10. It will be broadened near the origin but not at a point 90° around the sidelobe as measured from the source.

To estimate the magnitude of the suppression of distant sources, it is useful to calculate R_b , the peak response to a point source at a distance r_1 from the origin of the (l, m) plane, as a fraction of the response to the same source at the origin. Because the effect we are considering is a radial smearing, we need consider only the intensity along a radial line through the (l, m) origin, as shown in Fig. 6.11a. We use idealized parameters; the passband is represented by a rectangular function of width Δv and the synthesized beam by a circularly symmetrical Gaussian function of standard deviation $\sigma_b = \theta_b / \sqrt{8 \ln 2}$, where θ_b is the half-power beamwidth. For simplicity, the factor $(v/v_0)^2$ in the integral in the numerator of Eq. (6.74) is omitted. The convolution becomes a one-dimensional (radial) process, as shown in Fig. 6.11b. The radially elongated source is represented by a rectangular function from $r_1(1 - \Delta v/2v_0)$ to $r_1(1 + \Delta v/2v_0)$, normalized to unit area. The beam is represented by the function $e^{-r^2/2\sigma_b^2}$, which is normalized to unity on the beam axis. When the beam is centered on the source, as shown in Fig. 6.11, R_b is given by

$$\begin{aligned}
 R_b &= \frac{v_0}{r_1 \Delta v} \int_{r_1(1-\Delta v/2v_0)}^{r_1(1+\Delta v/2v_0)} e^{-(r-r_1)^2/2\sigma_b^2} dr \\
 &= \sqrt{2\pi} \frac{\sigma_b v_0}{r_1 \Delta v} \operatorname{erf} \left(\frac{r_1 \Delta v}{2\sqrt{2}\sigma_b v_0} \right) \\
 &= 1.0645 \frac{\theta_b v_0}{r_1 \Delta v} \operatorname{erf} \left(0.8326 \frac{r_1 \Delta v}{\theta_b v_0} \right). \tag{6.75}
 \end{aligned}$$

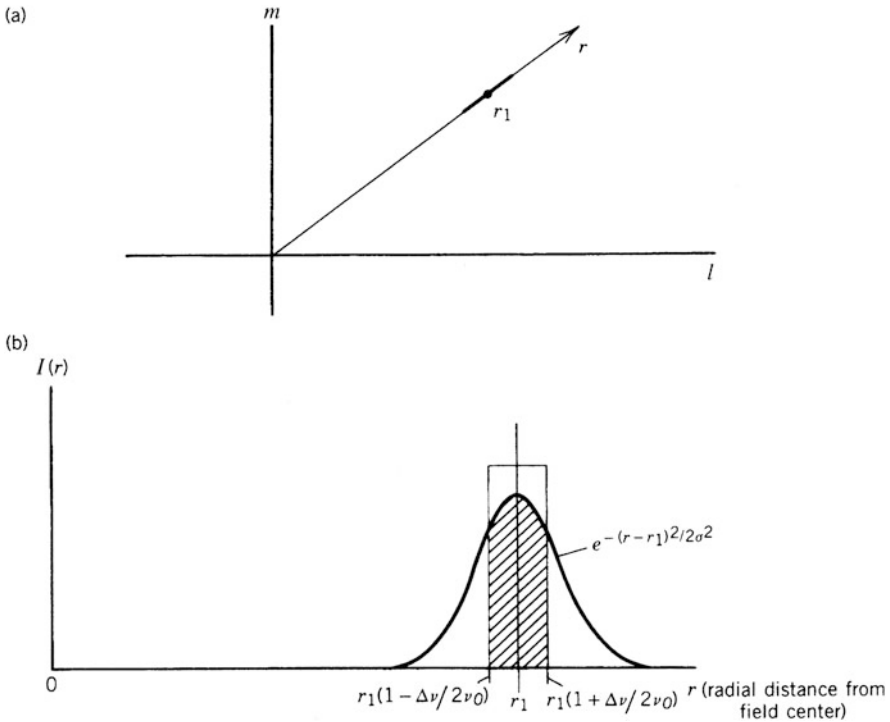


Fig. 6.11 Response of an array with a broadband receiving system to a point source at distance r_1 from the origin of the (l, m) plane. **(a)** The point source (delta function) at r_1 becomes radially broadened into a rectangular function of unit area indicated by the heavy line. **(b)** Cross section of the intensity distribution in the r direction. The synthesized beam is represented by the Gaussian function. The peak intensity of the response to the source is proportional to the shaded area.

