1977

A statistical analysis of ionospheric effects in radio interferometry

Ali Okatan
Iowa State University

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again – beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

A statistical analysis of ionospheric effects in radio interferometry

by

Ali Okatan

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1977

TABLE OF CONTENTS

I. INTRODUCTION

A. Radio Interferometry

In the last decade, one of the most significant achievements in radio astronomy was the development of the VLBI (Very-Long-Baseline-Interferometry) technique. With this new technique, it has become possible to study the universe more thoroughly.

The radio interferometer is not new in radio astronomy, however. The first interferometer was introduced in the late 1940's by Ryle and his coworkers in England (Fomalont and Wright, 1974). Since then radio interferometry has been developing to reach its full potential.

There are many types of radio interferometers, all based on the same principles. They differ more in constructional details than in principle.

In the beginning interferometers and lunar occultations were used to study discrete radio sources.

Today, radio interferometry is used in many different areas, e.g. in geodesy, navigation, and satellite trackings (Salzberg, 1973; Burke, 1969; Counselman, 1973).

In principle a radio interferometer is simply an array of antennas which are used to pick up the radiation from so-called radio sources. The signals from the antenna outputs are combined and processed appropriately. The type of combination depends on the type of the interferometer, e.g., in a correlation interferometer the signals from antenna elements are cross-correlated to obtain the correlated power (Cohen, 1973).

There are many excellent review articles on the history, development, and practical applications of radio interferometry (Fomalont and Wright, 1974; Meeks, 1976, Chapter 5; Cohen, 1973; Fomalont, 1973; Burke, 1969; Moran, 1973). Before giving the theory of the two-element interferometer, a brief review of coherence theory will be given in the next section.

B. The Coherence Theory

In this section a brief description of the coherence theory will be given. The theory of coherence is important for an understanding of what an interferometer measures. The subject is very complex and the full treatment is not the purpose of this section. The reader is referred to the literature for details of the theory (Ko, 1967; Carter and Somers, 1976).

Electromagnetic fields encountered in radio astronomy vary rapidly and irregularly. We characterize these fields by their frequency spectra, polarizations, and average powers. To describe these fields, let $E_1(P_1,t_1)$ and $E_2(P_2,t_2)$ be two similar electric-field components of a wave at two points P_1 and P_2, respectively, oriented along the z-axis as shown in Figure 1.b.1.

The incident wave at P_1 or P_2 may be considered to be a superposition of plane waves represented by complex phasors $E(f_x,f_y,t)$ coming from different directions which form an angular spectrum of plane waves (Booker and Clemmow, 1950). The arguments f_x and f_y are the direction cosines of the wave normal with respect to the x and y axes, respectively.

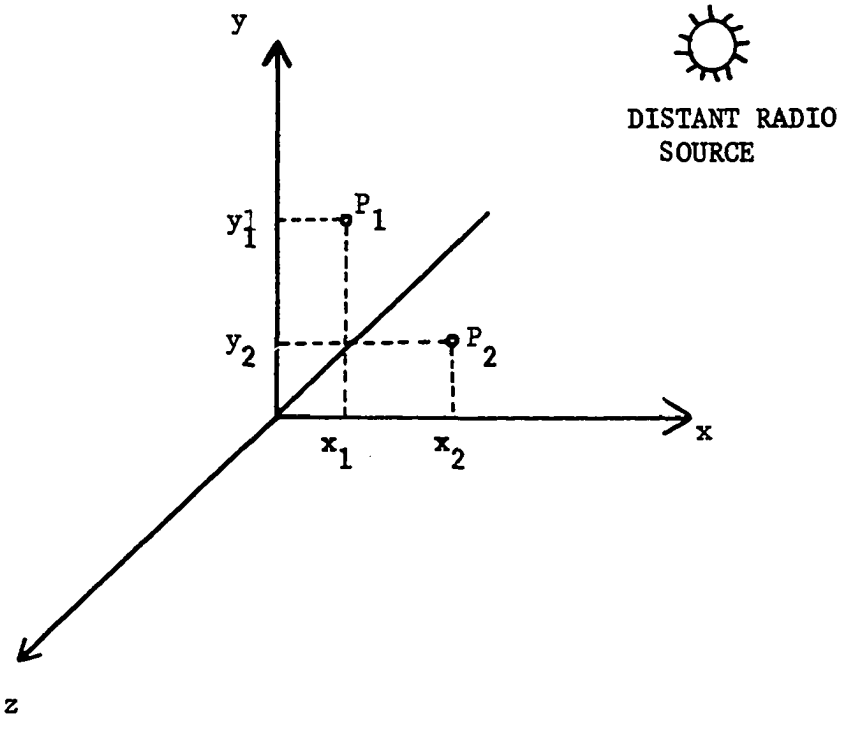


Figure 1.b.1. The geometry of P_1, P_2, and the radio source

It can be shown that the field distribution on the x - y plane and the angular spectrum have the relation (Booker and Clemmow, 1950)

$$E(\frac{x}{\lambda}, \frac{y}{\lambda}, t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} E(f_x, f_y, t) \cdot \exp\left[j \frac{2\pi}{\lambda}(xf_x + yf_y)\right] df_x df_y \qquad (1.B.1)$$

where λ is the wavelength. Since the solid angle formed by an angular element $df_x df_y$ is $df_x df_y / f_z$, the brightness of the source is proportional to $f_z \langle |E(f_x, f_y, t)|^2 \rangle$ in the direction of (f_x, f_y). f_z is the direction cosine of the wave normal with respect to the z-axis, and t is the time variable.

If we consider the case for which $f_z \simeq 1$, then the brightness distribution of the source is $\langle |E(f_x, f_y, t)|^2 \rangle$. The angular brackets denote ensemble averaging.

Now, let us assume that we have two spaced antennas on the x - y plane, each situated at (x_1, y_1) and $(x_2 = x_1 + \Delta x,\ y_2 = y_1 + \Delta y)$ coordinates, respectively, as shown in Fig. 1.b.1. Each antenna output is the voltage proportional to $E(x_1, y_1, t)$ and $E(x_2, y_2, t)$, respectively. If we take the proportionality constants as unity, we have

$$V\left(\frac{x_1}{\lambda} + u,\ \frac{y_1}{\lambda} + v, t\right) = E\left(\frac{x_1}{\lambda} + u,\ \frac{y_1}{\lambda} + v, t\right), \tag{1.B.2}$$

$$V\left(\frac{x_1}{\lambda},\ \frac{y_1}{\lambda}, t\right) = E\left(\frac{x_1}{\lambda},\ \frac{y_1}{\lambda}, t\right), \tag{1.B.3}$$

where $u = \Delta x/\lambda$, $v = \Delta y/\lambda$.

Using the Fourier transform of Eq. (1.B.1) and the Wiener-Khinchine theorem (Ko, 1967) one finds

$$\begin{aligned} &\left\langle V\left(\frac{x_1}{\lambda} + u,\ \frac{y}{\lambda} + v, t\right) V^*\left(\frac{x_1}{\lambda},\ \frac{y_1}{\lambda}, t\right)\right\rangle \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle |E(f_x, f_y, t)|^2 \rangle \\ &\cdot \exp[2\pi j(uf_x + vf_y)]\, df_x\, df_y. \end{aligned} \tag{1.B.4}$$

where the asterisk denotes the complex conjugate, $j = \sqrt{-1}$. Eq. (1.B.4) is a very important equation because it relates the quantity measured by an interferometer to the brightness distribution of a radio source. The quantity $\left\langle V\left(\frac{x_1}{\lambda} + u,\ \frac{y}{\lambda} + v, t\right) V^*\left(\frac{x_1}{\lambda},\ \frac{y_1}{\lambda}, t\right)\right\rangle$ is usually denoted by $\Gamma_{12}(u,v)$ and called the <u>partial coherence</u> function.

In practice, the angular brackets in Eq. (1.B.4) are replaced by the time average operators since we do not have an infinite number of stochastic field functions.

Using time averaging instead of the ensemble averaging implies that we are assuming the electromagnetic waves from a radio source are ergodic (Carter and Somers, 1976).

Now we define a quantity, which is called complex visibility by radio astronomers and complex degree of coherence in the area of coherence theory. The complex visibility is defined as (Cronyn, 1972)

$$\gamma(u,v) \triangleq \frac{\langle V(\frac{x_1}{\lambda}+u,\ \frac{y_1}{\lambda}+v,t)\, V^*(\frac{x_1}{\lambda},\ \frac{y_1}{\lambda},t)\rangle}{\langle |V(\frac{x_1}{\lambda}+u,\ \frac{y_1}{\lambda}+v,t)|^2\rangle^{\frac{1}{2}} \langle |V(\frac{x_1}{\lambda},\ \frac{y_1}{\lambda},t)|^2\rangle^{\frac{1}{2}}} \qquad (1.B.5)$$

Using Schwartz's inequality , it can be shown that

$$0 \le |\gamma(u,v)| \le 1. \qquad (1.B.6)$$

The complex visibility, which is the quantity measured by an interferometer, is a complex function of u and v. We can express $\gamma(u,v)$ as

$$\gamma(u,v) = |\gamma(u,v)| \exp[j\Phi(u,v)], \qquad (1.B.7)$$

where $|\gamma(u,v)|$ and $\Phi(u,v)$ are called the amplitude and phase of the complex visibility function, respectively.

Let us express Eq. (1.B.5) in a more practical form. Using time averages instead of the ensemble averages in Eq. (1.B.5) we obtain

$$\hat{\gamma}(u,v) = \frac{\overline{V_1^* V_2}}{\left[\overline{V_1 V_1^*}\right]^{\frac{1}{2}} \left[\overline{V_2 V_2^*}\right]^{\frac{1}{2}}} , \tag{1.B.8}$$

where $\hat{\gamma}(u,v)$ = measured complex visibility function,

$$V_1 = V\left(\frac{x_1}{\lambda}, \frac{y_1}{\lambda}, t\right) ,$$

$$V_2 = V\left(\frac{x_1}{\lambda} + u, \frac{y_1}{\lambda} + v, t\right) .$$

The overbars in Eq. (1.B.7) denote the time average, i.e.

$$\overline{V_1^* V_2} = \frac{1}{T} \int_{t'}^{t'+T} V_1^* V_2 \, dt \tag{1.B.9}$$

where T is the averaging time, and t' is the time at which the complex visibility function is measured. In the ideal case, the averaging time T should be very large in order for the complex visibility function to be measured accurately. But, T is limited due to reasons which will be discussed later.

Thus, to obtain the brightness distribution of a source, we use an interferometer which measures the complex visibility function. Using different antenna spacings $\gamma(u,v)$ is obtained for different u and v's.

C. Theory of the Two-element Interferometer

In the previous section we showed that the complex visibility function of a radio source can be measured by using two spaced antennas. We will give a brief discussion of a two-element interferometer in this section. Figure 1.c.1 shows the geometry of a simple two-element interferometer.

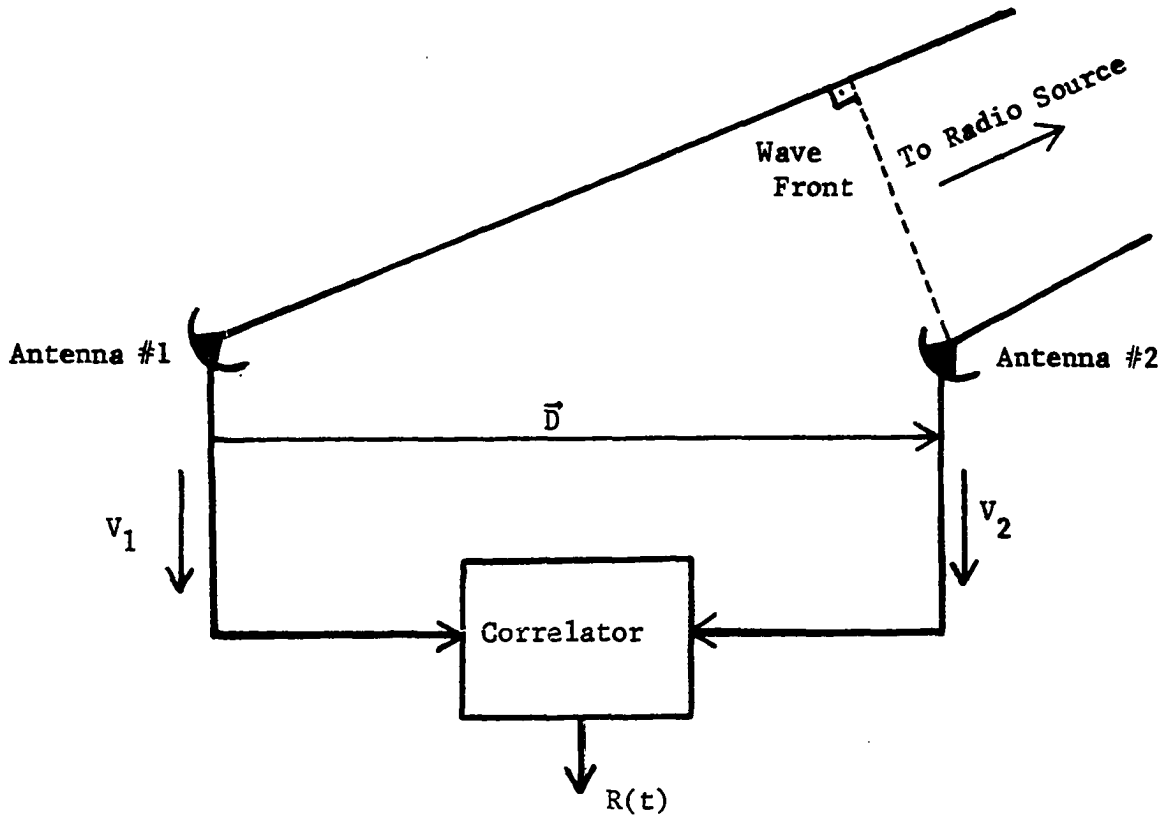


Figure 1.c.1. A simple two-element correlating interferometer

Let us find the response of the interferometer, shown in Figure 1.c.1., to a plane wave of frequency ω, radiated by a distant point source.

$$V_1 = g_1 E \exp(j\omega t), \tag{1.C.1}$$

$$V_2 = g_2 E \exp[j\omega(t+\tau_g)] \tag{1.C.2}$$

where g_1 and g_2 are the proportionality constants between electric field intensity and the terminal voltage of the first and second antennas, respectively. E is the complex electric field amplitude of the wave. The τ_g is called the <u>geometric delay</u> and given by (Rogers, 1976)

$$\tau_g = \frac{D}{c}\left[\sin\delta_B \sin\delta_s + \cos\delta_B \cos\delta_s \cos(L_s - L_B)\right]$$

$$= \frac{2\pi D}{\omega\lambda}\left[\sin\delta_B \sin\delta_s + \cos\delta_B \cos\delta_s \cos(L_s - L_B)\right] \qquad (1.C.3)$$

where c = speed of light,

D = length of the baseline vector $\vec{D}$ shown in Figure 1.c.1.

δ_B = declination of the baseline,

δ_s = declination of the source,

L_s = hour angle of the source,

L_B = hour angle of the baseline, and

λ = wavelength of the radiation.

For the definitions of δ_B, δ_s, L_s and L_B see Fomalont and Wright (1974).

As the earth rotates the hour angle of the radio source varies, thus, the geometric delay τ_g is a function of time.

At the output of the multiplier we have the complex voltage

$$R(t) = kg_1g_2\, V_1V_2^*$$

$$= kg_1g_2\, EE^* \exp(j\omega t)\exp[-j\omega(t+\tau_g)]$$

$$= kg_1g_2\, EE^* \exp(-j\omega t_g)$$

$$= G|E|^2 \exp(-j\omega\tau_g). \qquad (1.C.4)$$

where $G = kg_1g_1$

k = multiplication factor of the multiplier.

Using Eq. (1.C.3) in Eq. (1.C.4) we obtain

$$R(t) = G|E|^2 \exp\{-j\,2\pi D_\lambda[\theta_0 + \theta_1 \cos(L_s - L_B)]\}, \qquad (1.C.5)$$

where $D_\lambda = D/\lambda$, $\theta_0 = \sin\delta_B \sin\delta_s$ and $\theta_1 = \cos\delta_B \cos\delta_s$. We write Eq. (1.C.5) as

$$R(t) = G|E|^2 \cos\{2\pi D_\lambda [\theta_0 + \theta_1 \cos(L_s - L_B)]$$

$$- j\, G\, |E|^2 \sin\{2\pi D_\lambda [\theta_0 + \theta_1 \cos(L_s - L_B)]\}\} \qquad (1.C.6)$$

From electromagnetic field theory we know that $|E|^2$ is proportional to the average power radiated by the source of waves. Since the hour angle of the radio source varies with time, the interferometer output is a sinusoidal voltage with time varying phase. The amplitude of this sinusoidal output voltage is proportional to the average power radiated by the radio source. The variation of phase is determined by the source and interferometer geometry. Thus, the measurement of the phase yields information about the position of a radio source. Later we will show that the measurement of phase is quite difficult. Now, we define some terms used in interferometry.

Fringe phase: The fringe phase is defined as

$$\phi \triangleq \omega \tau_g , \qquad (1.C.7)$$

where the τ_g is as defined previously.

Fringe rate: We define the fringe rate as

$$F_R \triangleq \frac{1}{2\pi} \frac{\partial \phi}{\partial t}. \qquad (1.C.8)$$

From Eq. (1.C.7) we have

$$\frac{\partial \phi}{\partial t} = \omega \frac{\partial \tau_g}{\partial t} \qquad (1.C.9)$$

Substituting this into Eq. (1.C.8) yields

$$F_R = \frac{\omega}{2\pi} \frac{\partial \tau_g}{\partial t} .$$

Resolution of an interferometer: The angular resolution of an interferometer is defined by

$$\text{Resolution} \triangleq \frac{1}{D_p} = \frac{\lambda}{D} \tag{1.C.10}$$

where D_p is the length of the projected baseline in units of the wavelength, defind as the projection of the baseline onto the plane of the incident wavefront. The projected baseline can be represented by two components u,v, in the direction of increasing right ascension and declination, respectively. u and v are given by (Rogers, 1976).

$$u = \frac{2\pi D}{\lambda} \cos\delta_B \sin(L_s - L_B), \tag{1.C.11}$$

$$v = \frac{2\pi D}{\lambda} [\sin\delta_B \cos\delta_s - \cos\delta_B \sin\delta_s \cos(L_s - L_B)]. \tag{1.C.12}$$

Eqs. (1.C.11) and (1.C.12) are the parametric equations of an ellipse in the u - v plane. As the earth rotates, the projected baseline traces out this ellipse.

The response of an interferometer to an extended radio source can be derived by assuming that the source is a collection of many point sources radiating incoherently (Fomalont, 1973; Kraus, 1966, p. 183). The response is given by

$$R(u,v) = G P_c(u,v) \exp[j\omega_0 \tau_g(t)] , \tag{1.C.13}$$

where ω_0 is the center frequency of the band of interest, G is a factor representing the effects of antennas and receiver characteristics. It is a complex function of u, v, and ω (Moran, 1973). For clarity we assume

that it is constant and equal to unity. $P_c(u,v)$ is the correlated power of the source and called complex visibility function. In the derivation of Eq. (1.C.13) it is assumed that the radiation is not monochromatic as we assumed in the derivation of the response of a simple two-element interferometer to a point source. Moreover, a delay is inserted into one of the arms of the interferometer to compensate the geometric delay τ_g. If this procedure is omitted, one obtains (Rogers, 1976)

$$R(u,v) = GP_c(u,v)\left(\frac{\sin\frac{\Delta\omega}{2}\tau_g}{\frac{\Delta\omega}{2}}\right)\exp[j\omega_0\tau_g] \quad , \qquad (1.C.14)$$

where $\Delta\omega$ is the bandwidth of the system.

In practice the interferometers are not as simple as the one shown in Figure 1.c.1. In Figure 1.c.2 a working correlation interferometer is shown. The preamplifiers shown in the figure are low-noise maser or parametric amplifiers. The mixer stages are used to convert the frequency bands to a lower frequency band where the signal processing may be done more efficiently. A digital computer is used to track the delay. It computes the best estimate of the geometric delay by using the best estimates of the source and the system parameters. The I.F. stages are single or double sideband amplifiers depending upon the type of interferometer.

The response of an interferometer shown in Figure 1.c.2 is given (Rönnäng, 1971, p. 24)

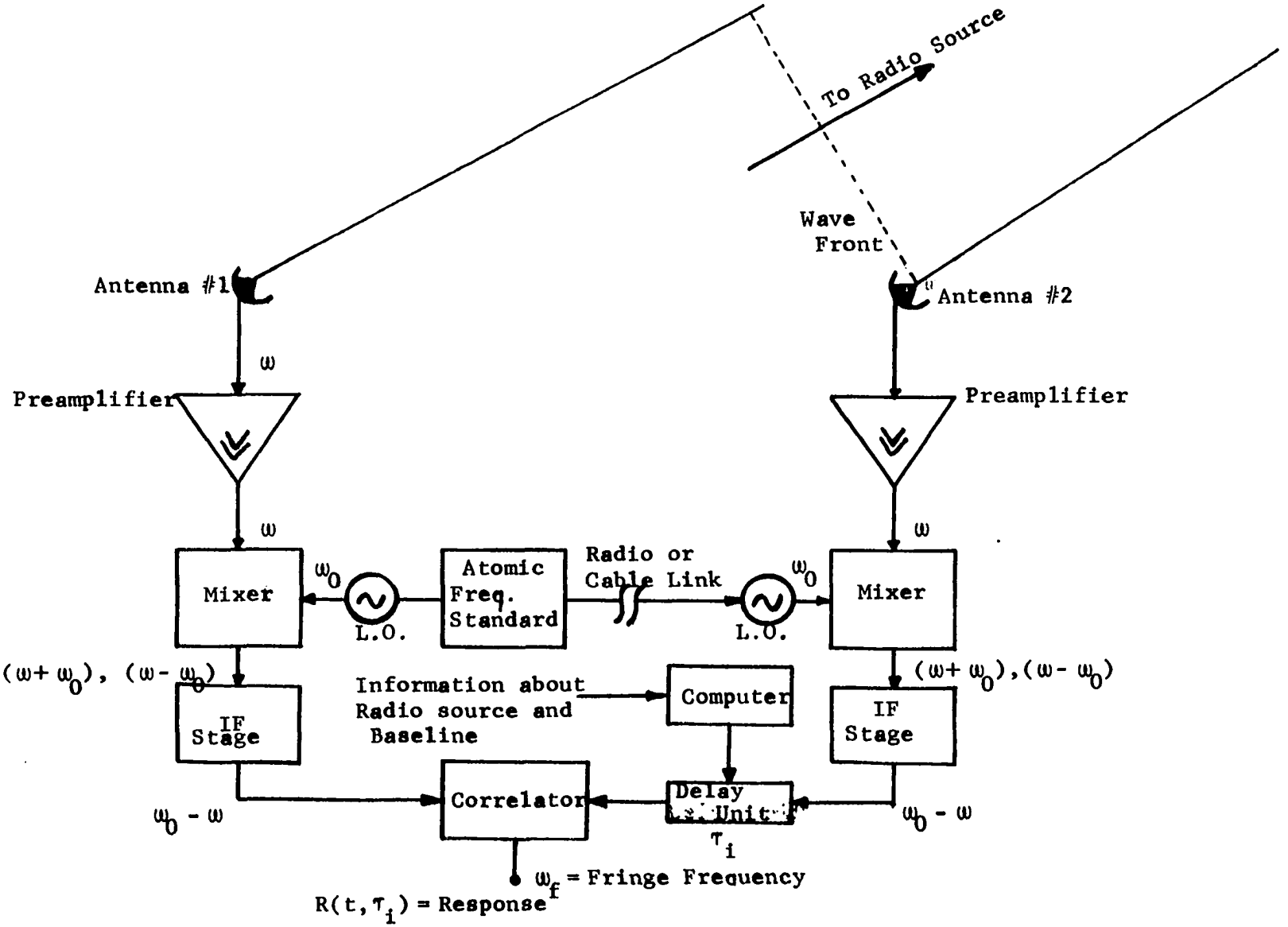


Figure 1.c.2. A block diagram of the basic elements in an interferometer system

$$R(t,\tau_i) = P_c(u,v)\exp\{j[\omega_0\tau_g(t) + \varphi_2(t) - \varphi_1(t) + \frac{\Delta\omega}{2}(\tau_g - \tau_i)]\}$$

$$\times \frac{\sin\frac{\Delta\omega}{2}(\tau_g - \tau_i)}{\frac{\Delta\omega}{2}(\tau_g - \tau_i)} , \qquad (1.C.15)$$

where $\varphi_2(t)$ and $\varphi_1(t)$ are slowly varying phases of the local oscillators 1 and 2, respectively. The two oscillators are phase locked to an atomic frequency standard so that the differential phase $\varphi_2(t) - \varphi_1(t)$ is kept small. In Eq. (1.C.15) τ_i is the time delay introduced by computer controlled delay unit. For a point source $\tau_i \simeq \tau_g$, but for an extended source τ_i is calculated for a reference position in sky, e.g. the centroid of the source being observed.

The resolution of an interferometer, given by Eq. (1.C.10), is related to the power of the interferometer to distinguish two close points in the sky. Thus, if the resolution is not good, e.g. D_p^{-1} is not small, then the fine structure of the sky will not be seen by the interferometer.

It is well-known that there are many radio sources having very small angular diameters, e.g. less than 10^{-3} arc-sec (Cohen, 1973). Therefore, in order for these compact radio sources to be observed, the baseline of the interferometer must be very large. Over several hundred kilometer, it is not possilbe to phase lock the local oscillators of the interferometer to the same atomic frequency standard. However, several thousand kilometers of baseline are required to obtain the necessary resolution. In 1967, two individual groups, one in Canada and the other in the USA, introduced Very-Long-Baseline-Interferometry (Burke, 1969). This technique

will be discussed in the next section. Hereafter, we will use the short notation VLBI for the Very-Long-Baseline-Interferometry.

D. VLBI (Very-Long-Baseline-Interferometry)

In the previous section, basic definitions in interferometry have been given. This section deals with the VLBI (Very-Long-Baseline-Interferometer). In many respects the theory of VLBI and of conventional interferometers are the same. In the VLBI technique, the two local oscillators are independent and there is no physical connection between the interferometer elements. Each local oscillator is phase locked to its own atomic standard. At each end of the interferometer, the signals are converted down to the video frequency and recorded on magnetic tapes along with the precise timing information. The recorded isgnals are brought together for further processing (Cohen, 1973). In conventional interferometers we do not attempt to track the geometric delay. This is done during processing of the signals recorded on magnetic tape.

In VLBI it is not unusual to obtain very high fringe rates, e.g. 15 kHz (Clark, 1973). To reduce the fringe rate the local-oscillator frequencies are offset by an angular frequency $\Delta\omega'$. Figure 1.d.1 shows a VLBI system schematically.

The details of a VLBI system are given in many papers (Clark, 1973; Rogers, 1976; Whitney, 1974; Fomalont and Wright, 1974). In the processing of the data, one of the signals is shifted an appropriate amount to align the signals; then, the two signals are multiplied and added together. This process is repeated until the maximum of the correlation function is obtained which corresponds to delay tracking in conventional interferometry.

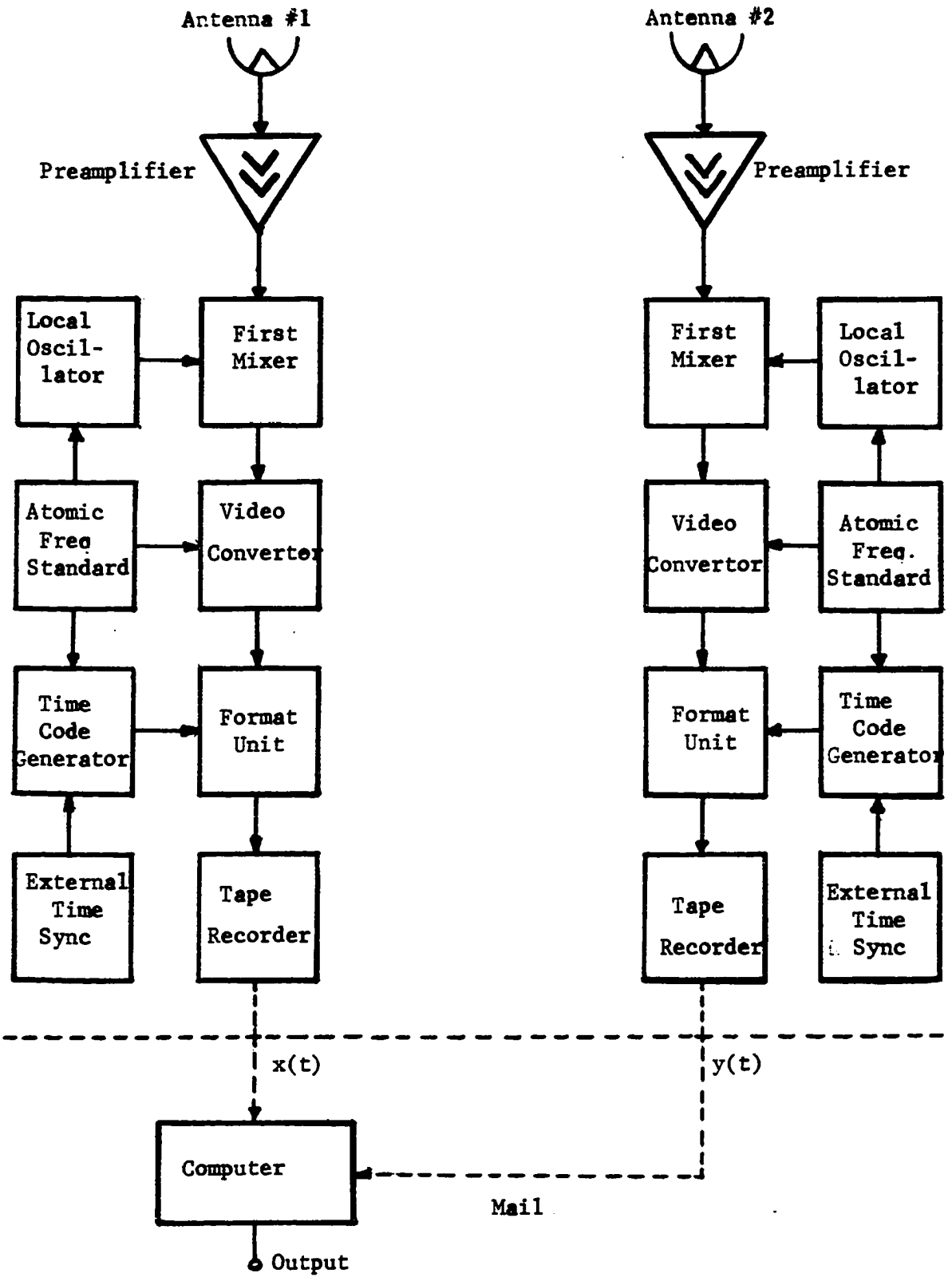


Figure 1.d.1. General block diagram of a VLBI system

The correlation function obtained after T_i seconds integration is given by (Moran, 1976)

$$R(t,\tau) = P_c(u,v)\exp\{j[\omega_0\tau_g(t) - \Delta\omega' t + \varphi_2(t) - \varphi_1(t) + \frac{\Delta\omega}{2}(\tau_g - \tau)]\} \frac{\sin \frac{\Delta\omega}{2}(\tau_g(t) - \tau)}{\frac{\Delta\omega}{2}(\tau_g(t) - \tau)} \qquad (1.D.1)$$

where ω_0 = center frequency of the band of interest,

$\Delta\omega$ = pre-detection bandwidth (or I.F. bandwidth),

P_c = correlated power,

τ = time delay introduced to one of the signals,

$\Delta\omega'$ = frequency offset, the difference between the two local oscillator frequencies, and

$\varphi_2(t) - \varphi_1(t)$ = differential clock phase noise term due to imperfections in the local oscillators.

In the derivation of Eq. (1.D.1) the effects of the receiving systems are not taken into account. This does not create any problems because in the calculation of the complex fringe visibility the quantities describing the frequency characteristics of the receiving systems do not appear. We assume that the IF filters have rectangular frequency characteristics.

To modify Eq. (1.D.1) for the ionospheric scintillation effects, we simply add another two terms, φ_1 and φ_2, to the argument of cosine function in the equation. Upon doing so we get

$$R(t,\tau) = P_c(u,v)\exp\{j[\omega_0\tau_g(t) - \Delta\omega' t + \varphi_2(t) - \varphi_1(t) + \phi_2(t) - \phi_1(t)$$

$$+ \frac{\Delta\omega}{2}(\tau_g(t) - \tau)]\} \frac{\sin \frac{\Delta\omega}{2}(\tau_g - \tau)}{\frac{\Delta\omega}{2}(\tau_g - \tau)} \qquad (1.D.2)$$

To obtain the real part of $R(t,\tau)$, the recorded signals $x(t)$ and $y(t)$ are multiplied and integrated over some time interval T_i. The outputs obtained every T_i seconds form a sinusoidal waveform. Shifting the phase of the resultant sinusoidal output yields the imaginary part of $R(t,\tau)$.

Figure 1.d.2 shows the block diagram of the process to obtain the real part of $R(t,\tau)$.

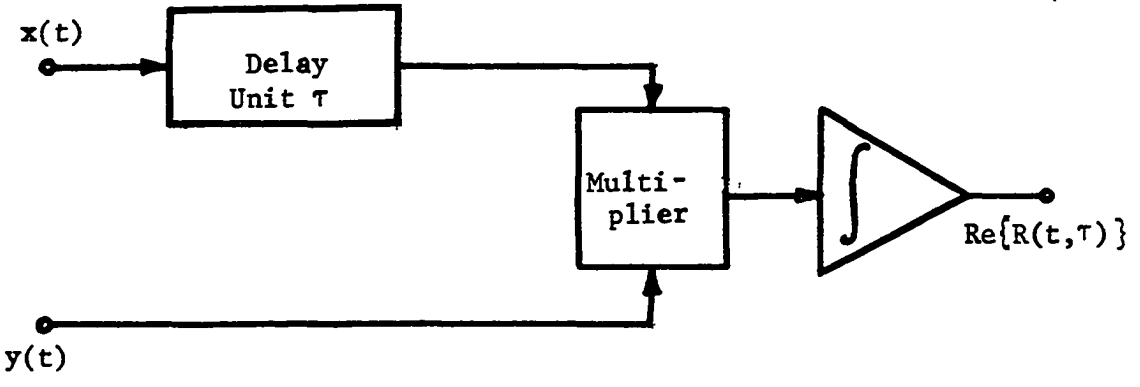


Figure 1.d.2. The block diagram of the correlation process

As seen from Eq. (1.D.2), the correlation function $R(t,\tau)$ is a function of time because of the time dependence of $\tau_g(t)$. The integration time of each interval of sampled data must be short compared to the time scale of the variations of the argument of the cosine function in Eq. (1.D.2). The uncertainty in $R(t,\tau)$ becomes large if the integration time is this small. Because of this, we like to have a large integration time, which is the reason for introducing the frequency offset between two local oscillators.

It is also possible to stop all systematic variations in $R(t,\tau)$ due to the source position and the frequency offset by using the process shown in Figure 1.d.3. Let us assume that $\tau = \tau_g$ is obtained before the fringe stopping process, and let $\Delta\varphi(t)$ denote the term $\varphi_2(t) - \varphi_1(t) + \phi_2(t) - \phi_1(t)$ in Eq. (1.D.2).

At the outputs of the integrators shown in Figure 1.d.3 we have

$$R_1(t) = P_c \cos \Delta\varphi(t), \tag{1.D.3}$$

$$R_2(t) = P_c \sin \Delta\varphi(t). \tag{1.D.4}$$

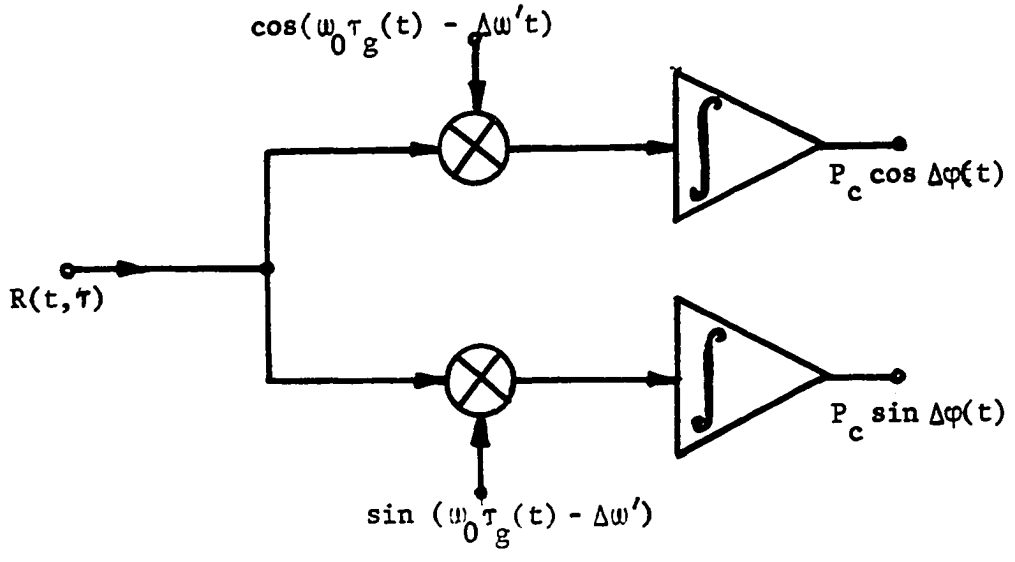


Figure 1.d.3. The block diagram of the process to stop the fringes

Now we define the complex correlation function as

$$R(t) = R_1(t) + jR_2(t) \tag{1.D.5}$$

where $j = \sqrt{-1}$, $R_1(t)$ and $R_2(t)$ have been defined by Eqs. (1.D.3) and (1.D.4) respectively. The amplitude of R(t) is

$$|A(t)| = [R_1^2(t) + R_2^2(t)]^{\frac{1}{2}}, \tag{1.D.6}$$

and called the fringe amplitude. The phase angle

$$\Delta\phi(t) = \tan^{-1}[R_2(t)/R_1(t)] \tag{1.D.7}$$

is called the residual fringe phase.

The measured complex correlation function given by Eq. (1.D.5) is in fact the correlation function given by Eq. (1.B.8) in section 1.B if we ignore the effects of the receiving systems. Thus,

$$R(t') = \frac{1}{T}\int_{t'}^{t'+T} V_1^*(t)V_2(t)dt \tag{1.D.8}$$

where $V_1(t)$ and $V_2(t)$ are complex voltages at the outputs of the antenna 1 and antenna 2, respectively. The asterisk denotes complex conjugation and t' is the time at which the correlation function is measured. To obtain the complex visibility, according to Eq. (1.B.7), we must measure the term $[\overline{(V_1V_1^*)}\ \overline{(V_2V_2^*)}]^{\frac{1}{2}}$ which is equal to $\sqrt{P_1P_2}$, where P_1 and P_2 are the measured powers of the source at the two antenna sites.

If the one-bit correlation technique is used, the correlator shown in Figure 1.d.2 gives an estimate of the normalized cross-correlation function of the signals x(t) and y(t). Where x(t) and y(t) are the outputs of the first and second IF filters, respectively. In the one-bit correlation technique, the signals x(t) and y(t) are sampled at the Nyquist rate and each sample is represented by either +1 or 0 according to

$$x_s(t) = \begin{cases} +1 & \text{if } x(t) > 0 \\ 0 & \text{if } x(t) \leq 0 \end{cases}, \tag{1.D.9}$$

$$y_s(t) = \begin{cases} +1 & \text{if } y(t) > 0 \\ 0 & \text{if } y(t) \leq 0 \end{cases}, \tag{1.D.10}$$

where $x_s(t)$ and $y_s(t)$ are sampled signals associated with x(t) and y(t), respectively. The correlation function of $x_s(t)$ and $y_s(t)$ is then given by

$$\delta_c(t_i,\tau) = \frac{1}{T_i} \sum_{t=t_i}^{t_i+T_i} [x_s(t) \oplus y_s(t+\tau)], \tag{1.D.11}$$

where t_i is the time at which the correlation is calculated, T_i the integration time, $\tau$the time delay. The symbol $\oplus$ in Eq. (1.D.11) denotes the exclusive=or operation. The multiplication and the exclusive-or are equivalent in the one-bit correlation technique. The actual normalized cross-correlation function is obtained from $\delta_c(t_i,\tau)$ by (Cohen, 1973)

$$\delta_{xy}(t_i,\tau) = \sin[\tfrac{\pi}{2}\,\delta_c(t_i,\tau)], \tag{1.D.12}$$

provided that x(t) and y(t) are normally distributed. For $\delta_c(t_i,\tau) \ll 1$, we have

$$\delta_{xy}(t_i,\tau) \simeq \tfrac{\pi}{2}\,\delta_c(t_i,\tau). \tag{1.D.13}$$

In the fringe stopping process, the $\cos[\omega_0\tau_g(t) - \Delta\omega' t]$ and $\sin[\omega_0\tau_g(t) - \Delta\omega' t]$ functions are represented by square wave functions with appropriate delays rather than phase shifts, and the normalized correlation function is found by the method described above. The correlated power is then given by (Moran, 1976)

$$P_c = [(P_1+N_1)(P_2+N_2)]^{\frac{1}{2}}\delta_{xy}(t_i,\tau) \tag{1.D.14}$$

where P_1, P_2, N_1, and N_2, are the average signal and receiver noise powers, respectively, at stations 1 and 2.

To measure the position of a point source, the phase of the complex fringe function is compared with the fringe phase of a point source whose position is well known (Cohen, 1973). The difference in phase is interpreted as a position error in the unknown source. The details of this technique are given by (Shapiro, 1976). The measurement of the geometric time delay τ_g for a point source also provides the source position information, provided the interferometer baseline parameters are known.

Using more than one frequency channel to widen the effective bandwidth has proven to be a good technique for measuring the geometric delay (Rogers, 1970).

The techniques described above give reasonably good results if the operating frequency is high, e.g. over several hundred MHz where the effects of the ionosphere are negligible.

The main source of the phase noise at small wavelengths is the atmosphere. The water vapor content in the lower atmosphere varies irregularly, thus causing random fluctuations in the refractive index of the medium. A typical RMS phase noise due to the atmosphere at 2695 MHz is about 30^{o} (Basart et al., 1970).

At decametric wavelengths, the ionosphere is the dominant source of phase noise. In the next section, the phase effects of the ionosphere will be discussed somewhat in detail. At decametric wavelengths the atmospheric effects are negligible compared with those of the ionosphere; therefore, we will not treat the atmospheric effects.

E. Effects of the Ionosphere

Because the earth's ionosphere plays a very significant role in the measurement of the positions and sizes of the radio sources by the VLBI technique, we will discuss some properties of this ionized medium. A more detailed treatment is given by Lawrence et al. (1963), Al'pert (1973), and Davies (1969).

The precision with which the positions and sizes of radio wources can be determined at low frequencies is ultimately limited not by the distance

between the two ends of a very-long-baseline interferometer but by the ionospheric effects. Therefore, the VLBI data must be corrected to account for the ionospheric effects. Several correction procedures exist that have been used by some workers. These techniques will be given in Chapter II.

Before we start investigating the interferometer phase variations due to the ionosphere, it is appropriate to give a brief background on the effects of the ionosphere on a monochromatic plane wave. A good treatment of the problem is given by Lawrence et al. (1963).

Because of the irregular nature of the ionosphere, radio waves passing through this medium are randomly disturbed.

A radio wave coming from a source in space has to travel through the ionosphere to reach an antenna on earth. Therefore, the ionosphere may affect the amplitude, phase, direction of propagation, and polarization of the wave. We will not make any attempt to discuss all effects the ionosphere imposes on a radio wave. Rather we will discuss only the effects on the phase of a radio wave.

Let us consider a monochromatic plane wave at the upper boundary of the ionosphere having a phase of zero radians. At the lower boundary the phase of the wave is given by (Davies, 1969, p. 229).

$$\phi = \frac{2\pi f}{c} \int_{\text{ray path}} \mu \, ds \qquad (1.E.1)$$

where f = frequency of the wave (Hertz),

μ = real part of the complex refractive index of the medium.