A curve of R_b as a function of the parameter $r_1 \Delta\nu / \theta_b \nu_0$, which is the distance of the source from the origin measured in beamwidths, multiplied by the fractional bandwidth, is shown in Fig. 6.12. Values of 0.2 and 0.5 for this parameter reduce the response by 0.9% and 5.5%, respectively.

If the receiving passband is represented by a Gaussian function of equivalent width $\Delta\nu$ (i.e., standard deviation = $\Delta\nu/2.5066$), the reduction factor becomes

$$R_b = \frac{1}{\sqrt{1 + (0.939 r_1 \Delta\nu / \theta_b \nu_0)^2}} \tag{6.76}$$

A curve of this function is also included in Fig. 6.12. Comparison of the two curves illustrates the dependence on the passband shape.

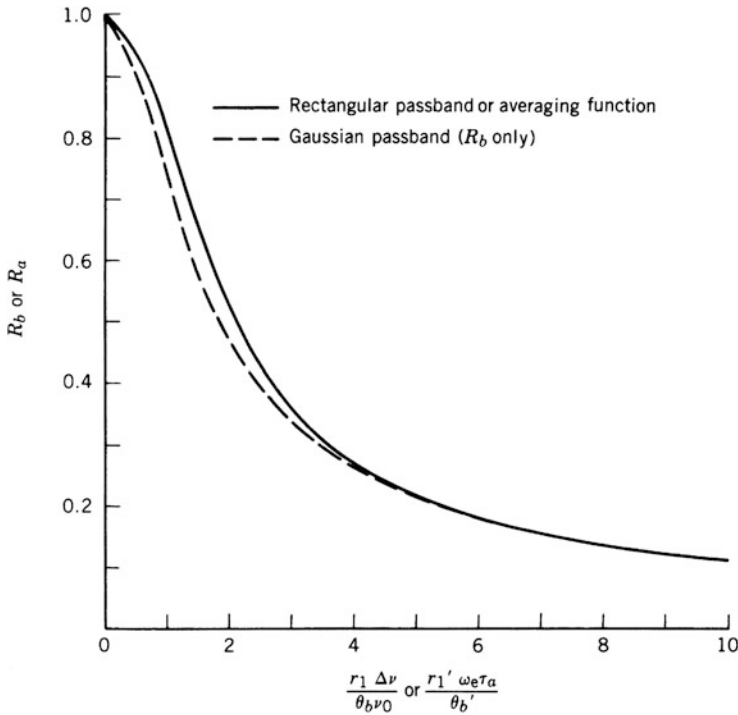


Fig. 6.12 Relative amplitude of the peak response to a point source as a function of the distance from the field center and either the fractional bandwidth or the averaging time.

6.3.2 Wide-Field Imaging with a Multichannel System

Broadband images can also be obtained by observing with a multichannel system (i.e., a spectral line system as described in Sect. 8.8.2). In this case, the passband is divided into a number of channels by using either a bank of narrowband filters or a multichannel digital correlator. The visibility is measured independently for each channel, so the values of u and v can be scaled correctly and an independent image obtained for each channel. This scaling causes the spatial sensitivity function to vary over the band, and at frequency ν , the synthesized beam is $(\nu/\nu_0)^2 b_0(l\nu/\nu_0, m\nu/\nu_0)$, where $b_0(l, m)$ is the monochromatic beam at frequency ν_0 . The images can be combined by summation, and if given equal weights, the result for N channels is represented by

$$I(l, m) ** \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{\nu_i}{\nu_0} \right)^2 b_0 \left(\frac{l\nu_i}{\nu_0}, \frac{m\nu_i}{\nu_0} \right) \right]. \tag{6.77}$$

In this case, there is no smearing of the intensity distribution, but the beam suffers a radial smearing that has the desirable effect of reducing distant sidelobes. Therefore, this mode of observation is well suited for imaging wide fields. The improvement in the beam results from the increase in the number of (u, v) points measured, an effect that is also used in multifrequency synthesis discussed in Sect. 11.6.