Complex refractive index is $n = \mu - j\chi$,

ϕ = phase of the wave at the lower boundary of the ionosphere.

c = velocity of light in a vacuum.

In the absence of collisions and of the magnetic field, μ of the ionized medium is given by (Davies, 1969, p. 92)

$$\mu = (1 - k \frac{N}{f^2})^{\frac{1}{2}} \tag{1.E.2}$$

where k = 80.5,

N = electron density ($\#/m^3$),

f = frequency of the wave (Hz).

We can expand Eq. (1.E.2) as

$$\mu = 1 - \frac{1}{2} (k \frac{N}{f^2}) + \frac{1}{8} (k \frac{N}{f^2})^2 - \ldots . \tag{1.E.3}$$

We keep the first two terms in Eq. (1.E.3) by assuming that

$$\frac{kN}{f^2} < 1 \tag{1.E.4}$$

Thus, we have

$$\mu \simeq 1 - \frac{kN}{2f^2} \tag{1.E.5}$$

Substituting this into Eq. (1.E.1) yields

$$\phi = \frac{2\pi f}{c} \int_{\text{ray path}} (1 - \frac{kN}{2f^2}) ds$$

$$\simeq \frac{2\pi f}{c} \int_{\text{ray path}} ds - \frac{2\pi f}{c} \int_{\text{ray path}} \frac{kN}{2f^2} ds$$

$$= \frac{2\pi f}{c} L - \frac{k\pi}{cf} \int_{\text{ray path}} N ds \tag{1.E.6}$$

where

$$L = \int_{\text{ray path}} ds \tag{1.E.7}$$

The electron density N is subject to changes with time. For a short period of time (e.g., for several minutes) we can model N as

$$N(t,s) = N_0(s) - \Delta N(t,s) \tag{1.E.8}$$

where s is the spatial variable along the ray path, t the time. $N_0(s)$ is the time average value of N(t,s) and $\Delta N(t,s)$ is the time varying irregular component of the electron density N(t,s).

Substituting Eq. (1.E.8) into Eq. (1.E.6) gives

$$\phi = \frac{2\pi f}{c} L - \frac{k\pi}{cf} \int_{\text{ray path}} N_0(s)ds + \frac{k\pi}{cf} \int_{\text{ray path}} \Delta N(t,s)ds$$

$$= \phi_0 + \Delta\phi(t) \tag{1.E.9}$$

where $$\phi_0 = \frac{2\pi f}{c} L - \frac{k\pi}{cf} \int_{\text{ray path}} N_0(s)ds, \tag{1.E.10}$$

$$\Delta\phi(t) = \frac{k\pi}{cf} \int_{\text{ray path}} \Delta N(t,s)ds. \tag{1.E.11}$$

We will call $\Delta\phi(t)$ the random phase component of a wave.

Let

$$\Delta\phi(t) = \frac{K}{f} \int_{\text{ray path}} \Delta N(t,s)\, ds \tag{1.E.12}$$

where $$K = \frac{k\pi}{c}$$

$$= 8.488 \times 10^{-7}.$$

We define $\int_{\text{ray path}} \Delta N(t,s)ds$ as the random total electron content denoted by $\Delta N_T(t)$. Thus, we write

$$\Delta\phi(t) = \frac{K}{f} \Delta N_T(t) \qquad (1.E.13)$$

The random total electron content and ϕ_0 are function of the position across the wavefront. Therefore, we no longer have a flat wavefront after a plane wave passes through the irregular ionosphere. Next, the relation between the interferometer phase and the electron density variations in the ionosphere will be derived.

The two elements of the interferometer observe two different parts of the ionosphere as shown in Figure 1.e.1.

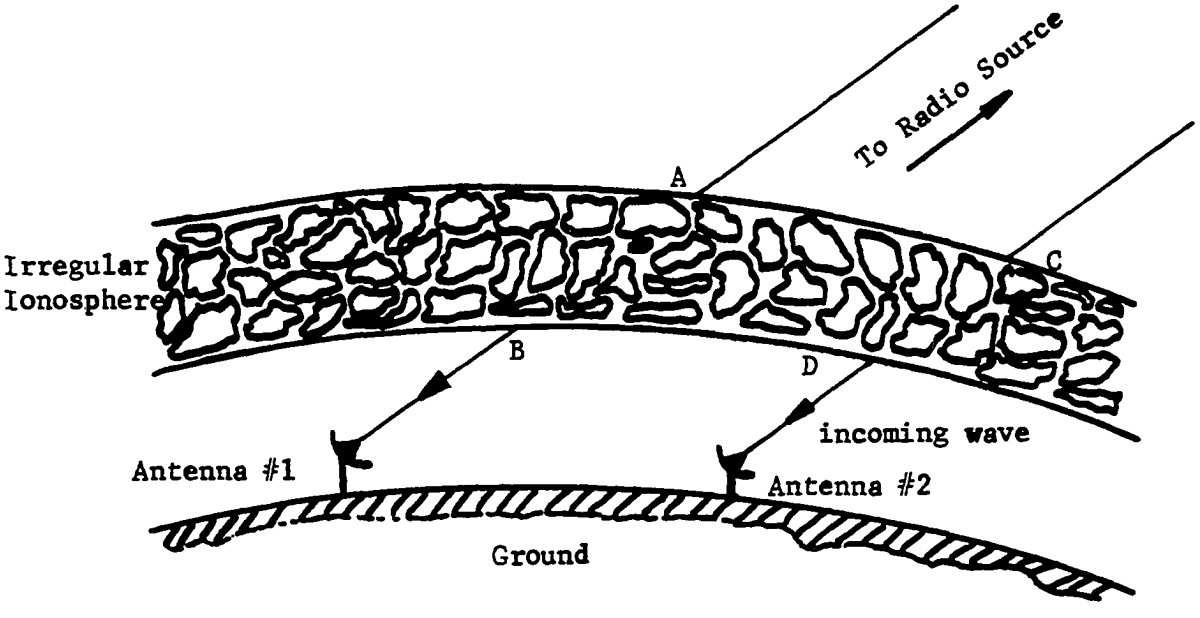


Figure 1.e.1. The position of the ionosphere in radio interferometry

It can be shown that one component of the interferometer phase is the difference of the phases introduced by the ionosphere at two sites. This term will be defined as the differential phase. In the previous section

it was shown that at any instant of time the phase change at each site was given by

$$\Delta\phi_i = \frac{K}{f} \int_{\substack{i^{th} \\ \text{ray path}}} \Delta N_i \, ds \quad , \; i = 1,2 \tag{1.E.14}$$

where i is the site number. For the first site we have

$$\Delta\phi_1 = \frac{K}{f} \int_{\text{ray path 1}} \Delta N_1(t,s) ds \; , \tag{1.E.15}$$

and for the second

$$\Delta\phi_2 = \frac{K}{f} \int_{\text{ray path 2}} \Delta N_2 ds \; . \tag{1.E.16}$$

The differential phase is

$$\Delta\phi = \Delta\phi_1 - \Delta\phi_2 = \frac{K}{f}\left(\int_A^B \Delta N_1 ds - \int_C^D \Delta N_2 ds \right)$$

$$= \frac{K}{f}(\Delta N_{1T} - \Delta N_{2T}) \tag{1.E.17}$$

where ΔN_{1T}, ΔN_{2T} are random total electron contents along the ray paths AB and CD, respectively.

We will develop statistical relations between differential phase, random total electron content, and the spatial distribution of electron density irregularities. Both ΔN_{1T} and ΔN_{2T} are functions of time. If we assume that the ionosphere has a structure consisting of blobs (different density volumes) and further that the electron density distribution within each individual blob is not changing with time, we can assume that the phase changes are primarily due to the rotation of the earth. With respect

to a constant coordinate system, a radio source is considered to be fixed while the earth is rotating. Therefore, a line joining the interferometer antenna and the radio source is not fixed in the same coordinate system. As the earth rotates, the line of sight passes through different blobs.

Under the assumptions we stated above, the spatial irregularity distribution is related to the time dependent differential interferometer phase. This relation allows us to infer some properties of the ionosphere from differential phase data.

F. Probability Density Distribution of Differential Phase

In the last section it was shown that the differential phase is expressed as

$$\Delta\phi = \frac{K}{f}(\Delta N_{1T} - \Delta N_{2T}) \qquad (1.F.1)$$

where $\Delta\phi$ = differential phase (in radians)

$K = 8.488 \times 10^{-7}$

f = the frequency of the wave (Hz)

ΔN_{1T} = random total electron content along the first ray path (m^{-2})

ΔN_{2T} = random total electron content along the second ray path (m^{-2})

To get the probability density distribution function of $\Delta\phi$, we write $\Delta\phi$ as

$$\Delta\phi = \Delta\phi_1 - \Delta\phi_2 \ , \qquad (1.F.2)$$

then $p(\Delta\phi)$ is given by (Papoulis, 1965, p. 189)

$$p(\Delta\phi) = p(\Delta\phi_1) \circledast p(\Delta\phi_2) \qquad (1.F.3)$$

where $p(\Delta\phi)$ is probability density distribution function of $\Delta\phi$ and the symbol $\circledast$ denotes the convolution.

For evaluation of the above expression we need to have explicit probability distribution functions for the random variables $\Delta\phi_1$ and $\Delta\phi_2$. Now, let us assume that the irregular components of the total electron contents along two paths joining the source and the two interferometer sites have the same form of probability distribution function. In addition, assume that the probability distribution function has a Gaussian form, e.g.,

$$(\Delta\phi_i) = \frac{1}{\sqrt{2\pi\sigma^2_{\Delta\phi_i}}} \exp\left(-\frac{(\Delta\phi_i)^2}{2\sigma^2_{\Delta\phi_i}}\right), \qquad i = 1,2 \qquad (1.F.4)$$

where $\sigma^2_{\Delta\phi_i}$ is the variance of the random variable $\Delta\phi_i$, and we assumed that $E[\Delta\phi] = 0$. Substituting Eq. (1.F.4) into Eq. (1.F.3) we find

$$p(\Delta\phi) = \left[\frac{1}{\sqrt{2\pi\sigma^2_{\Delta\phi_1}}} \exp\left(-\frac{(\Delta\phi_1)^2}{2\sigma^2_{\Delta\phi_1}}\right)\right] \circledast \left[\frac{1}{\sqrt{2\pi\sigma^2_{\Delta\phi_2}}} \exp\left(-\frac{(\Delta\phi_2)^2}{2\sigma^2_{\Delta\phi_2}}\right)\right]$$

$$= \frac{1}{\sqrt{2\pi\sigma^2_{\Delta\phi_1}}\sqrt{2\pi\sigma^2_{\Delta\phi_2}}} \left[e^{-\frac{(\Delta\phi_1)^2}{2\sigma^2_{\Delta\phi_1}}} \circledast e^{-\frac{(\Delta\phi_2)^2}{2\sigma^2_{\Delta\phi_2}}}\right]$$

$$= \frac{1}{\sqrt{(2\pi)^2\sigma^2_{\Delta\phi_1}\sigma^2_{\Delta\phi_2}}} \int_{-\infty}^{\infty} e^{-\frac{(\Delta\phi-\tau)^2}{2\sigma^2_{\Delta\phi_1}}} e^{-\frac{\tau}{2\sigma^2_{\Delta\phi_2}}} d\tau$$

$$= \frac{1}{\sqrt{(2\pi)^2 \sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}}} \int_{-\infty}^{\infty} \exp - \left[\frac{1}{2} \frac{\sigma^2_{\Delta\phi_1} + \sigma^2_{\Delta\phi_2}}{\sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}} \tau^2 - \frac{\Delta\phi}{\sigma^2_{\Delta\phi_1}} \tau + \frac{1}{2\sigma^2_{\Delta\phi_1}} (\Delta\phi)^2\right] d\tau$$

$$= A \int_{-\infty}^{\infty} \exp - (a\tau^2 + b\tau + c) d\tau \qquad (1.F.5)$$

where $a = \frac{1}{2} \frac{\sigma^2_{\Delta\phi_1} + \sigma^2_{\Delta\phi_2}}{\sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}}$, $b = - \frac{\Delta\phi}{\sigma^2_{\Delta\phi_1}}$, $c = \frac{(\Delta\phi)^2}{2\sigma^2_{\Delta\phi_1}}$,

$$A = \left[(2\pi)\ \sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}\right]^{-1}$$

Using integral tables we find

$$\int_{-\infty}^{\infty} \exp - (a\tau^2 + b\tau + c) d\tau = \sqrt{\frac{\pi}{a}} \exp\left[(b^2 - 4ac)/4a\right] \qquad (1.F.6)$$

Substituting this in Eq. (1.F.5) we get

$$p(\Delta\phi) = A\sqrt{\frac{\pi}{a}} \exp\left[(b^2 - 4ac)/4a\right]$$

$$= \frac{1}{\left[(2\pi)\ \sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}\right]} \left[\frac{\pi}{\frac{1}{2} \frac{\sigma^2_{\Delta\phi_1} + \sigma^2_{\Delta\phi_2}}{\sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}}}\right]^{\frac{1}{2}} \times$$

$$\exp \left\{\left[\frac{(\Delta\phi)^2}{4\sigma^2_{\Delta\phi_1}} - \frac{(\Delta\phi)^2}{2\sigma^2_{\Delta\phi_1}} \times \frac{\sigma^2_{\Delta\phi_1} + \sigma^2_{\Delta\phi_2}}{\sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}}\right]/2 \frac{\sigma^2_{\Delta\phi_1} + \sigma^2_{\Delta\phi_2}}{\sigma^2_{\Delta\phi_1} \sigma^2_{\Delta\phi_2}}\right\} \qquad (1.F.7)$$

Simplifying Eq. (1.F.7) yields

$$p(\Delta\phi) = \frac{1}{\sqrt{2\pi(\sigma^2_{\Delta\phi_1} + \sigma^2_{\Delta\phi_2})}} \exp\left[-\frac{(\Delta\phi)^2}{2(\sigma^2_{\sigma_{\Delta\phi_1}} + \sigma^2_{\Delta\phi_2})}\right] \qquad (1.F.8)$$

This derivation assumes that the random electron density variations on two sites are statistically independent. From Eq. (1.F.8) we see that the interferometer fringe phase, in the absence of any other effects, is a Gaussian random variable with the variance of $(\sigma^2_{\Delta\phi_1} + \sigma^2_{\Delta\phi_2})$. Where $\sigma^2_{\Delta\phi_1}$ is the variance of $\Delta\phi_1$ and $\sigma^2_{\Delta\phi_2}$ is the variance of $\Delta\phi_2$. $\Delta\phi$ has a zero mean. A sample histogram of the fringe phase is shown in Figure 1.e.2.

If there is a correlation between phase fluctuations at two interferometer antennas, we can find the relation between the mean square fringe phase deviation and the correlation coefficient of the phase fluctuations. Let δ be the correlation coefficient between $\Delta\phi_1$ and $\Delta\phi_2$, then

$$\text{Var}[\Delta\phi] = \text{Var}[\Delta\phi_1 - \Delta\phi_2] = \text{Var } \Delta\phi_1 + \text{Var } \Delta\phi_2 - 2 \text{ cov}(\Delta\phi_1, \Delta\phi_2) \qquad (1.F.9)$$

since

$$\text{cov}(\Delta\phi_1, \Delta\phi_2) = \delta[(\text{Var } \Delta\phi_1)(\text{Var } \Delta\phi_2)]^{\frac{1}{2}} \qquad (1.F.10)$$

we have

$$\text{Var } \Delta\phi = \text{Var } \Delta\phi_1 + \text{Var } \Delta\phi_2 - 2\delta[\text{Var } \Delta\phi_1 \text{ Var } \Delta\phi_2]^{\frac{1}{2}} \qquad (1.F.11)$$

If

$$\text{Var } \Delta\phi_1 = \text{Var } \Delta\phi_2 = \sigma^2 , \qquad (1.F.12)$$

then we get

$$\begin{aligned} \text{Var } \Delta\phi &= 2\sigma^2 - 2\delta\sigma^2 \\ &= 2(1 - \delta)\sigma^2 \end{aligned} \qquad (1.F.13)$$

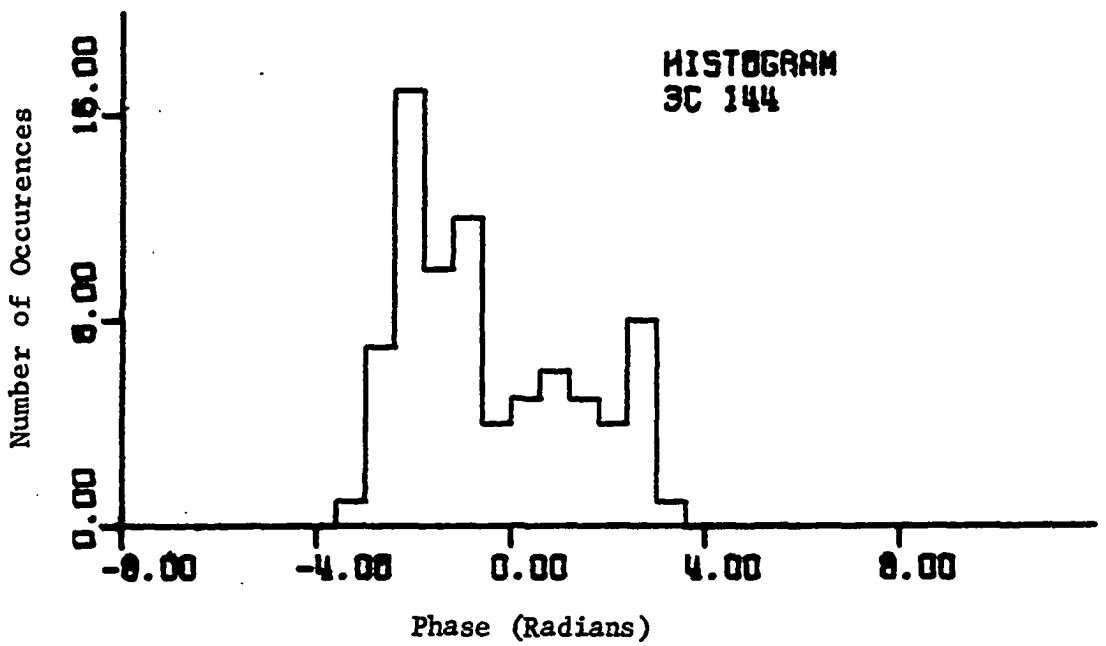


Figure 1.f.1. A sample histogram of the residual interferometer phase. Each phase record is 2 seconds in length

and we see that the presence of the correlation between $\Delta\phi_1$ and $\Delta\phi_2$ reduces the variance of the residual fringe phase. If the correlation is perfect; e.g., $\delta = 1$ we obtain Var $\Delta\phi = 0$. In other words if there is perfect correlation between the phase corruptions at two antennas, there is no residual phase variations due to the ionosphere. To see how the phase fluctuations affect the correlation of the fields at two antennas, we assume a monochromatic plane wave with unit amplitude. In complex notation we represent this wave by

$$E = \exp(j\omega t) \tag{1.F.14}$$

where ω is the angular frequency of the wave. Let this simple plane wave be present at the upper boundary of the ionosphere. Assuming that the ionosphere changes only the phase of the wave when the receivers are in the Fresnel Zone of the irregularities, we have

$$E_1 = \exp(j\omega t + \phi_1) \ , \tag{1.F.15}$$

$$E_2 = \exp(j\omega t + \phi_2) \ , \tag{1.F.16}$$

where E_1 is the electric field at the first antenna, E_2 is the electric field at the second antenna. ϕ_1 and ϕ_2 are random variables varying in time. This assumption is true if the receivers are in the Fresnel Zone of the irregularities. Now we form the covariance of E_1 and E_2.

$$\text{cov}(E_1,E_2) = \langle E_1 E_2^* \rangle \ , \tag{1.F.17}$$

where the asterisk denotes complex conjugate. We write $\text{cov}(E_1,E_2)$ as

$$\text{cov}(E_1,E_2) = \gamma_{12} = \langle e^{j\omega t + j\phi_1} e^{-j\omega t - j\phi_2} \rangle$$

$$= \langle e^{j(\phi_1 - \phi_2)} \rangle . \qquad (1.F.18)$$

Let $\Delta\phi = \phi_1 - \phi_2$, then Eq. (1.F.18) can be written as

$$\gamma_{12} = \langle \exp(j\Delta\phi) \rangle . \qquad (1.F.19)$$

Assuming that $\Delta\phi$ is a Gaussian random variable with zero mean and variance $\sigma^2_{\Delta\phi}$, we write

$$p(\Delta\phi) = \frac{1}{\sqrt{2\pi\,\sigma^2_{\Delta\phi}}} \exp\left[-\frac{(\Delta\phi^2)}{2\sigma^2_{\Delta\phi}}\right] , \qquad (1.F.20)$$

where $p(\Delta\phi)$ is the probability density distribution function of $\Delta\phi$. From the definition of the expected value of a random variable, we write Eq. (1.F.19) as

$$\gamma_{12} = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\,\sigma^2_{\Delta\phi}}} \exp\left[-\frac{(\Delta\phi)^2}{2\sigma^2_{\Delta\phi}}\right] \exp(j\Delta\phi)\, d(\Delta\phi)$$

$$= \exp[-\sigma^2_{\Delta\phi}/2]. \qquad (1.F.20)$$

From the above equations we see that if $\text{Var}(\phi_1 - \phi_2)$ is large enough, the covariance of two fields approaches to zero. That is to say, independent phase variations imposed by the ionosphere on radio waves received by the two antennas of the interferometer may destroy the correlation between the fields. Therefore, before obtaining the correlation function one must correct the phase according to the known characteristics of the physical process taking place in the F region of the ionosphere.

In the above discussion we mentioned the irregular component of the interferometer fringe phase. In addition to the irregular component there is a slowly varying differential phase component due to the changes in the mean electron densities of the ionosphere above two interferometer sites. Mean electron density depends on the time of day, e.g., mean electron density during the day is different from that of the night. These variations are quite slow compared to the phase fluctuations due to the irregularities in the ionosphere (Stannard et al., 1970).