6.4 Effect of Visibility Averaging

6.4.1 Visibility Averaging Time

In most synthesis arrays, the output of each correlator is averaged for consecutive time periods, τ_a , and thus consists of real or complex values spaced at intervals τ_a in time. It is advantageous to make τ_a long enough to keep the data rate from the correlator readout conveniently small. An upper limit on τ_a results from a consideration of the sampling theorem discussed in Sect. 5.2.1 and is briefly explained as follows. In discrete Fourier transformation of the visibility to intensity, the data points are often spaced at intervals Δu and Δv , as shown in Fig. 5.3. If the size of the field to be imaged is θ_f in the l and m directions, then $\Delta u = \Delta v = 1/\theta_f$. In time τ_a , the motion of a baseline vector within the (u, v) plane should not be allowed to exceed Δu ; otherwise, the visibility values will not fully represent the angular variation of the brightness function.

Consider the case in which the longest baseline is east–west in orientation and the source under observation is at a high declination, which results in the fastest motion of the baseline vector. If the baseline length is D_λ wavelengths, the vector in the (u, v) plane traces out an approximately circular locus, the tip of which moves at a speed of $\omega_e D_\lambda$ wavelengths per unit time, where ω_e is the angular velocity of rotation of the Earth. Thus, we require that $\tau_a \omega_e D_\lambda < 1/\theta_f$, which results, in practice, in $\tau_a \approx C/(\omega_e D_\lambda \theta_f)$, where C is a factor likely to be in the range 0.1–0.5. Note that $D_\lambda \theta_f$ is approximately the number of synthesized beamwidths across the field, and thus τ_a must be somewhat smaller than the time taken for the Earth to rotate through one radian, divided by this number. Although shorter baselines could be averaged for longer times, in most synthesis arrays, all correlator outputs are read at the same time, at a rate appropriate for the longest baselines. Another consideration is that sporadic interference can be edited out of the data with minimal information loss if τ_a is not too long. For large arrays, τ_a is generally in the range of tens of milliseconds to tens of seconds. Determining the visibility at the $(\Delta u, \Delta v)$ grid points from the sampled data on the (u, v) loci is discussed in Sect. 10.2.3.

6.4.2 Effect of Time Averaging

We now examine in more detail the effect of the averaging on the synthesized intensity distribution. In reducing the data, all visibility values within each interval

τ_a are treated as though they applied to the time at the center of the averaging period. Thus, for example, the measurements at the beginning of each averaging period enter into the visibility data with assigned values of u and v that apply to times $\tau_a/2$ later than the true values. In effect, the resulting image consists of the average of a large number of images, each with a different timing offset distributed progressively throughout the range $-\tau_a/2$ to $\tau_a/2$. These timing offsets apply only to the assignment of (u, v) values and do not resemble a clock error that would affect the whole receiving system.

Consider an unresolved source, represented by a delta function. To simplify the situation, we consider observations with east–west baselines and examine the effects in the (u', v') plane and the corresponding (l', m') sky plane (see Sect. 4.2). The spacing loci are circular arcs generated by vectors rotating at angular velocity ω_e , as shown in Fig. 6.13a. Consider first the case of an east–west linear array; then, of the antenna spacing components (X, Y, Z) defined in Fig. 4.1, only Y is nonzero. The circular arcs of the spacing loci are centered on the (u', v') origin as in Fig. 6.13b, and a timing offset δt is equivalent to a rotation of the (u', v') axes through an angle $\omega_e \delta t$. The visibility of the source is the combination of two sets of sinusoidal corrugations, one real and one imaginary:

$$\delta(l'_1, m'_1) \longleftrightarrow \cos 2\pi(u'l'_1 + v'm'_1) - j \sin 2\pi(u'l'_1 + v'm'_1) . \tag{6.78}$$

The angle of the corrugations is related to the position angle $\psi' = \tan^{-1}(m'_1/l'_1)$ of the point source, as shown in Fig. 6.14. A change in ψ' causes an equivalent rotation of the corrugations and vice versa. For an east–west array, time offsets therefore correspond to proportional rotations of the intensity in the (l', m') plane. It