For a short period of time, we can assume that the differential regular phase component at the output of the interferometer is constant. This regular phase term has to be removed to do further processing on the phase data.

To remove the constant phase term due to the mean electron density differences between two interferometer sites, we need to know average values of the total electron contents along two ray paths. The best way to obtain the total electron content information is the use of the ionospheric sounding equipment at each interferometer terminal while a VLBI experiment is being performed. This method is not preferred because of its complexity. Another method for removing the effects of mean-total-electron-density-difference over two VLBI sites is to use mathematical models for the ionosphere.

Although the mathematical models do not represent the mean electron density structure, they are used because of their simplicity. Having an electron density model is not enough to calculate the total electron content along radio wave paths, because the wave paths are not known very

well. Therefore a ray tracing procedure must be used to obtain exact ray paths (Mathur et al., 1970). If the observation period is long, e.g., several hours, then N_0 in Eq. (1.E.8) is no longer constant and may vary from $5 \times 10^{17} m^{-2}$ by day to 5×10^{16} by night (Hagfors, 1976). Therefore, care must be taken when using any correction method. But, because of highly disturbed conditions of the ionosphere, especially at frequencies below 100 MHz, VLBI observations are limited to several minutes at a time. Thus, we do not have to be concerned with variations in ϕ_0.

II. PREVIOUS WORK

In the previous chapter we discussed the basic radio interferometer and the effects of the ionosphere on the complex fringe visibility function. The ionosphere is the biggest problem in decametric wavelength interferometry. The scintillation effects can totally destroy the fringes thus prohibiting low-frequency interferometry.

There have been many attempts by various workers to overcome the problems caused by the ionosphere. If we knew the TECs along the wave paths at any instant of time we could correct all the effects of this highly dispersive medium, where TEC stands for "total electron content."

The proposed methods are listed below:

(a) Satellite Tracking Method: In this method, a satellite is tracked by two sites of the interferometer during an interferometric experiment. From Faraday rotation measurements total electron contents along the paths joining the two interferometer antennas and the satellite can be deduced. Figure 2.1 illustrates this method.

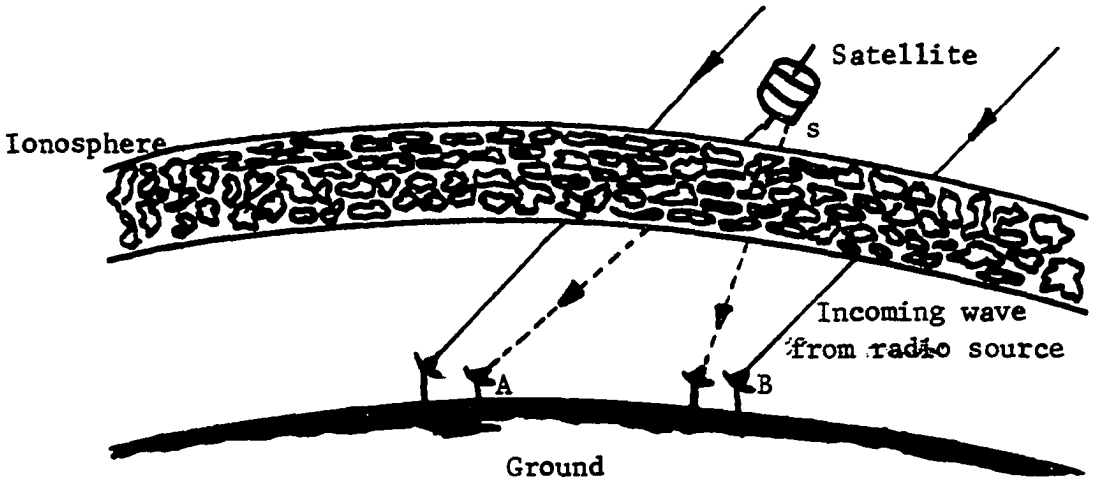


Figure 2.1 The geometry of satellite tracking method

To deduce TEC's along A - S and B - S, two signals with different frequencies are received and Faraday rotation angles are measured at the two interferometer sites. Using the relation between the integrated electron densities and the Faraday rotation angles, TECs are found. The details of Faraday rotation are given in Kraus (1966).

From Figure 2.1 we see that the wave paths A - S and B - S do not coincide with the wave paths to the source, A - ∞ and B - ∞. Therefore, computed TEC's from satellite monitoring are not exact values of TEC's along the paths to the source. Another disadvantage of this method is that it is not always possible to find a satellite at the right place and time.

If the satellite were in the same place as the radio source we would need two different frequency receivers operating simultaneously.

(b) Backscatter Method: This technique is described as follows. At each VLBI site a laser beam is sent toward the ionosphere in the source direction. From the reflected part of the beam the total electron content may be deduced. A disadvantage of this method is due to the fact that the wave path for the radiation coming from a radio source is not a simple straight line. The electron content measured by the back scatter technique does not give the TEC along the real wave path. This method is also very expensive to employ. Figure 2.2 illustrates the method.

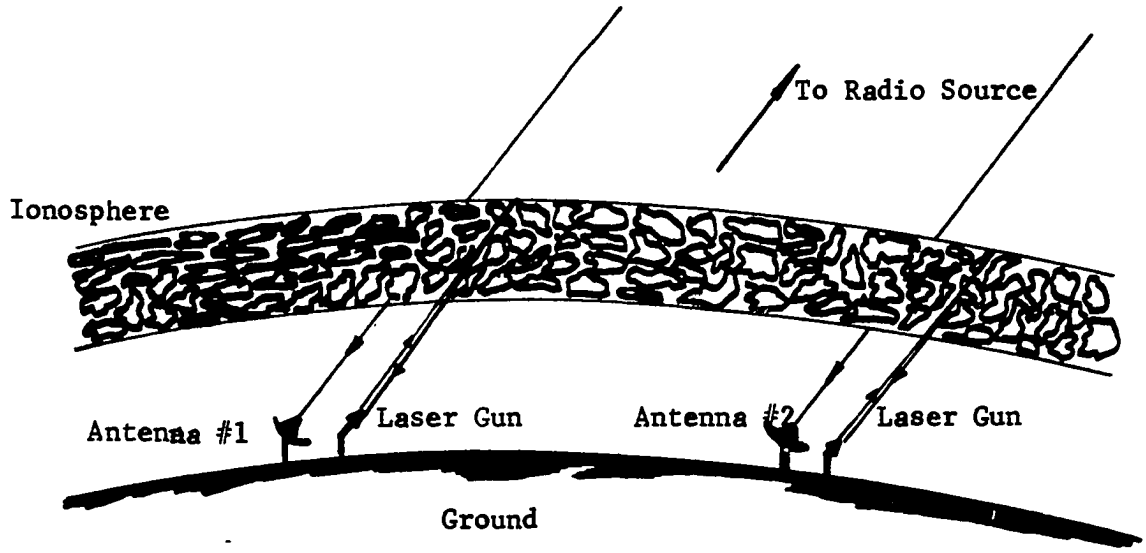


Figure 2.2. The illustration of the backscatter method

(c) Two-Frequency Method: In this technique, at each end of the interferometer two separate receivers operating at widely separated frequencies are employed. Owing to the dispersive character of the ionoshpere it is possible to deduce TEC's along the wave paths, or to remove the effects of the ionosphere by comparing the fringe phases obtained each from a different frequency channel. To show how this method works we take the Eq. (1.E.12) given in chapter I,

$$\phi(t) = \frac{K}{f} \int_{\text{wave path}} N(t,h)dh \qquad (2.1)$$

where $N(h,t)$ = electron density as a function of the distance along the ray path and time $[\#/m^3]$

K = a constant $[= 8.488 \times 10^{-7}]$

f = the frequency of the wave [Hz]

$\phi(t)$ = the phase of the wave, due to the presence of electrons in the medium.

If we denote the center frequencies of two channels by f_1 and f_2, the residual fringe phase due to the ionosphere for the frequency f_1 is

$$\Delta\phi^{(1)} = \frac{K}{f_1} \int_{\text{wave path}} \Delta N(h,t)dh \tag{2.2}$$

For the other frequency

$$\Delta\phi^{(2)} = \frac{K}{f_2} \int_{\text{wave path}} \Delta N(h,t)dh \tag{2.3}$$

From Eqs. (2.2) and (2.3) we have

$$\Delta\phi^{(2)} = \frac{f_1}{f_2} \Delta\phi^{(1)} \tag{2.4}$$

where $\Delta N(h,t) = N_1(h,t) - N_2(h,t)$, $N_1(h,t)$ and $N_2(h,t)$ are the electron densities along wave paths joining the source with two VLBI sites, respectively.

From the measurement of $\Delta\phi^{(1)}$ and using Eq. (2.4), $\Delta\phi^{(2)}$ is predicted. This value is subtracted from the measured residual fringe phase of the interferometer for the frequency f_2.

This technique seems to be promising but as with the other phase correction techniques it has its own disadvantage. A VLBI system is quite complicated and using another frequency channel adds more complexity and cost to the existing system. Whitney et al. (1976) used a similar technique for the group delay estimate. Instead of using two receivers at

each end, they switched the frequency bands at a high rate to obtain the same result. However, the disadvantage of this method is that less information is gathered from the radio source.

(d) Two-source Method: This technique is useful only for the position measurements by an interferometer. It employs an additional radio source whose position is known quite accurately. The requirement for this source is that the angular separation between the two sources must be small. The interferometer is switched rapidly back and forth from one source to the other source. Subtracting the fringe phases obtained from observations of the two sources, we obtain the difference phase function which is free of almost all propagation medium effects, the effects of any clock wandering (Shapiro, 1976). This method can be used if we have a radio source close to the one for which we want to estimate the position.

There are two other techniques, so called broken coherence, and coherent averaging techniques. Before we discuss these two methods, it is appropriate to give the signal processing methods in VLBI (Very-Long-Baseline-Interferometry).

A. Signal Processing in VLBI

For either a continuum or a spectral line radio source, the main objective in the processing of the signals recorded at each end of an interferometer is to compute the cross-correlation function of the signals. There are at least six VLBI recording and processing systems in existence today (Moran, 1976). All of them do the same thing, i.e., to compute correlation between two incoming waves.

The theoretical aspects of digital signal processing in VLBI have been discussed by several authors (Moran, 1976; Rogers, 1976; Clark, 1973). Since the data we have analyzed in this thesis have been recorded by a NRAO Mark I VLBI system, we will describe this system in some detail.

In the Mark I system, the data are recorded on 7-track computer tapes in a fashion that any general-purpose computer can read. The density of the data on computer tapes is low, which reduces the error rate to almost zero. The data at the two terminals are clipped and recorded at 720 k bits per second which corresponds to a 360 kHz video bandwidth. Moreover, the data are blocked into records each having .2 sec length.

No explicit timing information is recorded on the tapes and each bit is labeled by counting the number of the records from the beginning of the tape. Even though no explicit timing information is being written on the tapes, the starting time is known very accurately. The first bit is recorded 0.31 sec after a pre-set trigger starts the recorder.

The data can be processed by any general-purpose computer with the programs required. A disadvantage of this system is that the time required for performing the signal processing is long, e.g., a seven delay-correlation function of 1 sec of data requires 15 sec computation time if IBM 360 Model 50 is used. The NRAO Mark II system overcomes this problem by employing hardware digital processing equipment.

The sampling is done in two levels in the Mark I system because of its simplicity. The cost of this scheme is the loss of signal-to-noise ratio by a factor of $\pi/2$. The philosophy behind the idea of using 1-bit sampling is that assuming the signals are Gaussian random processes with zero means,

all information about the signals can be obtained from second moments of the processes. And for a Gaussian process, the second moment can be obtained from zero-crossing information of the process. Thus, we only take the sign of the signal at any instant of time.

Figure 2.a.1 shows a block diagram of the signal processing used in the Mark I system.

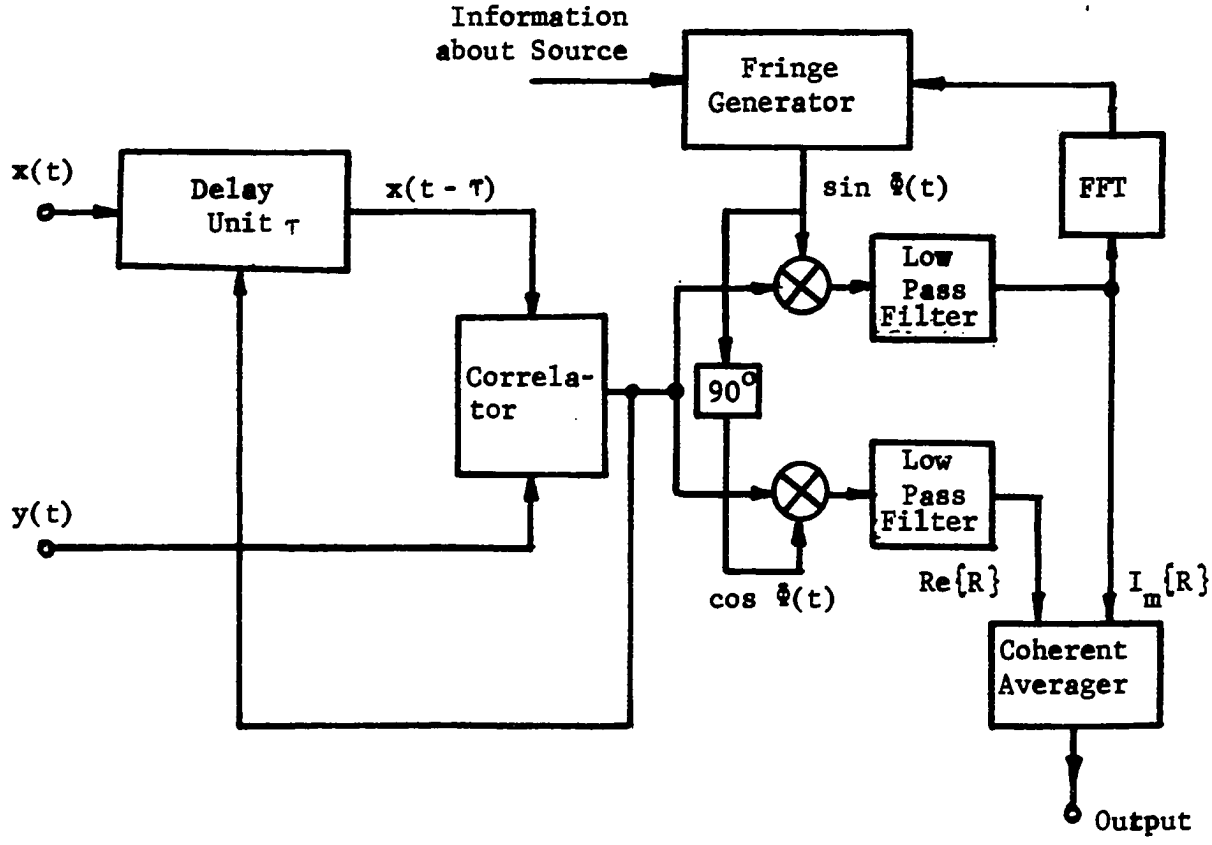


Figure 2.a.1. VLBI signal processing in block diagram form

The operations shown in Figure 2.a.1 can easily be performed by a digital computer. From the previous chapter, Eq. (1.D.2), we have

$$R(t,\tau) = P_c(u,v)\exp\{j[\omega_0\tau_g(t) + \varphi_2(t) - \varphi_1(t) + \phi_2(t) - \phi_1(t) - \Delta\omega' t + \frac{\Delta\omega}{2}(\tau_g - \tau)]\} \times \{\sin[\frac{\Delta\omega}{2}(\tau - \tau_g)]/[\frac{\Delta\omega}{2}(\tau - \tau_g)]\} \quad (2.A.1)$$

For convenience, let $\Delta\omega' = 0$. The natural fringe rate described by

$$F_F = \frac{1}{2\pi}\frac{d}{dt}(\omega_0\tau_g(t)) = \frac{\omega_0}{2\pi}\frac{d\tau_g(t)}{dt} \quad (2.A.2)$$

can take values up to 10 kHz or even more. Therefore to obtain $R(t,\tau)$ we can not use any arbitrary integration time. The integration time should be much smaller than the fringe period in order for the fringes not to be washed out. We notice that the finite bandwidth of the receivers affects the correlation function, therefore, the $\tau = \tau_g$ condition must be satisfied at all times. To do that, $R(t,\tau)$ is computed for more than one lag and the maximum of this function is taken as the value of $R(t,\tau)$. The time lag τ that gives maximum correlation is in fact the maximum likelihood estimate of the time delay between two wavefronts arriving at two VLBI terminals (Rogers, 1970).

As seen from Eq. (2.A.1), even when $\tau = \tau_g$ is satisfied, the cross correlation function varies with time because of the time variations of $\tau_g(t)$ and $\Delta\phi(t)$. We again note that $\Delta\phi(t)$ is the phase noise due to the propagation media and imperfect local oscillators. To remove the time variations due to the $\omega_0\tau_g^c(t)$ term, we multiply $R(t,\tau)$ obtained at the end of each integration by sine of $\omega_0\tau_g^c(t)$ and cosine of $\omega_0\tau_g^c(t)$, and further integrate the resultant functions over some time interval, where τ_g^c is the calculated geometric delay. This integration time is determined

by the expected fringe frequency offset $\frac{1}{2\pi}\frac{d}{dt}(\omega_0\tau_g(t) - \omega_0\tau_g^c(t))$.

Because of the $\Delta\phi(t)$ term in the correlation expression, we cannot stop the fringes even though we know the exact fringe phase. Suppose the fringe phase is known within very small uncertainty limits, then at the end of each integration we obtain

$$R_c(t) = A(t)\cos[\Delta\phi(t)] \tag{2.A.3}$$

$$R_s(t) = A(t)\sin[\Delta\phi(t)] \tag{2.A.4}$$

where $R_c(t)$ is the real part of the complex fringe visibility function, $R_s(t)$ is the imaginary part of the complex visibility function. $\Delta\phi(t)$ includes the complex fringe visibility phase plus the phase noise introduced by the local oscillators and the ionosphere. This last integration process corresponds to low-pass filtering. We obtain the complex correlation function from Eqs. (2.A.3) and (2.A.4) as

$$\begin{aligned} R(t) &= [R_c^2(t) + R_s^2(t)]^{\frac{1}{2}} \exp\{j[\arctan(R_c(t)/R_s(t))]\} \\ &= A(t)\exp[j\phi(t)]. \end{aligned} \tag{2.A.5}$$

For a 160 sec long observation we obtain 800 complex numbers for the complex fringe function.

Assuming that each telescope signal is represented by its sign at any instant of time, we can summarize a typical signal processing scheme as follows: One of the bit streams is first shifted an appropriate amount

until the bits to be correlated correspond to samples of the same wave front. The bit streams are then multiplied together and accumulated for 288 bits (0.4 msec). The numbers, obtained every 0.4 msec, are then multiplied by the sine and cosine of the calculated fringe phase and further accumulated for 0.2 sec. This process of multiplication and accumulation corresponds to matched-filter detection in communications. The process described above is repeated for each block of data. If there are no source position errors, baseline errors, propagation errors, and clock instability effects, the outputs obtained every 0.2 sec would be the same within the statistical errors.

The estimates of the complex fringe function have large sampling errors due to short integration times. To reduce the errors, the outputs obtained every 0.2 sec are time averaged coherently over some period of time. As we will discuss later, coherent averaging cannot be done over the whole observation interval.

As mentioned before, the phase of the complex fringe function is a random variable due to the fluctuations of the local oscillator phases and the propagation media, e.g., the ionosphere and the atmosphere.

Even if there were no ionosphere and oscillator phase corruptions, the fringe phase and amplitude would still be random variables because of the receiver noises, sky background noises, and statistical sampling errors. The analysis of these quantities have been made by Moran (1973). We will only give the results and remember that the analysis of Moran assumes no ionospheric propagation path effects.

Let us model the observed fringe phase as

$$\phi(t) = \phi_{source}(t) + \phi_{iono}(t) + \phi_{clock} + \phi_{stat}(t) \qquad (2.A.6)$$

where $\phi(t)$ = observed residual fringe phase

$\phi_{source}(t)$ = phase carrying information about the radio source

$\phi_{iono}(t)$ = noise introduced by the ionosphere

$\phi_{clock}(t)$ = the difference in the local oscillator phases

$\phi_{stat}(t)$ = statistical phase noise due to the receiver, skyback-ground noises and statistical sampling process.

If we are observing a point source with no position error, the first term on the right hand side of Eq. (2.A.6) equals zero. For simplicity, we will assume that a point source is being observed.

The ionospheric noise $\phi_{iono}(t)$ is our prime concern in this thesis because the other terms are negligible at decametric wavelengths. $\phi_{clock}(t)$ has two components: a linear drift, and a white noise component. As it will be shown later, the linear drift can be removed easily, and we assume that the white noise component is negligible compared to $\phi_{iono}(t)$.

Now we define the coherence time of an interferometer. The coherence time is defined as (Cohen, 1973) the maximum time interval T_c for which

$$\left[\overline{\phi^2(t)}\right]^{\frac{1}{2}}_{T_c} \leq 1 \text{ rad} \qquad (2.A.7)$$

where $\phi(t)$ is as in Eq. (2.A.6), and we assume that

$$E[\phi(t)] = 0. \qquad (2.A.8)$$

The overbar in Eq. (2.A.7) denotes time averaging.