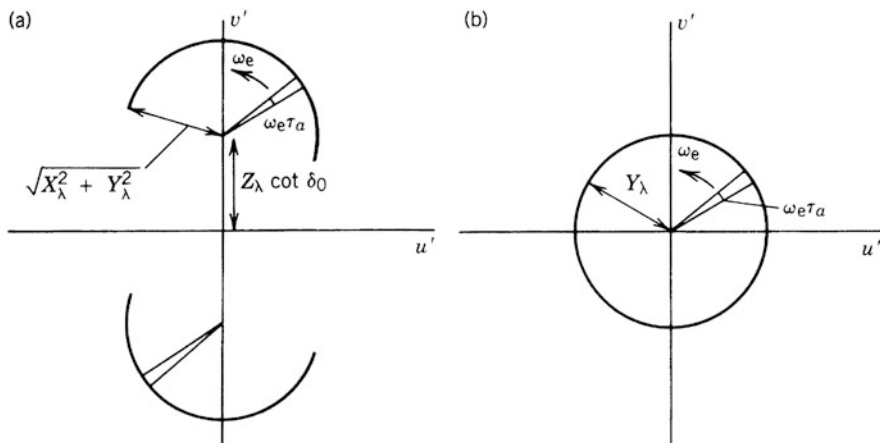


Fig. 6.13 Spacing loci in the (u', v') plane, (a) for the general case and (b) for an east–west baseline. The angle $\omega_e \tau_a$ over which the averaging takes place is enlarged for clarity: for example, with an averaging time of 30 s, the angle would be 7.5 arcmin.

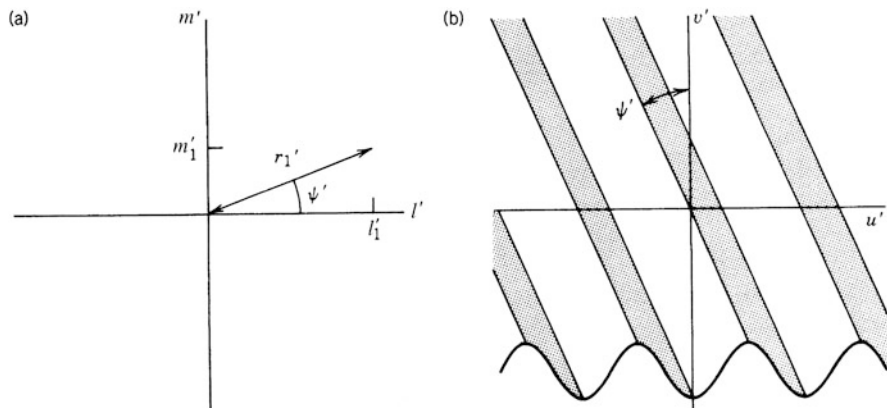


Fig. 6.14 (a) Point source at (l'_1, m'_1) and (b) the real part of the corresponding visibility function. The ridges of the sinusoidal corrugations that represent the visibility in the (u', v') plane are orthogonal to the radius vector r'_1 at the position of the source in the (l', m') plane.

follows that the effect of the time averaging is to produce a circumferential smearing similar to that resulting from the receiving bandwidth but orthogonal to it. If we express positions in the (l', m') plane in terms of the radial coordinates (r', ψ') shown in Fig. 6.14a, the image obtained from the averaged data can be expressed in terms of the sky brightness $I(r', \psi')$ by

$$I_a(r', \psi') = \left[\frac{1}{\omega_e \tau_a} \int_{-\omega_e \tau_a/2}^{\omega_e \tau_a/2} I(r', \psi') d\psi' \right] * * b_0(r', \psi'), \quad (6.79)$$

where b_0 is the synthesized beam.

The fractional decrease in the peak response to the point source is most easily considered in the (l', m') plane. With an east–west baseline, the contours of the synthesized beam are approximately circular in the (l', m') plane, as long as the observing time is approximately 12 h, which results in spacing loci in the form of complete circles in the (u', v') plane. If we assume that the synthesized beam can be represented by a Gaussian function, as in the calculations for the bandwidth effect, the curve for the rectangular bandwidth in Fig. 6.12 can also be used for the averaging effect. In one case, the spreading function is radial and of width $r_1 \Delta v / v_0$, and in the other, it is circumferential and of width $r'_1 \omega_e \tau_a$. Thus, for the averaging effect, we can replace $r_1 \Delta v / \theta_b v_0$ in Eq. (6.75) and Fig. 6.12 (solid curve) by $r'_1 \omega_e \tau_a / \theta'_b$, noting that $r'_1 = \sqrt{l_1^2 + m_1^2} \sin^2 \delta_0$ and θ'_b , the synthesized beamwidth in the (l', m') plane, is equal to the east–west beamwidth in the (l, m) plane. Hence, for the decrease in the response to a point source resulting from averaging, using