The coherence time of an interferometer is an important parameter of the system. It shows how good the phase stability of the system is. To illustrate the effect of the phase instability, let us assume that we want

to average the interferometer outputs obtained every 0.2 sec. Each output is a phasor with a random amplitude and phase, and denoted by R(t). Let A(t) and $\phi(t)$ be the amplitude and phase of the phasor R(t). Then, we write

$$R(t) = A(t) \exp[j\,\phi(t)]. \tag{2.A.9}$$

Assuming that $\phi(t)$ is a normal random process with zero mean and variance σ_ϕ^2, we find the average value of R(t) as follows: We write

$$\overline{R(t)} = \overline{A(t) \exp[j\,\phi(t)]}\,. \tag{2.A.10}$$

Assume that A(t) and $\phi(t)$ are independent process, then we have

$$\overline{R(t)} = \overline{A(t)}\ \overline{\exp[j\,\phi(t)]}\,. \tag{2.A.11}$$

Assuming that

$$\overline{\exp[j\,\phi(t)]} \simeq E\{\exp[j\,\phi(t)\} \tag{2.A.12}$$

we write,

$$\overline{R(t)} \simeq \overline{A(t)} \int_{-\infty}^{\infty} e^{j\phi} p(\phi)\,d\phi$$

$$= \overline{A(t)} \int_{-\infty}^{\infty} e^{j\phi} \frac{1}{\sqrt{2\pi\,\sigma_\phi^2}}\, e^{-\frac{\phi^2}{2\sigma_\phi^2}}\, d\phi$$

$$= \overline{A(t)}\ \frac{1}{\sqrt{2\pi}\,\sigma_\phi} \int_{-\infty}^{\infty} \cos\phi\, e^{-\frac{\phi^2}{2\sigma_\phi^2}}\, d\phi$$

$$+ \overline{A(t)} \frac{1}{\sqrt{2\pi}\,\sigma_\phi} \int_{-\infty}^{\infty} \sin\phi\, e^{-\frac{\phi^2}{2\sigma_\phi^2}}\, d\phi \qquad (2.A.13)$$

Since $\sin\phi$ is odd function of ϕ, the second term in the above expression is zero. Thus,

$$\overline{R(t)} \simeq \overline{A(t)} \frac{1}{\sqrt{2\pi}\,\sigma_\phi} \int_{-\infty}^{\infty} \cos\phi\, e^{-\frac{\phi^2}{2\sigma_\phi^2}}\, d\phi$$

$$= \overline{A(t)}\, \exp(-\sigma_\phi^2/2) \qquad (2.A.14)$$

Thus, the phase noise degrades the fringe amplitude by a factor of $\exp(-\sigma_\phi^2/2)$. The effects of the phase noise can be easily seen in the power spectrum of the fringe function. Figure 2.a.2 shows the effects of the phase instabilities.

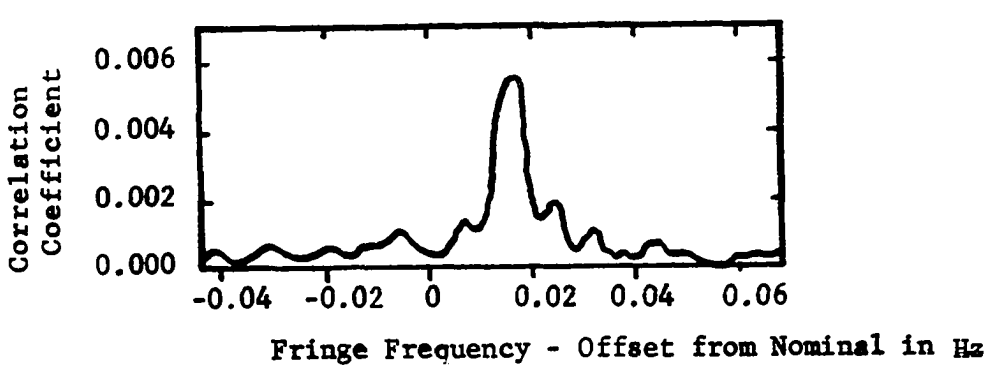


Figure 2.a.2. The effects of phase instabilities (After Clark et al., 1968)

B. Incoherent Averaging in the Frequency Domain

Let us define a phase instability function g(t)(Clark et al., 1968) satisfying

$$|g(t)| = 1 \tag{2.B.1}$$

and

$$f(t) = g(t)\, A, \tag{2.B.2}$$

where f(t) is the measured fringe function, A is the true complex fringe amplitude.

Taking the Fourier transform of both sides of Eq. (2.B.2) yields

$$F(\omega) = G(\omega) \circledast A\delta(\omega) \tag{2.B.3}$$

where $\circledast$ denotes convolution, $\delta(\omega)$ is Dirac delta function, and

$$G(\omega) = \int_{-\infty}^{\infty} g(t) e^{-j\omega t} dt \tag{2.B.4}$$

Parseval's equation states that

$$\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega = \frac{1}{T} \int_{0}^{T} |Ag(t)|^2 dt = |A|^2 \tag{2.B.5}$$

where T is the length of the observation. Thus,

$$|A| = \left[\int_{-\infty}^{\infty} |F(\omega)|^2 d\omega \right]^{\frac{1}{2}} \tag{2.B.6}$$

Even in the absence of the signal the integral in Eq. (2.B.6) will be nonzero; therefore, the estimate of the fringe amplitude given by Eq. (2.B.6) will be biased.

C. Coherent Averaging

Assume that we obtained the complex fringe function after integrating the two recorded interferometer signals, x(t) and y(t) (see Figure 1.d.1). Let the integration time be much shorter than the interferometer coherence time. That is to say, the phase of interferometer does not change considerably during the integration time. Then, the output of the interferometer can be represented as a complex phasor shown in Figure 2.c.1.

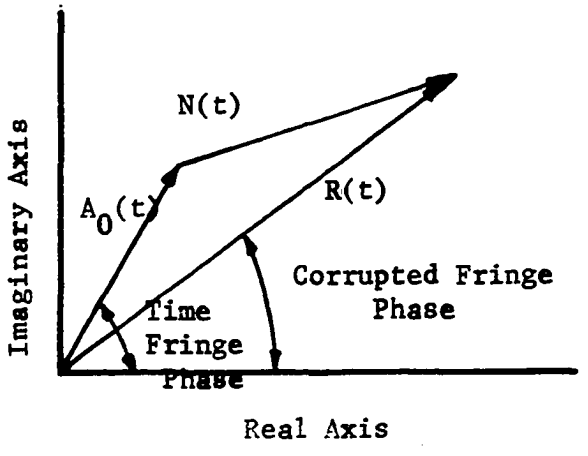


Figure 2.c.1. Phasor representation of the interferometer output voltage

The phasor R(t) has two components: A(t) and N(t). We write

$$R(t) = A(t) + N(t), \tag{2.C.1}$$

where A(t) is the true fringe phasor and N(t) is the noise phasor. N(t) originates from the receiver, sky-background noises, and the statistical sampling. If T_{a1}, T_{a2} are antenna temperatures and T_{s1}, T_{s2} are system temperatures, then the amplitudes of the true fringe phasor and noise phasor are given by (Rogers, 1970)

$$|A(t)| = (T_{a1}\, T_{a2})^{\frac{1}{2}}, \tag{2.C.2}$$

$$[|N(t)|^2]^{\frac{1}{2}} = (T_{s1}\, T_{s2}/2\, B\, T_c)^{\frac{1}{2}}\ , \qquad (2.C.3)$$

where B and T_c are the bandwidth of the receivers and coherent integration time, respectively. When the 1-bit correlation scheme is used, $[|N(t)|^2]^{\frac{1}{2}}$ increases by a factor of $\frac{\pi}{2}$ (Mutel, 1975). In the derivation of Eq. (2.C.2) it is assumed that $T_a \ll T_s$.

From Eq. (2.C.3) it is easy to see that as the integration time T_c increases the RMS value of the noise vector becomes smaller. But, we cannot integrate beyond the period of the coherence time of the system since the phase noise degrades the complex fringe function severely. As will be discussed in the next section, it is possible to use integration times much larger than the coherence time of the interferometer system by using the so-called broken-coherence technique.

D. Broken-Coherence Technique

The broken-coherence technique can be applied if the probability density distribution function of the fringe amplitude is known. Moran (1973) has calculated the probability density distribution functions of the fringe amplitude and the fringe phase for the weak and strong signal cases. We will only give the results for these two cases.

(a) Weak signal case: For this case $T_A \ll T_R$ and the probability density distribution function of the measured fringe amplitude R is given by

$$p(R) = \frac{R}{\sigma^2} \exp\left(-\frac{R^2+A^2}{2\sigma^2}\right) I_0\left(\frac{RA}{\sigma^2}\right),\ R>0 \qquad (2.D.1)$$

and for the measured fringe phase

$$p(\phi) = \frac{1}{2\pi} \exp\left(-\frac{A^2}{2\sigma^2}\right)\left\{1+\left(\frac{\pi}{2}\right)^{\frac{1}{2}} \frac{A\cos\phi}{\sigma} \exp\left(\frac{A^2\cos\phi}{2\sigma^2}\right) \times \left[1+\mathrm{Erf}\left(\frac{A\cos\phi}{\sqrt{2}\sigma}\right)\right]\right\}, \qquad (2.D.2)$$

where R = measured fringe amplitude

ϕ = measured fringe phase

$I_0(\cdot)$ = modified Bessel function of order zero and given by

$$I_0(X) = (1/2\pi)\int_0^{2\pi} \exp(X\cos\psi)d\psi \tag{2.D.3}$$

$$Erf(X) = (2/\sqrt{\pi})\int_0^{X} \exp(-\psi^2)d\psi \tag{2.D.4}$$

$$\sigma^2 = T_R/2BT_c \tag{2.D.5}$$

The expected values of R and R^2 are given by

$$E[R] = (\frac{\pi}{2})^{\frac{1}{2}}\sigma\exp(\frac{-A^2}{4\sigma^2})[(1+\frac{A^2}{2\sigma^2}) \times I_0(\frac{A^2}{4\sigma^2}) + \frac{A^2}{2\sigma^2} I_1(\frac{A^2}{4\sigma^2})] \tag{2.D.6}$$

$$E[R^2] = A^2 + 2\sigma^2 \tag{2.D.7}$$

where $I_1(\cdot)$ is the modified Bessel function of order 1. $E[\cdot]$ denotes expectation operation. The expected value of R is not A as seen from Eq. (2.D.6); therefore, a simple averaging does not provide an estimate for the true fringe amplitude A.

A technique used in statistical estimation is the maximum likelihood method. A detailed description of the method is given in Appendix A. If we have N measurements of the fringe amplitude, then the maximum likelihood function of N measurements is

$$L = \prod_{k=1}^{N} p(R_k), \tag{2.D.8}$$

or

$$L = \prod_{k=1}^{N} \frac{R_k}{\sigma^2}\exp(-\frac{R_k^2+A^2}{2\sigma^2})I_0(\frac{R_kA}{\sigma^2}). \tag{2.D.9}$$

The optimum estimate of the true fringe amplitude is the value of A which makes the derivative of the maximum likelihood function L with respect to A a maximum. The value of A which makes $\frac{\partial L}{\partial A}$ maximum, called A_e, is given approximately by

$$A_e^2 = [\frac{1}{N} \sum_{k=1}^{N} R_n^2] - 2\sigma^2 \qquad (2.D.10)$$

when the signal coming from a radio source is not present, σ^2 can be measured. The $2\sigma^2$ term is the bias and has to be removed. The signal-to-noise ratio for this case is given by (Moran, 1973)

$$\text{SNR}\Big|_{\text{incoherent}} \simeq (\frac{T_A}{T_R})^2 [BT(\frac{B}{B_f})]^{\frac{1}{2}}, \quad T_A \ll T_R(2BT)^{\frac{1}{2}} \qquad (2.D.11)$$

where $T_A = \sqrt{T_{a1}T_{a2}}$, $B_f = T_c^{-1}$, $T_R = \sqrt{T_{R_1}T_{R_2}}$, and T is the total integration time. For the 1-bit correlation case the signal-to-noise ratio is decreased by $\frac{\pi}{2}$. As seen from Eq. (2.D.11) B_f must be made as small as possible.

The minimum detectable fringe amplitude for the weak signal case can be found easily. Assuming a 5:1 signal-to-noise ratio for detection, we obtain the minimum detectable fringe amplitude by setting SNR = 5 in Eq. (2.D.11). This gives

$$\text{SNR}\Big|_{\text{incoherent}} = (\frac{T_A}{T_R})^2 [BT(\frac{B}{B_f})]^{\frac{1}{2}} = 5. \qquad (2.D.12)$$

Solving Eq. (2.D.12) for T_A yields

$$T_A(\text{min}) = \sqrt{5}\, T_R [BT(\frac{B}{B_f})]^{-\frac{1}{4}}$$

$$= \sqrt{5}\, T_R N^{-\frac{1}{4}} (BT_c)^{-\frac{1}{2}}, \qquad (2.D.13)$$

where $N = T/T_c$, the number of samples to be averaged incoherently.

The signal-to-noise ratio for the coherent averaging is given by (Rogers, 1970)

$$\left.SNR\right|_{\text{coherent}} = \frac{T_A}{T_R} (BT)^{\frac{1}{2}}. \tag{2.D.14}$$

Comparing Eqs. (2.D.11) and (2.D.14) yields

$$\left.SNR\right|_{\text{coherent}} = \sqrt{2}\ \left.SNR\right|_{\text{incoherent}}, \tag{2.D.15}$$

thus coherent averaging gives a higher signal-to-noise ration than does incoherent averaging.

(b) Strong signal case: For this case $A \gg \sigma$, the probability distributions of the fringe amplitude and the fringe phase are given by

$$p(R) \simeq \frac{1}{(2\pi)^{\frac{1}{2}}} \frac{1}{\sigma} \left(\frac{R}{A}\right)^{\frac{1}{2}} \exp\left[-\frac{(R-A)^2}{2\sigma^2}\right], \tag{2.D.16}$$

$$p(\phi) \simeq \frac{1}{(2\pi)^{\frac{1}{2}}} \frac{A}{\sigma} \exp\left(-\frac{A^2\phi^2}{2\sigma^2}\right). \tag{2.D.17}$$

From Eq. (2.D.16) we have

$$E[R] = A[1 + (\sigma^2/2A^2)], \tag{2.D.18}$$

and

$$\sigma_R \simeq \sigma. \tag{2.D.19}$$

From Eq. (2.D.17) we have

$$E[\phi] = 0, \tag{2.D.20}$$

and

$$\sigma_{\phi} \simeq \sigma/A \ . \tag{2.D.21}$$

If $A \gg \sigma$, clearly

$$E[R] \simeq A. \tag{2.D.22}$$

Thus, for the strong signal case the optimum estimate of the true fringe amplitude is simply the average of the measurements. Since for this case the p.d.f. of R is approximately normal, the average of the measurements is also a maximum likelihood estimate.

The signal-to-noise ratio is given by (Moran, 1973)

$$SNR \simeq \frac{1}{\sqrt{2}} \frac{T_A}{T_R} (BT)^{\frac{1}{2}}, \quad T_A \gg \frac{T_R}{(2BT)^{\frac{1}{2}}} \ . \tag{2.D.23}$$

The broken coherence method described before is not very good especially for the weak signal case. If one does not have an independent measurement of the noise temperatures of the interferometer receivers, the bias introduced by this method cannot be removed. In addition to the bias problem, the sensitivity is not very good compared to the coherent averaging method.

From Eq. (2.A.14) we also notice that, if the phase is noisy, the coherent averaging underestimates the fringe function. All the analysis for the incoherent averaging assumes that the phase noise originates from the receiver noises and the sampling errors. The real data we have analyzed do not have a zero mean and shows nonwhite properties. Therefore, the phase of the complex correlation function must be examined and the effects of the local oscillators and the ionosphere have to be removed before we

do coherent averaging. This process reduces the variance of the phase, and thus it increases the coherent integration time.

In the next chapter we will give a technique for removing the bias introduced by the phase noise during coherent integration. Then, the developed methods will be applied to the real interferometric data.

III. PROPOSED SOLUTION

A. Bias Recovery

As previously mentioned, one problem which limits performance of a VLB (Very-Long-Baseline) interferometer is the nonhomogenity of the atmosphere, ionosphere and the interstellar medium. At short wavelengths the atmosphere is more important, and at long wavelengths the ionosphere and interstellar medium dominate.

The inhomogenities add different random phase changes to the signals received by different antennas of an interferometer system. The more widely spaced the antennas, the more the two unknown phase changes will differ.

The phase of an interferometer output is the difference of the phases of the signals. In addition to the phase noise introduced by the ionosphere, interstellar medium and the atmosphere there is another phase noise due to imperfect local oscillators used at each antenna site.

We will attempt to correct the phase of an interferometer corrupted by the propagation media and the local oscillators. In order to make any corrections we must first detect the phase noise.

In the preceding chapter we described the NRAO Mark I recording system. In this system a record length is 0.2 second and the length of a typical observation is approximately 3 minutes. In this system we have the following phase noises (Whitney, 1974).

(a) Short-term noise: This noise has a time scale of 0.2 sec or less. It is impossible to detect because in the NRAO Mark I system the shortest data record is 0.2 sec. This type of noise reduces the estimate

of the complex fringe amplitude obtained every 0.2 sec.

Here we assume that effects of this type of noise are negligible. In practical systems this assumption may not hold due to the short-term local oscillator phase noises.

(b) Medium-term noise: This type of noise has a time scale of 3 minutes or less. It can be detected, and it is the phase noise for which we will attempt to make corrections. It has the same effect on the complex fringe amplitude as does the short-term phase noise.

(c) Long term noise: We define this type of phase noise as the noise having time scales of more than 3 minutes. This type of noise cannot be detected by using only one observation.

In the following discussion we will use the term "phase noise" for the medium-term phase noise.

In the previous chapter we saw that the presence of the finge phase noise reduces the amplitude of the complex fringe amplitude estimated by the coherent average.

The methods for reducing the phase noise effects on the estimation of the fringe amplitude were given in the previous chapter. Of these methods, the incoherent averaging is the most widely used because it does not require any additional equipment added to the standard interferometer. However, this method overestimates the complex fringe amplitude by a factor of $1+0.5(SNR^{-2})$ where SNR is the signal-to-noise ratio of the interferometer.

Now let us assume that we have n estimates of the complex fringe amplitude each obtained by coherent averaging over short time intervals. Further, we assume that the bias on each estimate is negligible.

We want to average these n estimates coherently while we take phase values of each estimate. These phase samples will be used to identify the fringe phase noise so that we can recover the bias.

Obtaining n coherent estimates of the complex fringe amplitude and later averaging them coherently is the same as if we average the two interferometer signals for the time nT_c, where T_c is the coherent averaging time to obtain each one of n estimates.

Later it will be clear why we first obtain n estimates for the complex fringe amplitude. When the 1-bit correlation scheme is used, the interferometer output is the normalized cross-correlation. Therefore, in the following discussion we will use the term "fringe amplitude" for the normalized correlation.

In the presence of a fringe phase noise the complex fringe amplitude is underestimated by a factor of $\exp(-\sigma_\phi^2/2)$ when the coherent average is used, where σ_ϕ^2 is the variance of the phase noise process. We also assume that the phase noise process is normally distributed with zero mean.

Let $\hat{R}$ be an estimator for the complex fringe amplitude, obtained by coherent averaging. We previously said that this estimate is biased. Now let us define a new estimator as

$$Z = \hat{R} \exp(\sigma_\phi^2/2) \tag{3.A.1}$$

where $\hat{R}$ and σ_ϕ^2 were defined previously. The estimator Z defined by Eq. (3.A.1) is not practical since we do not know the exact value of σ_ϕ^2. If we replace σ_ϕ^2 in Eq. (3.A.1) by $\hat{\sigma}_\phi^2$, the sample variance of the phase noise process, we obtain

$$\hat{Z} = \hat{R} \exp(\hat{\sigma}_{\phi}^2/2) \tag{3.A.2}$$

Before we use $\hat{Z}$ as an estimator for the complex fringe amplitude we must check its bias and variance properties.

To obtain the bias of $\hat{Z}$ let us take the expectation of both sides of Eq. (3.A.2). Upon doing so we obtain

$$E[\hat{Z}] = E[\hat{R} \exp(\hat{\sigma}_{\phi}^2/2)] \tag{3.A.3}$$

where $E[\cdot]$ denotes the expectation operator. For the time being, let us assume that $\hat{R}$ and $\hat{\sigma}_{\phi}^2$ are independent random variables (an estimator is a random variable). Then Eq. (3.A.3) can be written as

$$E[\hat{Z}] = E[\hat{R}] E[\exp(\hat{\sigma}_{\phi}^2/2)] \tag{3.A.4}$$

From Chapter II, Eq. (2.A.14) we have

$$E[\hat{R}] = R \exp(-\sigma_{\phi}^2/2) \tag{3.A.5}$$

where R is the true fringe amplitude. Substituting Eq. (3.A.5) into Eq. (3.A.4) yields

$$\begin{aligned} E[\hat{Z}] &= R\, E[\exp(-\sigma_{\phi}^2/2)\, \exp(\hat{\sigma}_{\phi}^2/2)] \\ &= R\, E[\exp(\frac{\hat{\sigma}_{\phi}^2 - \sigma_{\phi}^2}{2})] \end{aligned} \tag{3.A.6}$$

We assumed previously that the fringe phase noise process is a normal, ergodic process with zero mean. Therefore, we can use time-averaging instead of expectation operators to find the variance of the process, i.e.,

$$\hat{\sigma}_{\phi}^2 = \frac{1}{n-1} \sum_{i=0}^{n-1} \phi_i^2 , \quad i = 0, 1, 2, \ldots, n-1$$

where ϕ_i is the phase sample taken at time $i\Delta T$, ΔT is the sampling period, and n is the number of samples.

The variance of the sample variance $\hat{\sigma}^2_\phi$ is given (Meyer, 1975, p. 273) by

$$\mathrm{Var}(\hat{\sigma}^2_\phi) = \frac{2(\sigma^2_\phi)^2}{n-1} \tag{3.A.7}$$

where n is the number of samples used to calculate $\hat{\sigma}^2_\phi$. As seen from Eq. (3.A.7) as $n \to \infty$, $\mathrm{Var}(\hat{\sigma}^2_\phi) = 0$. This means that $E[\hat{\sigma}^2_\phi] \to \sigma^2_\phi$ as $n \to \infty$. Therefore, we write,

$$\lim_{n \to \infty} E\left[\exp\left(\frac{\hat{\sigma}^2_\phi - \sigma^2_\phi}{2}\right)\right] = 1 \tag{3.A.8}$$

Thus,

$$E[\hat{Z}] = R \tag{3.A.9}$$

A more rigorous proof of Eq. (3.A.9) is the following. If the $\hat{\sigma}^2_\phi$ is a sample variance of a normal, zero mean process $\phi(t)$, then the random variable $[(n-1)\hat{\sigma}^2_\phi/\sigma^2_\phi]$ is χ^2 distributed with $(n-1)$ degrees of freedom (Meyer, 1975, p. 272). In terms of the new variable we write Eq. (3.A.6) as

$$E[\hat{Z}] = R\, e^{\sigma^2_\phi/2}\, E[\exp(kx)] \tag{3.A.10}$$

where

$$x = \frac{(n-1)\hat{\sigma}^2_\phi}{\sigma^2_\phi}, \tag{3.A.11}$$

$$k = \sigma^2_\phi/2(n-1) \tag{3.A.12}$$

By the definition of the expected value we have

$$E[\exp(kx)] = \int_0^\infty e^{kx} p(x)dx \tag{3.A.13}$$

where p(x) is the probability density function of x and is given by

$$p(x) = \frac{x^{\frac{1}{2}(\nu-2)}}{\Gamma(\nu/2)2^{\nu/2}} e^{-x/2} \tag{3.A.14}$$

where $\Gamma(\cdot)$ is the Gamma function, ν is the number of degrees of freedom.