Eq. (6.75), we can write

$$R_a = 1.0645 \frac{\theta'_b}{r'_1 \omega_e \tau_a} \operatorname{erf} \left(0.8326 \frac{r'_1 \omega_e \tau_a}{\theta'_b} \right). \quad (6.80)$$

Generally, one chooses τ_a so that R_a is only slightly less than unity at any point in the image, in which case we can approximate the error function by the integral of the first two terms in the power series for a Gaussian function:

$$R_a \simeq 1 - \frac{1}{3} \left(\frac{0.8326 \omega_e \tau_a}{\theta'_b} \right)^2 (l_1^2 + m_1^2 \sin^2 \delta_0). \quad (6.81)$$

This formula can be used for checking that τ_a is not too large.

Two aspects of the behavior predicted by Eq. (6.81) should be mentioned. First, if the source is near the m' axis and at a low declination, the averaging has very little effect. This is because the ridges of the sinusoidal corrugations of the visibility function then run approximately parallel to the u' axis, and in the transformation $u' = u \operatorname{cosec} \delta_0$, the period of the variations in the v direction is expanded by a large factor. In comparison, the arc through which any baseline vector moves in time τ_a is small, and hence, the averaging has only a small effect on the visibility amplitude. Second, for a source on the l' axis, R_a is independent of δ_0 . In this case, the ridges of the corrugations run parallel to the v axis, and the expansion of the scale in the v direction has no effect on the sinusoidal period.

For arrays that contain baselines other than east–west, the centers of the corresponding loci in the (u', v') plane are offset from the origin, as in Fig. 6.13a, and a time offset is no longer equivalent to a simple rotation of axes. However, this may not increase the smearing of the visibility, so the effect may be no worse than for an east–west array with baselines of similar lengths.

6.5 Speed of Surveying

The requirement to maximize the efficiency of use of large instruments requires consideration of the best procedures for surveying, i.e., searching large areas of the sky for radio sources of various types including transient sources. In the frequency range below about 2 GHz, four of the key science applications of the proposed Square Kilometre Array that require imaging of a significant fraction of the sky are as follows.

1. Searching for pulsars in binary combinations with neutron stars or black holes, for gravitational studies.
2. Measurement of Faraday rotation in very large numbers of radio galaxies to determine the structure of galactic and intergalactic magnetic fields.

3. Imaging of very large numbers of HI galaxies out to redshifts of $z \simeq 1.5$ to study galactic evolution and provide further constraints on the nature of dark energy.
4. Detection of transient events such as afterglows of gamma-ray bursts.

The choice of parameters for optimization of speed in survey observations is not the same as for optimization of sensitivity in targeted studies (Bregman 2005). Consider first the case of targeted observations of individual continuum sources that have angular dimensions small compared with a station⁷ beam. We can adapt the expression for the rms noise [Eq. (6.62)] as a measure of the minimum detectable flux density S_{\min} observed with two stations in a time τ :

$$S_{\min} = \frac{2kT_s}{A\sqrt{(\Delta\nu\tau)}}, \quad (6.82)$$

where A is here the collecting area of a station, equal to $A\eta_Q$ of Eq. (6.62). For an array with n_s stations, there are $n_s(n_s - 1)/2 \approx n_s^2/2$ correlated pairs of signals, so the right side of Eq. (6.82) is multiplied by a factor $\sqrt{2}/n_s$. The observing speed, i.e., the number of observations per unit time, is

$$1/\tau = \frac{A^2\Delta\nu S_{\min}^2 n_s^2}{8k^2 T_s^2}. \quad (6.83)$$

Next, consider the case of a survey in which we are concerned with the speed of coverage of a specified solid angle of sky down to a sensitivity level S_{\min} . Since A is the area of a station, the solid angle of a station beam is λ^2/A sr, where λ is the wavelength. If each station forms n_{sb} simultaneous beams, the instantaneous field of view is

$$F_v = \lambda^2 n_{sb}/A. \quad (6.84)$$

The reciprocal of the time τ required to cover solid angle F_v down to flux density level S_{\min} is given by Eq. (6.83), and the corresponding survey speed is

$$F_v/\tau = \frac{\lambda^2 A \Delta\nu S_{\min}^2 n_s^2 n_{sb}}{8k^2 T_s^2} \text{ sr per unit time}. \quad (6.85)$$