Substituting Eq. (3.A.1) into Eq. (3.A.1) gives

$$E[\exp(kx)] = \int_0^\infty \exp(kx) \frac{x^{\frac{1}{2}}}{\Gamma(\nu/2)2^{\nu/2}} e^{-x/2}dx$$

$$= \frac{1}{\Gamma(\nu/2)^{\nu/2}} \int_0^\infty e^{-cx} x^{(\frac{\nu}{2}-1)} dx \tag{3.A.15}$$

where

$$c = \frac{1}{2} - k \tag{3.A.16}$$

Evaluating the integral in Eq. (3.A.16) yields

$$E[\exp(kx)] = \frac{\Gamma^{(\frac{\nu}{2})}}{\Gamma(\nu/2)2^{\nu/2}c^{\nu/2}}. \tag{3.A.17}$$

Substituting Eqs. (3.A.13) and (3.A.17) into Eq. (3.A.18) yields

$$E[\exp \frac{\hat{\sigma}^2_\phi}{2}] = \frac{(n-1)^{\frac{1}{2}(n-1)}}{[(n-1) - \sigma^2_\phi]^{\frac{1}{2}(n-1)}}, \tag{3.A.18}$$

where $n - 1 = \nu$. Eq. (3.A.18) can be written as

$$E[\exp(\hat{\sigma}_\phi^2/2)] = (n-1)^{\frac{1}{2}(n-1)} [(n-1) - \sigma_\phi^2]^{-\frac{1}{2}(n-1)} \tag{3.A.19}$$

If we let

$$m = n - 1 \tag{3.A.20}$$

Equation (3.A.19) becomes

$$E[\exp(\hat{\sigma}_\phi^2/2)] = m^{\frac{1}{2}m} (m - \sigma_\phi^2)^{-\frac{1}{2}m}. \tag{3.A.21}$$

By using the binomial series expansion we write Eq. (3.A.21) as

$$E[\exp(\hat{\sigma}_\phi^2/2)] = [1 + (\frac{\sigma_\phi^2}{2}) + \frac{1}{2}(\frac{\sigma_\phi^2}{2})^2 + \frac{1}{6}(\frac{\sigma_\phi^2}{2})^3 + \ldots\ldots]$$
$$+ [\frac{1}{2m}\sigma_\phi^2 + (\frac{1}{2m} + \frac{3}{4m} + \frac{1}{m^2})\sigma_\phi^2 + \ldots\ldots] \tag{3.A.22}$$

For a bounded σ_ϕ^2, as $m \to \infty$ Eq. (3.A.22) becomes

$$E[\exp(\hat{\sigma}_\phi^2/2)] = 1 + \frac{\sigma_\phi^2}{2} + \frac{1}{2}(\frac{\sigma_\phi^2}{2}) + \frac{1}{6}(\frac{\sigma_\phi^2}{2})^3 + \ldots\ldots \tag{3.A.23}$$

We recognize that the right hand side of the above equation is the series expansion of $\exp(\sigma_\phi^2/2)$. Therefore, we write

$$E[\exp(\hat{\sigma}_\phi^2/2)] = \exp(\sigma_\phi^2/2). \tag{3.A.24}$$

Substituting this into Eq. (3.A.6) yields $E[\hat{Z}] = R$.

Thus, we proved that the estimator $\hat{Z}$ given by Eq. (3.A.2) is an unbiased estimator. Next we will find the variance of $\hat{Z}$. The variance of $\hat{Z}$ is the mean square error made in the estimation of the complex fringe amplitude R. From the well-known error propagation formula (Meyer, 1975), and Eq. (3.A.2)

$$\mathrm{Var}(\hat{Z}) = \left(\frac{\partial \hat{Z}}{\partial \hat{R}}\right)^2 \mathrm{Var}(\hat{R}) + \left(\frac{\partial \hat{Z}}{\partial \hat{\sigma}_\phi^2}\right)^2 \mathrm{Var}(\hat{\sigma}_\phi^2) \tag{3.A.25}$$

Taking the partial derivatives of the right hand side of Eq. (3.A.2) with respect to $\hat{R}$ and $\hat{\sigma}_\phi^2$, respectively, yields

$$\frac{\partial \hat{Z}}{\partial \hat{R}} = \exp(\hat{\sigma}_\phi^2/2) \ , \tag{3.A.26}$$

and

$$\frac{\partial Z}{\partial \hat{\sigma}_\phi^2} = \frac{1}{2} \hat{R} \exp(\hat{\sigma}_\phi^2/2). \tag{3.A.27}$$

We substitute Eqs. (3.A.27) and (3.A.28) to get

$$\mathrm{Var}(\hat{Z}) = \exp(\hat{\sigma}_\phi^2)\ \mathrm{Var}(\hat{R}) + \frac{1}{4} \hat{R}^2 \exp(\hat{\sigma}_\phi^2)\mathrm{Var}(\hat{\sigma}_\phi^2). \tag{3.A.28}$$

$\mathrm{Var}(\hat{\sigma}_\phi^2)$ was given by Eq. (3.A.7) as $2(\sigma_\phi^4)/(n-1)$. If the 1-bit correlation technique is used $\mathrm{Var}(\hat{R})$ is given by (Rogers, 1976)

$$\mathrm{Var}(\hat{R}) = \left(\frac{\pi}{2}\right)^2 \frac{1}{2BT} \tag{3.A.29}$$

where T is the coherent integration time and equal to $n\Delta T$, and B is the bandwidth of the receivers used. Substituting the expressions for $\mathrm{Var}(\hat{\sigma}_\phi^2)$ and $\mathrm{Var}(\hat{R})$ into Eq. (3.A.29) yields

$$\mathrm{Var}(\hat{Z}) = \exp(\hat{\sigma}_\phi^2)\ \left(\frac{\pi}{2}\right)^2 \frac{1}{2BT} + \frac{\sigma_\phi^4\, \hat{R}^2}{2(n-1)} \exp(\hat{\sigma}_\phi^2). \tag{3.A.30}$$

If we let σ_ϕ^2 be equal to zero, Eq. (3.A.31) becomes

$$\mathrm{Var}(\hat{Z}) = \left(\frac{\pi}{2}\right)^2 \frac{1}{2BT} \quad , \qquad \sigma_\phi^2 = 0. \tag{3.A.31}$$

which is the same as Eq. (3.A.30).

From Eq. (3.A.30) the uncertainty on the estimate of the fringe amplitude is

$$\sigma_{\hat{Z}} = \left[\frac{\pi^2}{8BT}\exp(\hat{\sigma}_\phi^2) + \frac{\sigma_\phi^4}{2(n-1)}\hat{R}^2\exp(\hat{\sigma}_\phi^2)\right]^{\frac{1}{2}}. \qquad (3.A.32)$$

Substituting Eq. (3.A.2) into Eq. (3.A.32) yields

$$\sigma_{\hat{Z}} = \left[\frac{\pi^2}{8BT}\exp(\hat{\sigma}_\phi^2) + \sigma_\phi^4\hat{Z}^2/2(n-1)\right]^{\frac{1}{2}}. \qquad (3.A.33)$$

To apply the bias recovery technique, an estimate of the variance σ_ϕ^2, and the coherent average $\hat{R}$, are obtained. Then, by using Eq. (3.A.2) the new estimate of the fringe amplitude $\hat{Z}$ is obtained. As seen from Eq. (3.A.33), when the bias recovery technique is used, the uncertainty on the final estimate $\hat{Z}$ is larger than the uncertainty on $\hat{R}$. Therefore, when σ_ϕ^2 is large the technique should not be used.

B. Phase Noise Modeling

Let the measured fringe phase at time t be $\phi(t)$. To remove some predictable components from $\phi(t)$ we model $\phi(t)$ as

$$\phi(t) = \phi_{source}(t) + \phi_{iono.}(t) + \phi_{clock}(t) + \phi_{stat}(t). \qquad (3.B.1)$$

The terms on the right hand side of Eq. (3.B.1) were defined in the previous chapter. For simplicity, we assume that the source observed is a point source, therefore the first term of the right hand side of Eq. (3.B.1) is zero when there is no source position and baseline errors. Then, Eq. (3.B.1) reduces to

$$\phi(t) = \phi_{iono.}(t) + \phi_{clock}(t) + \phi_{stat}(t) \qquad (3.B.2)$$

The components $\phi_{iono}(t)$, $\phi_{clock}(t)$ and $\phi_{stat}(t)$ are random stochastic

processes of time. In order to model each component we have to understand whether this process is ergodic so that time average properties of the process may be used in the modeling.

Now we will consider each component separately.

The clock noise model: We define the clock phase noise as the difference between the phases of the atomic time standards used in a VLBI experiment. It is well-known that a rubidium Time Standard has a phase drift in addition to a white noise component. Let $\phi_{c1}(t)$ the phase of the standard at the first VLBI site. Then we write

$$\phi_{c1}(t) = c_1 t + \epsilon_1(t) \tag{3.B.3}$$

where c_1 is the slope of the linear phase drift and $\epsilon_1(t)$ is a white noise process. As the phase of the second standard we write

$$\phi_{c2}(t) = c_2 t + \epsilon_2(t) \tag{3.B.4}$$

where c_2 is the slope of the drift and $\epsilon_2(t)$ is the white noise process. At the interferometer output we observe the differential phase

$$\phi_{clock}(t) = \phi_{c1}(t) - \phi_{c2}(t) \tag{3.B.5}$$

Substituting Eqs. (3.B.3) and (3.B.4) into Eq. (3.B.5) yields

$$\phi_{clock}(t) = (c_1 - c_2)t + \epsilon_1(t) - \epsilon_2(t) \tag{3.B.6}$$

If we let

$$c = c_1 - c_2, \tag{3.B.7}$$

$$\epsilon(t) = \epsilon_1(t) - \epsilon_2(t) \tag{3.B.8}$$

Then Eq. (3.B.6) becomes

$$\phi_{clock}(t) = ct + \epsilon(t) \tag{3.B.9}$$

Let us assume that c_1 and c_2 are Gaussian random variables. This assumption can be justified by using the central limit theorem. In an atomic standard, the phase drift is caused by many factos. No matter in what manner an individual factor affects the phase of the standard, over all effects will be distributed normally.

We saw in Chapter I that at decametric wavelengths the phase noise component due to the ionosphere dominates all other phase noise components. Therefore, we will assume that the random process $\epsilon(t)$ is almost zero, i.e., $e(t) \simeq 0$ for all t. Then we write Eq. (3.B.9) as

$$\phi_{clock}(t) = ct \tag{3.B.10}$$

Since c is a random variable, then, $\phi_{clock}(t)$ is a random process. The mean square value of the process is given by

$$\begin{aligned} \langle \phi_{clock}(t) \cdot \phi_{clock}(t) \rangle &= \langle ct \cdot ct \rangle \\ &= \langle c^2 \rangle \, t^2 \end{aligned} \tag{3.B.11}$$

where angular brackets mean ensemble averaging. The mean square value of $\phi_{clock}(t)$ is not the same for all t; therefore, this process is not stationary.

The random process defined by Eq. (3.B.10) is called a random ramp process. It can be thought of as the output of an integrator driven by a random constant c. At time zero the initial condition of the integrator is taken to be zero. This process may be represented by a set of two dif-

ferential equations

$$\dot{\phi}_{clock}(t) = \phi_{aux\ 2}(t) \tag{3.B.12}$$

$$\dot{\phi}_{aux\ 2}(t) = 0 \tag{3.B.13}$$

where the dots denote the time derivative operation, $\phi_{aux\ 2}(t)$ is an auxiliary random process whose initial condition $\phi_{aux\ 2}(t)$ determines the slope of the random ramp process. The initial condition $\phi_{aux\ 2}(0)$ is equal to c in Eq. (3.B.10).

The value of c may be approximated by comparing the two frequency standards after the experiment. The slope c may be found by fitting the phase noise to a linear function of time, i.e.,

$$\phi(t) = c_0 + \hat{c}t + \Delta\phi(t)$$

where c_0 is a constant, $\hat{c}$ is the estimate of the slope and $\Delta\phi(t)$ is the residual noise.

Ionospheric noise: The electron density irregularities in the ionosphere have been observed to travel across the wavefront with velocities up to 200 m/sec (Al'pert, 1973, p. 177).

Let us assume that the electron density irregularities are formed by the spherical blobs with a random spatial distribution as shown in Figure 3.b.1.

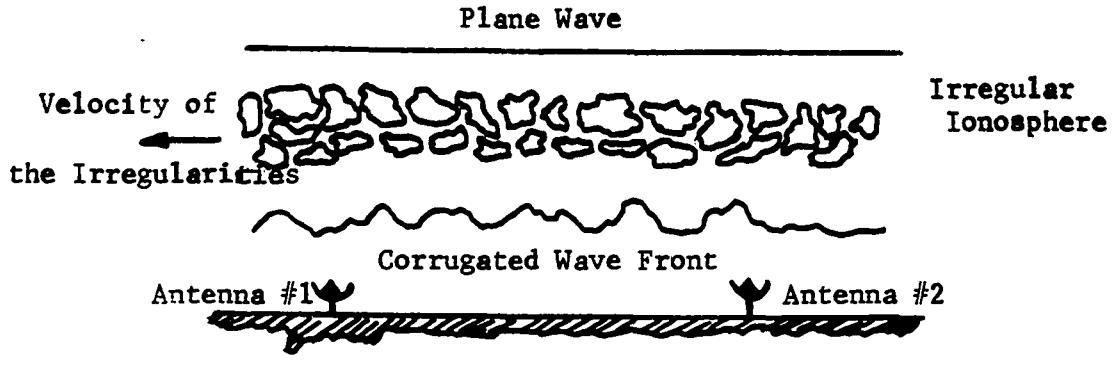


Figure 3.b.1. A typical irregularity structure in the ionosphere

The antennas in Figure 3.b.1 serves as phase detectors. The difference between antenna phases changes periodically as the structure moves across the wavefront. The data taken during VLBI experiments between Ames, Iowa and Boulder, Colorado in August 1971 showed a periodic behavior.

During these experiments the ionosphere was the primary source of phase noise (Clark and Ericson, 1973). Therefore, we attributed the periodic phase behavior to the ionosphere as predicted by the irregularity model given previously.

By periodically we do not mean a perfect periodic behavior. Since the electron density blobs are not distributed uniformly across the wavefront we would not expect a perfect periodicity.

We model the differential phase noise $\phi_{iono}(t)$ as a periodic random process defined by the equation

$$\phi_{iono}(t) = \varphi(t) \cos [\omega t + \theta(t)] \tag{3.B.14}$$

where $\varphi(t)$ and $\theta(t)$ are random, time varying amplitude and phase, respec-

tively, and ω is the angular frequency of the random periodic process $\phi_{iono}(t)$.

Eq. (3.B.14) can be written as

$$\phi_{iono}(t) = \mathrm{Re}\{\varphi(t)\ e^{j[\omega t + \theta(t)]}\} \tag{3.B.15}$$

where $\mathrm{Re}\{\cdot\}$ implies that the real part of the complex expression inside the parentheses is to be taken.

By definition the autocorrelation function of $\phi_{iono}(t)$ is given by

$$R_{\phi}(\tau) = E[\phi_{iono}(t)\ \phi_{iono}(t+\tau)] \tag{3.B.16}$$

where $E[\cdot]$ is the expected value operator. Let

$$\tilde{\phi}_{iono}(t) = \varphi(t)\ e^{j[\omega t + \theta(t)]} \tag{3.B.17}$$

where the tilde over $\phi_{iono}(t)$ means that $\tilde{\phi}_{iono}(t)$ is a complex function. By the use of the equality

$$E[\phi_{iono}(t)\ \phi_{iono}(t+\tau)] = \frac{1}{2}\,\mathrm{Re}\{\tilde{\phi}_{iono}(t)\ \tilde{\phi}^{*}_{iono}(t+\tau)\} \tag{3.B.18}$$

we write

$$R_{\phi}(\tau) = \frac{1}{2}\,\mathrm{Re}\{E[\varphi(t)\varphi(t+\tau)\ e^{j[\omega t + \theta(t)]} \cdot e^{-j[\omega(t+\tau) + \theta(t+\tau)]}]\} \tag{3.B.19}$$

Assuming that $\varphi(t)$ and $\theta(t)$ are independent random processes we write Eq. (3.B.19) as

$$R_{\phi}(\tau) = \frac{1}{2} \mathrm{Re}\{E[\varphi(t)\varphi(t+\tau)]E\,[e^{j[\theta(t) - \omega\tau - \theta(t+\tau)]}]\}$$

$$= \frac{1}{2} R_{\varphi}(\tau) R_e\{e^{-j\omega\tau} E[e^{j[\theta(t) - \theta(t+\tau)]}]\} \qquad (3.B.20)$$

where

$$R_{\varphi}(\tau) = E[\varphi(t)\varphi(t+\tau)] \qquad (3.B.21)$$

If $\theta(t)$ is an ergodic process with zero mean we write

$$\sigma^2_{\theta(t)} = \sigma^2_{\theta(t+\tau)} = \sigma^2_{\theta} \qquad (3.B.22)$$

where σ^2_{θ} is the variance of the process $\theta(t)$. Assuming that $\theta(t)$ is a normal stochastic process, we use Eqs. (1.F.13) and (1.F.20) to get

$$E[e^{j[\theta(t) - \theta(t+\tau)]}] = e^{-\sigma^2_{\theta}[1 - r(\tau)]}, \qquad (3.B.23)$$

where $r(\tau)$ is the normalized correlation function of $\theta(t)$.

Substituting Eq. (3.B.23) into Eq. (3.B.20) yields

$$R_{\phi}(\tau) = \frac{1}{2} R_{\varphi}(\tau)\ \mathrm{Re}\{e^{-j\omega\tau\ -\sigma^2_{\theta}(1 - r(\tau))}\}$$

$$= \frac{1}{2} R_{\varphi}(\tau)\ e^{-\sigma^2_{\theta}[1 - r(\tau)]} \cos \omega\tau\ . \qquad (3.B.24)$$

We assumed that the process $\phi_{iono}(t)$ was ergodic and Gaussian, therefore the autocorrelation function of $\phi_{iono}(t)$ must satisfy the condition (Davenport and Root, 1958, p. 67)

$$\int_{-\infty}^{+\infty} |R_{\phi}(\tau)|\,d\tau < +\infty \qquad (3.B.25)$$

or

$$\int_{-\infty}^{\infty} \left|\frac{1}{2} R_{\varphi}(\tau)\, e^{-\sigma_\theta^2[1 - r(\tau)]} \cos \omega\tau\right| d\tau < +\infty \qquad (3.B.26)$$

To satisfy the condition just given we set

$$\frac{1}{2} R_{\varphi}(\tau) e^{-\sigma_\theta^2[1 - r(\tau)]} = \varphi_0^2 e^{-\beta\tau} \qquad (3.B.27)$$

where φ_0^2 is the variance of the process $\phi_{ion}(t)$, and β is a positive parameter determined by the process.

We substitute Eq. (3.B.27) into Eq. (3.B.26) and get

$$\int_{-\infty}^{+\infty} \left|\varphi_0^2 e^{-\beta\tau} \cos \omega\tau\right| d\tau < +\infty,\ \beta > 0 \qquad (3.B.28)$$

For $\varphi_0^2 < \infty$ the above inequality is satisfied.

The parameters ω, β and φ_0^2 may be obtained empirically. We compute the autocorrelation function of the phase noise process $\phi(t)$ and by a simple curve fitting technique obtain approximate values of φ_0^2, ω and β.

The process given by Eq. (3.B.14) may be represented by the following stochastic differential equations (Gelb, 1974, p. 82)

$$\dot{\phi}_{iono}(t) = \phi_{aux\ 1}(t) + w(t) \qquad (3.B.29)$$

$$\dot{\phi}_{aux\ 1}(t) = -\alpha^2 \phi_{iono}(t) - 2\beta\, \phi_{aux\ 1}(t) + (\alpha - 2\beta) w(t) \qquad (3.B.30)$$

where β is the parameter defined by Eq. (3.B.27), α is equal to $\sqrt{\beta^2 + \omega^2}$, ω is the angular frequency of the random periodic process $\phi_{iono}(t)$. $w(t)$ is the white noise process whose variance is $2\beta\varphi_0^2$ where φ_0^2 was given by Eq. (3.B.27). $\phi_{aux\ 1}(t)$ is the auxiliary random variable.

A block diagram of the periodic random process is given in Figure 3.b.2.