For surveys to detect spectral line features, $\Delta\nu$ in Eqs. (6.83) and (6.85) represents the bandwidth of the line. Then, if it is necessary to search in frequency, the

⁷For an array with a large total collecting area, it may be more practical to use a large number of small antennas rather than a smaller number of large antennas. To limit the number of cross-correlated signal pairs, the small antennas are located in groups. These groups are commonly referred to as *stations*, and the signals from the antennas at a station are combined to provide a number of beams within the main beam of the individual antennas. For pairs of stations, the signals from corresponding beams are cross-correlated to provide the visibility data.

bandwidth of the receiving system can be included as an additional factor in the expression for the speed in Eq. (6.85).

Comparison of Eqs. (6.83) and (6.85) shows the effect of the field-of-view dependence in the survey case. The survey speed is proportional to the number of simultaneous beams and is less strongly dependent on the station aperture area A . The wavelength-squared factor results from the increased beamwidth with decreasing frequency, but the benefit of lower frequency (increasing λ) on the survey speed applies only so long as the effect of the galactic background radiation on the system temperature is small. From the galactic background model of Dulk et al. (2001), the brightness temperature in the range 10–1,000 MHz is approximately proportional to $\nu^{-2.5}$, so for frequencies at which this background is the dominant contributor to T_s , the frequency dependence of the survey speed is approximately proportional to ν^3 . For directions that are not close to the galactic plane, the background temperature is about 20 K at 500 MHz and 2 K at 1 GHz, so if the receiver contribution to T_s is ~ 20 K, there is a broad maximum in survey speed between these two frequencies.

Note that the discussion above involves the assumption that the sensitivity is limited only by the system noise. If dynamic range is the limiting factor, then the density of the (u, v) coverage, which improves with increasing n_s and τ , may become the most important consideration. In either case, performance improves with increasing number of stations.

Survey speed can be increased by increasing the number of stations as well as the number of station beams. However, the size of the correlator system for the full array is proportional to n_s^2 and to n_{sb} , so increasing the number of stations, or the number of station beams, requires an increase in the size of the correlator. Increasing the station aperture A is likely to require adding more antennas to the station subarray and thus increases the station beamforming hardware. The only way of increasing the observing speed that does not increase the signal-processing requirements is reducing the system temperature T_s . However, the complexity of phased-array feeds for the formation of multiple beams from a single parabolic antenna can degrade the system temperature. If cryogenic cooling is necessary, the required cooling capacity is considerably greater for multiple-beam systems than for single-beam ones. Thus, optimization of the array performance for a given overall cost requires a broad consideration of the performance of various parts of the receiving system.

Appendix 6.1 Partial Rejection of a Sideband

In an SSB system using a mixer as the input stage, the unwanted (image) sideband may be rejected by one of several schemes. These include use of a waveguide filter, a Martin–Puplett interferometer (Martin and Puplett 1969; Payne 1989), a tuned backshort, or a sideband-separating configuration of two mixers (as in Appendix 7.1). Practical considerations, particularly at millimeter wavelengths, can limit the effectiveness of the rejection of the image sideband. Let the response to the

Fig. A6.1 Vectors in the complex plane representing the parameters in Eqs. (A6.1) and (A6.2). The constant gain factor G_{mn} is omitted. If ρ is known and C_0 and $C_{\pi/2}$ are measured, Eq. (A6.3) gives the optimum estimate of \mathcal{V} .

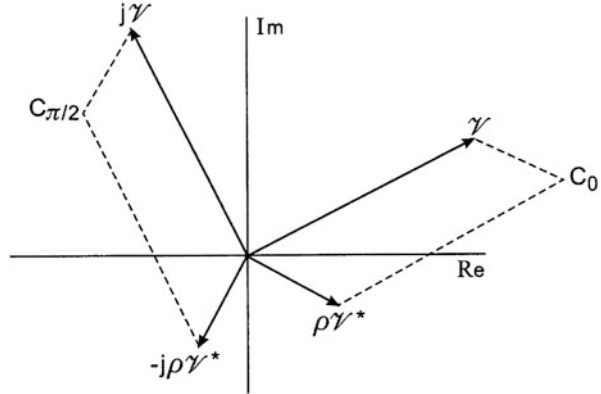


image sideband, in terms of the power gain of the receiver, be ρ times the response to the wanted (signal) sideband, where $0 < \rho < 1$. In practice, ρ could be as large as $\sim 1/10$.

In the case of spectral line observation, where the wanted line occurs only in the signal sideband, the effect of the noise introduced by the image sideband is to increase the rms noise at the correlator output by a factor $(1 + \rho)$. Thus, the sensitivity is correspondingly reduced (Fig. A6.1).

In the case of a continuum observation, the image sideband also introduces a component of signal and noise at the correlator. Assume that the visibility is the same in both sidebands, the fringes are stopped, and $\pi/2$ phase switching of the first LO allows measurement of the complex visibility. A complex correlator is used, and for simplicity, we consider that the instrumental phase is adjusted so that the line AB in Fig. 6.5b is coincident with the real axis. We can represent the complex correlator output with zero phase shift of the LO as

$$C_0 = G_{mn}(\mathcal{V} + \rho\mathcal{V}^*), \quad (\text{A6.1})$$

and with the $\pi/2$ phase switch as

$$C_{\pi/2} = G_{mn}(j\mathcal{V} - j\rho\mathcal{V}^*). \quad (\text{A6.2})$$

Here, G_{mn} is the gain in the signal sideband, so ρG_{mn} is the gain in the image sideband. Note that in the expression for $C_{\pi/2}$, the j factors have opposite signs for the two sidebands, because the $\pi/2$ phase shift causes the corresponding vectors in the complex plane to rotate through $\pi/2$ in opposite directions, as in Fig. A6.1. The optimum estimate of the visibility is then found to be

$$\mathcal{V} = \frac{1}{2G_{mn}} \left[\frac{1}{1 + \rho^2} (C_0 - jC_{\pi/2}) + \frac{\rho}{1 + \rho^2} (C_0 + jC_{\pi/2})^* \right]. \quad (\text{A6.3})$$

The first term within the square brackets represents the response of the signal sideband, and the second term represents the response of the image. The total noise power delivered to the correlator input [i.e., the sum of the two terms in the square brackets in Eq. (A6.3)] is proportional to $(1 + \rho)$, so the noise associated with the first term in the square brackets is proportional to $(1 + \rho)/(1 + \rho^2)$. Similarly, for the second term, the associated noise is proportional to $\rho(1 + \rho)/(1 + \rho^2)$. Thus, the total noise in the estimate of \mathcal{V} from Eq. (A6.3) is proportional to $(1 + \rho)^2/(1 + \rho^2)$. In terms of the rms noise, the sensitivity is proportional to the square root of the reciprocal of the last term, i.e., $\sqrt{(1 + \rho^2)/(1 + \rho)}$. For $\rho \approx 1/10$ or less, the ρ^2 term is very small, and the sensitivity degradation factor is approximately $(1 + \rho)^{-1}$ (Thompson and D’Addario 2000).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.



References

- Abramowitz, M., and Stegun, I.A., Eds., *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, DC (1968)
- Anantharamaiah, K.R., Ekers, R.D., Radhakrishnan, V., Cornwell, T.J., and Goss, W.M., Noise in Images of Very Bright Sources, in *Synthesis Imaging in Radio Astronomy*, Perley, R.A., Schwab, F.R., and Bridle, A.H., Eds., Astron. Soc. Pacific Conf. Ser., **6**, 431–442 (1989)
- Blum, E.J., Sensibilité des Radiotélescopes et Récepteurs à Corrélation, *Ann. Astrophys.*, **22**, 140–163 (1959)
- Bracewell, R.N., *The Fourier Transform and Its Applications*, McGraw-Hill, New York (2000) (also see earlier eds. 1965, 1978)
- Bregman, J.D., System Optimization of Multibeam Aperture Synthesis Arrays for Survey Performance, in *The Square Kilometre Array: An Engineering Perspective*, Hall, P.J., Ed., Springer, Dordrecht (2005)
- Colvin, R.S., “A Study of Radio Astronomy Receivers,” Ph.D. thesis, Stanford Univ., Stanford, CA (1961)
- Douglas, J.N., Bash, F.N., Ghigo, F.D., Moseley, G.F., and Torrence, G.W., First Results from the Texas Interferometer: Positions of 605 Discrete Sources, *Astron. J.*, **78**, 1–17 (1973)
- Dulk, G.A., Erickson, W.C., Manning, R., and Bougeret, J.-L., Calibration of Low-Frequency Radio Telescopes Using Galactic Background Radiation, *Astron. Astrophys.*, **365**, 294–300 (2001)
- Goldstein, S.J., Jr., The Angular Size of Short-Lived Solar Radio Disturbances, *Astrophys. J.*, **130**, 393–399 (1959)