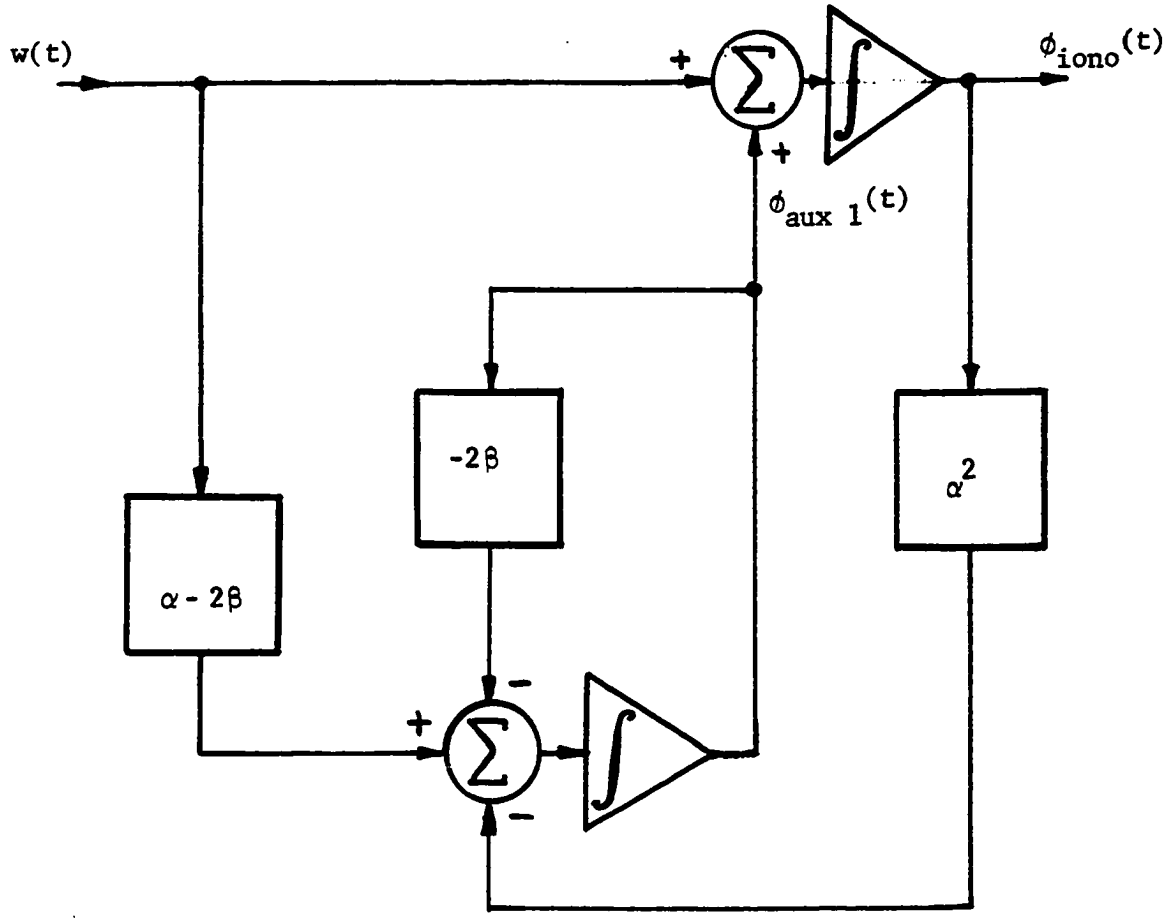


Figure 3.b.2. The block diagram for the generation of a periodic random process

Statistical noise model: The statistical noise in Eq. (3.B.2) may be modeled as a white noise process.

The variance of this process is given by (Moran, 1973)

$$\sigma^2_{\phi_{stat}} \simeq [T_1 T_2/2BT]/R^2, \tag{3.B.31}$$

provided that

$$T_A \gg \sqrt{T_1 T_2}/2BT \tag{3.B.32}$$

where T_1 and T_2 are the receiver temperatures of the interferometer, respectively, T_A is the antenna temperature, B is the bandwidth of the receivers and T is the integration time.

In cases where the ionosphere is a main contributor to the fringe phase noise, $\sigma^2_{\phi_{stat}}$ may be taken zero.

We assumed that the radio source observed was a point source. If the source is not a point source, then $\phi_{source}(t)$ in Eq. (3.B.1) cannot be taken zero. For short time periods we may take $\phi_{source}(t)$ as a constant. The value of this constant is determined by the structure of a radio source.

Since we never know the exact value of $\phi_{source}(t)$ it can be considered as a constant random process.

A constant random process may be represented by the equation

$$\dot{\phi}_{source}(t) = 0 \ . \tag{3.B.33}$$

C. Phase Noise Removal

In the previous section each phase noise component was modeled in terms of stochastic differential equations. In these equations the dependent variables are called state variables.

We may think each noise component in Eq. (3.B.1) as the output of a linear system driven by a white noise input. Then the differential equations given for each phase noise component are in fact state equations of the corresponding linear system.

Now we are in a position to attempt removing the phase noise components from the measured fringe phase. First, we will formulate the problem for Kalman filtering as follows.

Before we proceed let us rewrite differential equations describing different phase noise components

$$\dot{\phi}_{source}(t) = 0, \tag{3.C.1}$$

$$\dot{\phi}_{iono}(t) = \phi_{aux\ 1}(t) + w(t), \tag{3.C.2}$$

$$\dot{\phi}_{aux\ 1}(t) = -\alpha^2\phi_{iono}(t) - 2\beta\,\phi_{aux\ 1}(t) + (\alpha - 2\beta)w(t) \tag{3.C.3}$$

$$\dot{\phi}_{clock}(t) = \phi_{aux\ 2}(t), \tag{3.C.4}$$

$$\dot{\phi}_{aux\ 2}(t) = 0. \tag{3.C.5}$$

The measured fringe phase noise is given by

$$\phi(t) = \phi_{source}(t) + \phi_{iono}(t) + \phi_{clock}(t) + \phi_{stat}(t) \tag{3.C.6}$$

Let us assume that various parameters describing each noise component are estimated. We write Eq. (3.C.6) as a matrix form as

$$\phi(t) = (1\ 1\ 0\ 1\ 0) \begin{bmatrix} \phi_{source}(t) \\ \phi_{iono}(t) \\ \phi_{aux\ 1}(t) \\ \phi_{clock}(t) \\ \phi_{aux\ 2}(t) \end{bmatrix} + \phi_{stat}(t) \qquad (3.C.7)$$

We inserted the two auxiliary variables into the vector for convenience. Let us write Eqs. (e.C.1) - (3.C.5) in a matrix form

$$\begin{bmatrix} \dot{\phi}_{source}(t) \\ \dot{\phi}_{iono}(t) \\ \dot{\phi}_{aux\ 1}(t) \\ \dot{\phi}_{clock}(t) \\ \dot{\phi}_{aux\ 2}(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -\alpha^2-2\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \phi_{source}(t) \\ \phi_{iono}(t) \\ \phi_{aux\ 1}(t) \\ \phi_{clock}(t) \\ \phi_{aux\ 2} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ \alpha-2\beta \\ 0 \\ 0 \end{bmatrix} w(t) \qquad (3.C.8)$$

The statistical noise $\phi_{stat}(t)$ was modeled as white noise. We want to estimate $\phi_{source}(t)$ given the noisy measurement $\phi(t)$. Eqs. (3.C.7) and (3.C.8) are written in a way that the Kalman filter algorithm may be applied (see Appendix B). Therefore, we can use this filter to estimate $\phi_{source}(t)$. The Kalman filter algorithm is quite sensitive if we do not give right initial values of the signal to be estimated. Thus in the absence of any a priori information about $\phi_{source}(t)$, we estimate $\phi_{iono}(t)$ and $\phi_{clock}(t)$ as if they were the desired signals, we then obtain the estimate for $\phi_{source}(t)$ by

$$\hat{\phi}_{source}(t) = \phi(t) - (\phi_{iono}(t) - \phi_{clock}(t)). \qquad (3.C.9)$$

The Kalman filter is a recursive algorithm used in many areas. It is discussed in Appendix B. To be consistent with the notation given in Appendix B, let us change variables as follows:

$$\phi_{source}(t) = X_1(t) \qquad \phi_{clock}(t) = X_4(t)$$
$$\phi_{iono}(t) = X_2(t) \qquad \phi_{aux\ 2}(t) = X_5(t)$$
$$\phi_{aux\ 1}(t) = X_3(t) \qquad \phi_{stat}(t) = n(t)$$
$$\phi(t) = Z(t)$$

Substituting these new variables into Eq. (3.C.8) yields the matrix equation.

$$\begin{bmatrix} \dot{X}_1(t) \\ \dot{X}_2(t) \\ \dot{X}_3(t) \\ \dot{X}_4(t) \\ \dot{X}_5(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -\alpha^2 & -2\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \\ X_4(t) \\ X_5(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ \alpha - 2\beta \\ 0 \\ 0 \end{bmatrix} w(t) \qquad (3.C.10)$$

or

$$\dot{\underline{X}}(t) = \underset{\sim}{F}\, X(t) + \underset{\sim}{G}\, \underline{w}\,(t) \qquad (3.C.11)$$

where $$\underline{X}^T(t) = [X_1(t) \quad X_2(t) \quad X_3(t) \quad X_4(t) \quad X_5(t)]\,, \qquad (3.C.12)$$

$$\underset{\sim}{G}^T = [0 \quad 1 \quad \alpha - 2\beta \quad 0 \quad 0]\,, \qquad (3.C.13)$$

$$\underset{\sim}{F} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -\alpha^2 & -2\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} . \qquad (3.C.14)$$

Substituting new variables into Eq. (3.C.7) yields

$$\underline{Z}(t) = (1\ 1\ 0\ 1\ 0) \begin{bmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \\ X_4(t) \\ X_5(t) \end{bmatrix} + n(t) \qquad (3.C.15)$$

or

$$\underline{Z}(t) = \underset{\sim}{S}\,\underline{X} + \underline{n}(t) \qquad (3.C.16)$$

where $\underline{Z}(t)$ = 1×1 measurement vector, having $Z(t)$ as the only element

$\underline{n}(t)$ = 1×1 measurement noise vector, has only $n(t)$ as its element.

$$\underset{\sim}{S} = (1\ 1\ 0\ 1\ 0) \qquad (3.C.17)$$

Now we have $\underline{X}(t)$ as the state vector, $\underline{Z}(t)$ as measurement vector and $\underline{n}(t)$ as the measurement noise vector.

We obtain the state transition matrix by using Eq. (9.19) given in Appendix B. The transition matrix is found to be

$$\underset{\sim}{H}(t,0) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & h_{22}(t) & h_{23}(t) & 0 & 0 \\ 0 & h_{32}(t) & h_{33}(t) & 0 & 0 \\ 0 & 0 & 0 & 1 & t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.C.18}$$

where $$h_{22}(t) = 1 + \frac{at^2}{2} + \frac{abt^3}{6} \tag{3.C.19}$$

$$h_{23}(t) = t + \frac{bt^2}{2} + \frac{(a+b^2t^3)}{6} \tag{3.C.20}$$

$$h_{32}(t) = at + ab\,\frac{t^2}{2} + \frac{a(a+b^2)t^3}{6} \tag{3.C.21}$$

$$h_{33}(t) = 1 + bt + (a+b^2)\,\frac{t^2}{2} + \frac{2ab+b^3}{6}\,t^3 \tag{3.C.22}$$

where $a = -\alpha^2$, $b = -2\beta$, α and β have defined before in this chapter.

The transition matrix for the discrete case is obtained from Eq. (8) in Appendix C as

$$\underset{\sim}{H}_k = \underset{\sim}{H}(\Delta t,0) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & h_{22}(\Delta t) & h_{23}(\Delta t) & 0 & 0 \\ 0 & h_{32}(\Delta t) & h_{33}(\Delta t) & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.C.23}$$

where Δt = sampling period

$$h_{22}(\Delta t) = 1 + \frac{a}{2}(\Delta t)^2 + \frac{ab}{6}(\Delta t)^3 \tag{3.C.24}$$

$$h_{23}(\Delta t) = \Delta t + \frac{b}{2}(\Delta t)^2 + \frac{a+b^2}{6}(\Delta t)^3 \tag{3.C.25}$$

$$h_{32}(\Delta t) = a(\Delta t) + \frac{ab}{2}(\Delta t)^2 + \frac{a(a+b^2)}{6}(\Delta t)^3 \tag{3.C.26}$$

$$h_{33}(\Delta t) = 1 + b\,\Delta t + (a+b^2)\frac{(\Delta t)^2}{2} + \frac{2ab+b^3}{6}(\Delta t)^3 \tag{3.C.27}$$

The discrete input noise vector is defined as (see Appendix B, Eq. (9.16))

$$\underline{w}_k = \int_0^{\Delta t} \underset{\sim}{H}(\Delta t,\tau)\,\underset{\sim}{G}(\tau)\,\underline{w}(\tau)\,d\tau \tag{3.C.28}$$

Since $\underset{\sim}{H}$ = 5 x 5 matrix

$\underset{\sim}{G}$ = 5 x 1 matrix

$\underline{w}(t)$ = 1 x 1 vector

Then $\underline{w}_k$ is a vector of length 5. The covariance matrix of $\underline{w}_k$ is given by (Gelb, 1974, p. 121)

$$\underline{Q}_k = \underline{G}\,\underline{Q}\,\underline{G}^T\,\Delta t \tag{3.C.29}$$

where $\underline{Q} = E[w(t)w^T(t)] = (2\beta\,\varphi_0^2)$ (3.C.30)

$$\underline{G}^T = (0\ 1\ \alpha - 2\beta\ 0\ 0) \tag{3.C.31}$$

Substituting $\underline{G}$, G^T, and $\underline{Q}$ into Eq. (3.C.29) yields

$$\underline{Q}_k = \begin{bmatrix} 0 \\ 1 \\ \alpha - 2\beta \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & \alpha - 2\beta & 0 & 0 \end{bmatrix} (2\beta\varphi_0^2)\Delta t$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \alpha - 2\beta & 0 & 0 \\ 0 & \alpha - 2\beta & (\alpha - 2\beta)^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} q\Delta t$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & q\Delta t & (\alpha - 2\beta)q\Delta t & 0 & 0 \\ 0 & (\alpha - 2\beta)q\Delta t & (\alpha - 2\beta)^2 q\Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (3.C.32)$$

where

$$q = 2\beta\,\varphi_0^2 \qquad (3.C.33)$$

Now we can write Eq. (3.C.11) and (3.C.16) in discrete form as

$$\underline{X}_k = \underset{\sim}{H}\,\underline{X}_{k-1} + \underline{w}_k \qquad (3.C.34)$$

$$\underline{Z}_k = \underset{\sim}{S}\,\underline{X}_k + \underline{n}_k \qquad (3.C.35)$$

This X_k should not be confused with the signal components $X_i(t)$ as defined by the equations preceding Eq. (3.C.10). We want to estimate $\underline{X}_k$ since it has the ionospheric phase noise, the clock phase noise and the source visibility phase as its elements.

Using Kalman filter equations given in Appendix B, starting with Eq. (9.24) the vector $\underline{X}_k$ is estimated recursively at t_k. As an output of the filtering process we get the estimates of $\phi_{\text{aux }1}(t_k)$ and $\phi_{\text{aux }2}(t_k)$ along with the estimates of $\phi_{\text{source}}(t_k)$, $\phi_{\text{clock}}(t_k)$ and $\phi_{\text{iono}}(t_k)$. We really do not need $\phi_{\text{aux }1}$ and $\phi_{\text{aux }2}$, but as we pointed out earlier that the initial condition for $\phi_{\text{aux }2}(t_k)$ determines the slope of linear phase drift.

In using Kalman filter equations, we must specify the initial value of the vector process $\underline{X}_k$. If we look at the definition of the a priori error covariance matrix for the first step

$$\underset{\sim}{P}_1(-) = E[(\hat{\underline{X}}_1(-) - \underline{X}_1)(\hat{\underline{X}}_1(-) - \underline{X}_1)^T] \,, \tag{3.C.36}$$

where $\underline{X}_1$ is the true state vector at time t_1, and $\hat{\underline{X}}_1(-)$ is a priori estimate of the state vector $\underline{X}_1$. Let $\tilde{\underline{X}}_1$ be the error vector defined by

$$\tilde{\underline{X}}_1 = \hat{\underline{X}}_1(-) - \underline{X}_1 = \begin{bmatrix} \hat{X}_1(-) - X_1 \\ \hat{X}_2(-) - X_2 \\ \hat{X}(-) - X_3 \\ \hat{X}_4(-) - X_4 \\ \hat{X}_5(-) - X_5 \end{bmatrix} = \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \\ \tilde{X}_3 \\ \tilde{X}_4 \\ \tilde{X}_5 \end{bmatrix} \tag{3.C.37}$$

We substitute Eq. (3.C.37) into Eq. (3.C.36) and get

$$\underset{\sim}{P}_1(-) = E\left\{ \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \\ \tilde{X}_3 \\ \tilde{X}_4 \\ \tilde{X}_5 \end{bmatrix} \begin{bmatrix} \tilde{X}_1 & \tilde{X}_2 & \tilde{X}_3 & \tilde{X}_4 & \tilde{X}_5 \end{bmatrix} \right\}$$

$$= E\left\{ \begin{bmatrix} \tilde{X}_1\tilde{X}_1 & \tilde{X}_1\tilde{X}_2 & \tilde{X}_1\tilde{X}_3 & \tilde{X}_1\tilde{X}_4 & \tilde{X}_1\tilde{X}_5 \\ \tilde{X}_2\tilde{X}_1 & \tilde{X}_2\tilde{X}_2 & \tilde{X}_2\tilde{X}_3 & \tilde{X}_2\tilde{X}_4 & \tilde{X}_2\tilde{X}_5 \\ \tilde{X}_3\tilde{X}_1 & \tilde{X}_3\tilde{X}_2 & \tilde{X}_3\tilde{X}_3 & \tilde{X}_3\tilde{X}_4 & \tilde{X}_3\tilde{X}_5 \\ \tilde{X}_4\tilde{X}_1 & \tilde{X}_4\tilde{X}_2 & \tilde{X}_4\tilde{X}_3 & \tilde{X}_4\tilde{X}_4 & \tilde{X}_4\tilde{X}_5 \\ \tilde{X}_5\tilde{X}_1 & \tilde{X}_5\tilde{X}_2 & \tilde{X}_5\tilde{X}_3 & \tilde{X}_5\tilde{X}_4 & \tilde{X}_5\tilde{X}_5 \end{bmatrix} \right\}$$

$$= \begin{bmatrix} e_1 & & & & \\ & e_2 & & \underset{\sim}{0} & \\ & & e_3 & & \\ & & & e_4 & \\ & \underset{\sim}{0} & & & e_5 \end{bmatrix} \qquad (3.C.38)$$

where

$$e_i = E[\tilde{X}_i\tilde{X}_i] , \qquad i = 1, 2, 3, 4, 5.$$

Off-diagonal elements are zero since the elements of the state vector $\underline{X}_1$ are independent. For $k \neq 1$ this is not true, however. We will not discuss this here since it is not the purpose of this chapter.

It is easily seen from Eq. (3.C.36) that the diagonal elements of $\underset{\sim}{P}_1(-)$ are the initial mean square errors of the state variables. According to the degree of our knowledge about $\underline{X}_1(-)$ we specify the $\underset{\sim}{P}_1(-)$ matrix. This is done once, later the filter algorithm automatically updates the covariance matrix, thus providing the mean square error for each element of the state vector at every step. As the filter gets more data we expect the diagonal elements of the covariance matrix to become smaller.

If there is no a priori information about the source visibility phase angle, the author suggest taking the initial mean square error and the initial value for $X_1(t_k) = \phi_{source}(t_k)$ to be zero. Then, the filter assumes that the value of $\phi_{soutce}(t_k)$ is perfectly known to be zero and does not attempt to change the initial value of this variable. In this case we find the estimates only for $X_2(t_k) = \phi_{ion}(t_k)$ and $X_4(t_k) = \phi_{clock}(t_k)$ at each time t_k, $k = 1, 2, \ldots, N$. To remove $\phi_{ion}(t_k)$ and $\phi_{clock}(t_k)$ from the original data sequence $\phi(t_k)$ we form

$$\phi'(t_k) = \phi(t_k) - [\hat{\phi}_{ion}(t_k) + \hat{\phi}_{clock}(t_k)] \tag{3.C.39}$$

$\phi'(t_k)$ is the new phase data to be used in the coherent or incoherent signal detection methods described in Chapter II. To see how the phase fluctuation decreases we substitute Eq. (3.B.2) in Eq. (3.C.39) and obtain

$$\begin{aligned}\phi'(t_k) &= \phi_{iono}(t_k) + \phi_{clock}(t_k) - \hat{\phi}_{iono}(t_k) - \hat{\phi}_{clock}(t_k) + \phi_{stat}(t_k) \\ &= [\phi_{iono}(t_k) - \hat{\phi}_{iono}(t_k)] + [\phi_{clock}(t_k) - \hat{\phi}_{clock}(t_k)] + \phi_{stat}(t_k) \\ &= \tilde{X}_2(t_k) + \tilde{X}_4(t_k) + \phi_{stat}(t_k)\end{aligned} \tag{3.C.40}$$

Since $\tilde{X}_2(t_k)$, $\tilde{X}_4(t_k)$ and $\phi_{stat}(t_k)$ are independent, the variance of $\phi'(t_k)$ is

$$\sigma^2_{\phi'(t_k)} = \mathrm{Var}[\tilde{X}_2(t_k)] + \mathrm{Var}[\tilde{X}_4(t_k)] + \sigma^2_{\phi_{stat}(t_k)} \qquad (3.C.41)$$

From Eq. (3.B.2) the variance of $\phi(t_k)$ is

$$\sigma^2_{\phi(t_k)} = \mathrm{Var}[\phi_{ion}(t_k)] + \mathrm{Var}[\phi_{clock}] + \sigma^2_{\phi_{stat}(t_k)} \qquad (3.C.42)$$

$$= \mathrm{Var}[X_2(t_k)] + \mathrm{Var}[X_4(t_k)] + \sigma^2_{\phi_{stat}(t_k)} \qquad (3.C.42)$$

Comparing Eq. (3.C.41) and Eq. (3.C.42) we obtain

$$\mathrm{Var}[\phi'(t_k)] < \mathrm{Var}[\phi(t_k)], \qquad k = 1, 2, \ldots, N .$$

If we take the initial estimates of ϕ_{ion} and ϕ_{clock} as zero, the mean square errors made in the initial estimates are the variances of the variables ϕ_{ion} and ϕ_{clock}, respectively. For a stable Kalman filter, error covariances always decrease as the filter operates on each successive measurement. The decrease of the variance of the residual fringe phase allows one to average complex correlation functions for a longer time, thus increasing the sensitivity of the interferometer (Chapter II).

D. Implementation of the Methods

In the last three sections the bias recovery and the phase noise removal techniques were described. Both of these techniques are quite sensitive to model errors. To do modeling accurately, the phase data must be sampled fast enough so that the samples represent the phase data adequately.

If the NRAO Mark I system is used, the phase sampling period must be taken as 0.2 sec (one record length).