- Lawson, J.L., and Uhlenbeck, G.E., *Threshold Signals*, Radiation Laboratory Series, Vol. 24, McGraw-Hill, New York (1950), p. 68
- Lo, W.F., Dewdney, P.E., Landecker, T.L., Routledge, D., and Vaneldik, J.F., A Cross-Correlation Receiver for Radio Astronomy Employing Quadrature Channel Generation by Computed Hilbert Transform, *Radio Sci.*, **19**, 1413–1421 (1984)
- Martin, D.H., and Puplett, E., Polarized Interferometric Spectrometry for the Millimeter and Submillimeter Spectrum, *Infrared Phys.*, **10**, 105–109 (1969)
- Mickelson, R.L., and Swenson, G.W., Jr., A Comparison of Two Correlation Schemes, *IEEE Trans. Instrum. Meas.*, **IM-40**, 816–819 (1991)
- Middleton, D., *An Introduction to Statistical Communication Theory*, McGraw-Hill, New York (1960), p. 343
- Moran, J.M., Very Long Baseline Interferometric Observations and Data Reduction, in *Methods of Experimental Physics*, Vol. 12, Part C (*Astrophysics: Radio Observations*), Meeks, M. L., Ed., Academic Press, New York (1976), pp. 228–260
- Papoulis, A., *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York (1965)
- Payne, J.M., Millimeter and Submillimeter Wavelength Radio Astronomy, *Proc. IEEE*, **77**, 993–1017 (1989)
- Read, R.B., Two-Element Interferometer for Accurate Position Determinations at 960 Mc, *IRE Trans. Antennas Propag.*, **AP-9**, 31–35 (1961)
- Rice, S.O., Mathematical Analysis of Random Noise, *Bell Syst. Tech. J.*, **23**, 282–332 (1944); **24**, 46–156, 1945 (repr. in *Noise and Stochastic Processes*, Wax, N., Ed., Dover, NY, 1954)
- Rogers, A.E.E., Theory of Two-Element Interferometers, in *Methods of Experimental Physics*, Vol. 12, Part C (*Astrophysics: Radio Observations*), Meeks, M.L., Ed., Academic Press, New York (1976), pp. 139–157 (see Table 1)
- Swenson, G.W., Jr., and Mathur, N.C., On the Space-Frequency Equivalence of a Correlator Interferometer, *Radio Sci.*, **4**, 69–71 (1969)
- Thompson, A.R., and D'Addario, L.R., Relative Sensitivity of Double- and Single-Sideband Systems for Both Total Power and Interferometry, ALMA Memo 304, National Radio Astronomy Observatory (2000)
- Tiuri, M.E., Radio Astronomy Receivers, *IEEE Trans. Antennas Propag.*, **AP-12**, 930–938 (1964)
- Tucker, J.R., and Feldman, M.J., Quantum Detection at Millimeter Wavelengths, *Rev. Mod. Phys.*, **57**, 1055–1113 (1985)
- Vander Vorst, A.S., and Colvin, R.S., The Use of Degenerate Parametric Amplifiers in Interferometry, *IEEE Trans. Antennas Propag.*, **AP-14**, 667–668 (1966)
- M. Vinokur, Optimisation dans la Recherche d'une Sinusoïde de Période Connue en Présence de Bruit, *Ann. d'Astrophys.*, **28**, 412–445 (1965)
- Wozencraft, J.M., and Jacobs, I.M., *Principles of Communication Engineering*, Wiley, New York (1965) p. 205
- Wright, M.C.H., Clark, B.G., Moore, C.H., and Coe, J., Hydrogen-Line Aperture Synthesis at the National Radio Astronomy Observatory: Techniques and Data Reduction, *Radio Sci.*, **8**, 763–773 (1973)