To implement the techniques one should follow the following steps.

(a) Remove the natural fringe rate and the frequency offset from the VLBI data. This can be done by using standard VLBI data reduction programs.

(b) Analyze the fringe phase data to obtain the ionospheric and clock phase models. Since the natural fringe phase has already been taken out and all the corrections for the baseline errors have been made, the phase data includes only the differential ionospheric phase and the differential clock phase noises. Before

Before applying any process to the phase data the "2π" ambiguities have to be removed. Use a standard regression program to remove the trend from the phase data from which the "2π" ambiguities have been removed. The slope of the trend gives an estimate for the parameter c defined by Eq. (3.B.10). After removing the trend, look for the random periodic component in the residual phase data. Compute the covariance function of the data. If the periodic component exists, the autocovariance function will be periodic.

Simple curve fitting to the autocovariance function yields the parameters φ_0^2, β, and ω defined in Section C of this chapter.

If there is no evidence that a periodic component exists in the phase data, set the initial values of the variables $\phi_{iono}(t)$ and $\phi_{aux\_1}(t)$, and the initial values of the covariances asso-

ciated with these variables to zero. This procedure simply tells the Kalman filter that these variables are perfectly known, therefore it must not operate on these two variables. During the filtering these two variables keep their initial values, i.e., zero. Thus, the filter estimates only the clock phase noise component.

To compute the variance of the statistical phase noise component remove the periodic component from the phase data from which the linear trend is removed. The variance of the residual phase data is an estimate of the statistical noise.

The initial values of the noise components may be obtained from the regression analysis outlined above.

The diagonal elements of the initial error covariance matrix are the mean square values of the uncertainties in the initial values of the state variables. For the initial covariance matrix, the off-diagonal elements are taken to be zero.

(c) In the previous step all the necessary parameters were estimated, and the initial error covariance matrix and the initial values of the state variables were determined.

Now, use the algorithm given in Section C of this chapter to obtain the fringe phase from which the ionospheric and clock noises are removed. Using this new phase data apply the bias recovery technique given in the beginning of this chapter to calculate the new correlation between the two interferometer signals.

IV. RESULTS AND DISCUSSION

A. Data Acquisition and Analysis

In order to implement the proposed method for reducing the effects of the ionosphere and the atomic frequency standards the data taken during VLBI experiments between Ames, Iowa and Boulder, Colorado in 1971 August were used.

Because the ionosphere causes severe problems in decametric wavelength VLBI, the data were so noisy that very meaningful conclusions could not be drawn from the data. However, the data were suitable for testing the proposed technique.

Prior to the use of the phase noise removal method proposed in this thesis the data were preprocessed by using standard VLBI data reduction programs. The details of these programs were given by various authors (Whitney et al., 1976; Whitney, 1974; Clark, 1973; Clark et al., 1968).

In the data gathering the antennas used for the experiments were arrays of 160 pairs of crossed full-wave dipoles, operating at 26.3 MHz connected for left circular polarization. Each array had an effective aperture of 4,000 m^2. The NRAO Mark I recording system was used for recording the data on the magnetic tapes.

A standard data reduction process on tapes was made at Goddard Space Flight Center. The NRAO Mark I system has a maximum bandwidth of 360 kHz. The signals from IF outputs at each VLBI terminal were clipped, sampled at 720 kbits/sec and recorded on the magnetic tapes. From the sampling rate we see that the bandwidth used was 360 kHz. A 1-bit recording scheme was used to process the data more effectively.

In the previous chapters, we stated that in this system the data were recorded on magnetic tapes in blocked form with no timing information except the starting time of the recording. Each data record was 0.2 sec long.

To obtain the fringes, the signals recorded at each site were cross-correlated on a record-by-record basis. One of the bit streams was shifted according to the best known baseline, source parameters, and time offset introduced during the experiment. Then two records, each from one VLBI site, were multiplied together and accumulated for 0.4 msec. The multiplication process was done quite easily since the 1-bit multiplication is the same as the "exclusive or" process in the computer. The numbers obtained every 0.4 msec, were then multiplied by the sine and cosine of the computed fringe phase and further accumulated for 0.2 sec. This process was equivalent to shifting the fringe function in frequency by the expected fringe rate. The result of the process described above was a complex number with real and imaginary parts. By the use of six time shifts about the best predicted shift six additional cross-correlations were obtained for each 0.2 sec record. The analysis of this cross-correlation function in the frequency domain was made to check the predicted fringe rate and time offset were correct. If they were not correct, different fringe rate and time offset values were tried. This process was repeated for each record. Since a typical observation period was 160 sec long, a total of 800 records were analyzed for each observation.

After time offset and fringe rate errors were removed, the complex cross-correlation value with the maximum amplitude was selected. For 800

records 800 complex values of the cross-correlation function were obtained.

If there were no ionospheric and clock phase noises, the phase of each complex number would be a random phase with the probability density function given by Eq. (2.D.17). These random phases would be due to the receiver, sky-background, and sampling noises. The probability density function of the amplitudes is given by Eq. (2.D.16).

For a point source the expected value of the phase is known to be zero (Moran, 1973). But, in the presence of the ionospheric and clock phase noises, the phase of each complex number obtained every 0.2 sec has three components: ionospheric component, clock component and statistical component.

During the data processing at Goddard Space Flight Center the complex correlations obtained every 0.2 sec were averaged coherently for ten records. Therefore, the data we had for this analysis consisted of 80 pairs of real and imaginary numbers.

We must recognize that these complex numbers obtained by preprocessing do not represent the complex fringe amplitudes, because after the 1-bit correlation process we only obtain the normalized correlation. The true correlation function is recovered by multiplying the normalized correlation by a constant number determined by the receiver and sky-background noises.

This coherent averaging limits our analysis of the phase data since the sampling interval is 2 seconds. Thus, we could not see ionospheric phase variations having time scales less than 2 seconds. However, the

variance of the statistical noise can still be calculated by using the variance formula given by Eq. (2.D.21) in Chapter II.

The optimum estimator of the fringe amplitude is the coherent average, provided that the phase noise is at an acceptable level.

The phase data we had could not be used directly because of "2π" ambiguities. The corrected phase data were derived from the raw phase data by the procedure which attempts to remove these ambiguities. Sudden changes in the electron density can cause differential phase to jump by many radians. Since a typical interferometer measures the principal value of the phase and loses any $\pm 2\pi$ increments, the jumps in the phase data are seen as discontinuities.

The discontinuities are detected and corrected by adding or subtracting an integer multiple of 2π from the data at the point of discontinuity. The procedure is explained below.

Let $\{\phi_k, k = 1, 2, \ldots, n\}$ be a set of phase angles as given by the raw phase data, and let $\{\psi_k, k = 1, 2, \ldots, n\}$ be a set of phase angles corrected for "2π" ambiguities. Where n is the number of phase samples. To start the correction procedure, set

$$\psi_1 = \phi_1$$

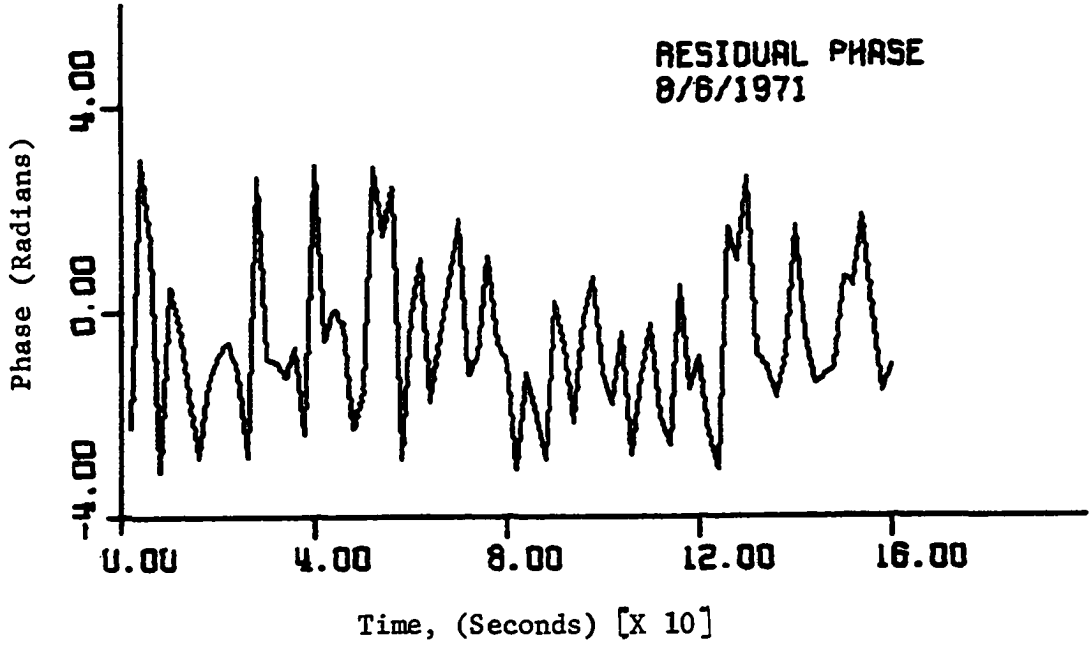
As a second step, choose the angle in the set $\{\phi_{k+1} + i\,2\pi, i = 0, 1, 2, \ldots\}$ that differs least from the angle ψ_i. The algorithm described above is recursive and quite simple. A FORTRAN IV implementation of the algorithm is given at the end of this thesis. The algorithm was tested as follows: A set of 80 random numbers, uniformly distributed between $-\pi$ and $+\pi$ were

generated. The algorithm was applied to these numbers, and there was no change in the numbers after the correction process. Several other independent sets of random numbers were used for testing the algorithm. The results were the same as the first test. Thus, we concluded that the phase correction algorithm was not changing the character of the phase data unless the data had special features, e.g., periodic behavior or a linear trend.

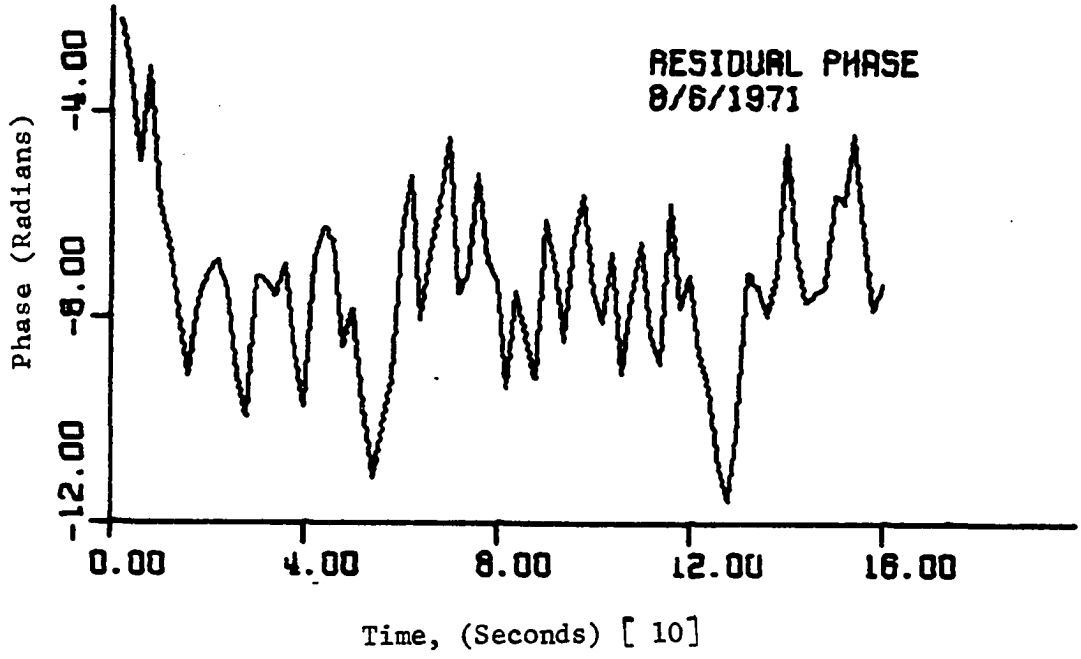
Firuges 4.a.1 - 4.a.10 show the raw and corrected phase data for the radio source 3C 144 (Crab nebula). These 10 data records were recorded on August 6, 7, 10, 19, 20, 21, 22, and 23 in 1971 with a Boulder-Ames baseline. The top panel in each figure shows the data before removing the $2i\pi$ ambiguities (called the raw data) and the lower panel shows the result after removing the $2i\pi$ ambiguities. There is no unique way to resolve the phase ambiguity. The procedure shown here is the simplest technique which minimizes phase discontinuities.

From the corrected phase data we notice that the residual fringe phase has mainly three types of variations. These are: 1) the linear drift, which is attributed to the differential local oscillator phase drift; 2) the slowly varying component which may be due to the ionosphere; and 3) the rapidly changing white noise component. The latter is due to the sky-background, the receiver noises, and the statistical sampling.

The character of the phase drift changes considerably from one figure to another. The August 6, 1971 record has little overall drift but larger phase changes and a noisier appearance than the other records. The data taken during August 7, 1971 show less drift and ionospheric disturbance

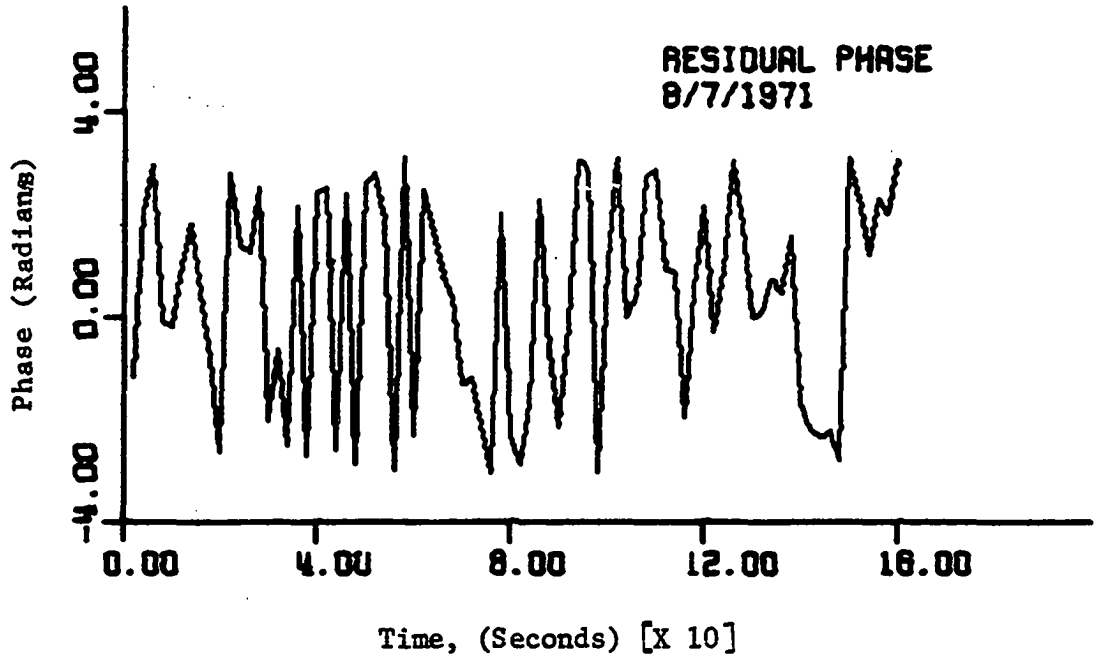


(a) Original phase data before 2 i π ambiguities were removed.

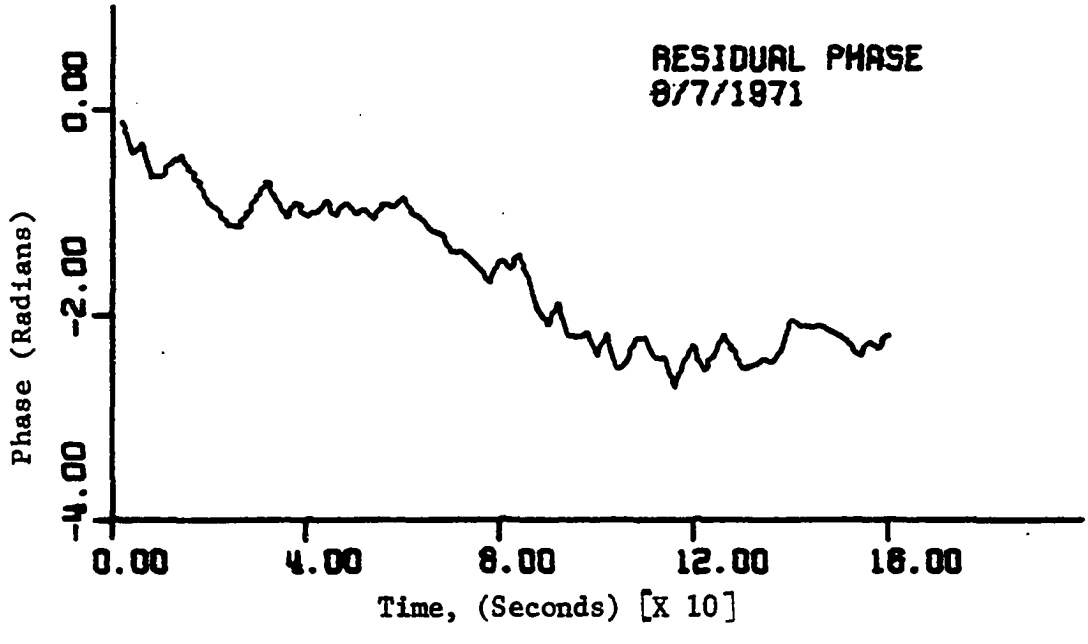


(b) Phase data after 2 i π ambiguities were removed.

Figure 4.a.1. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI and 2 Programs. Data collected on August 6, 1971.

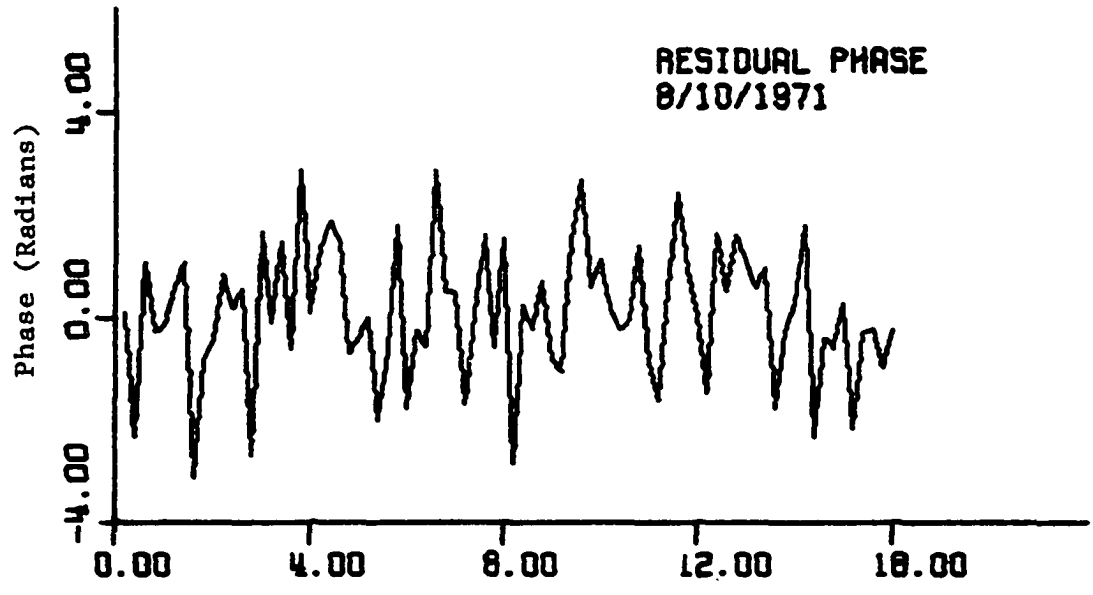


(a) Original phase data before 2 i π ambiguities were removed.

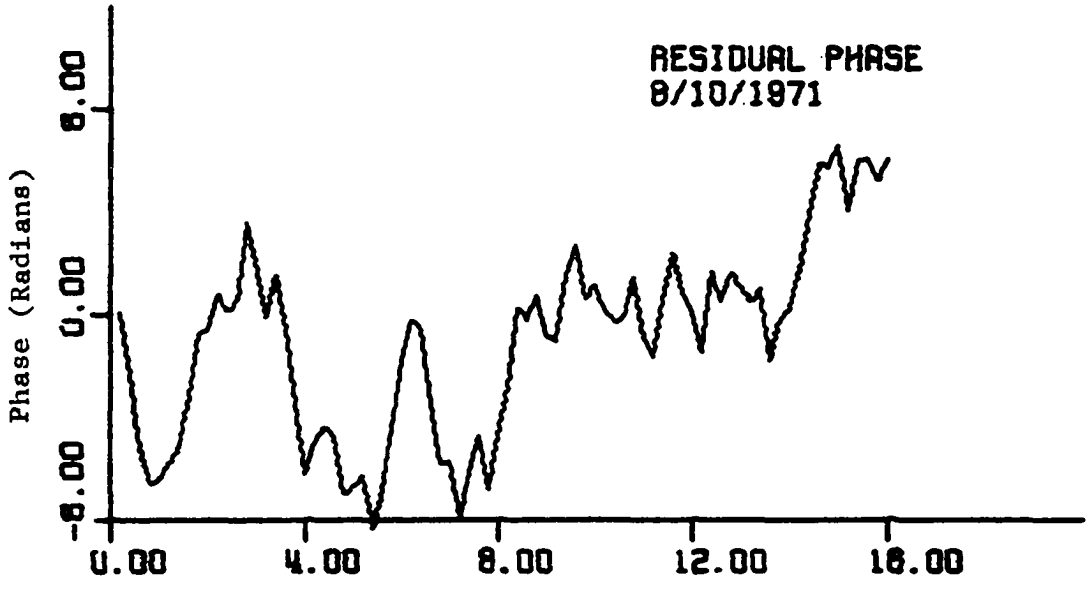


(b) Phase data after 2 i π ambiguities were removed.

Figure 4.a.2. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI and 2 Programs. Data collected on August 7, 1971.

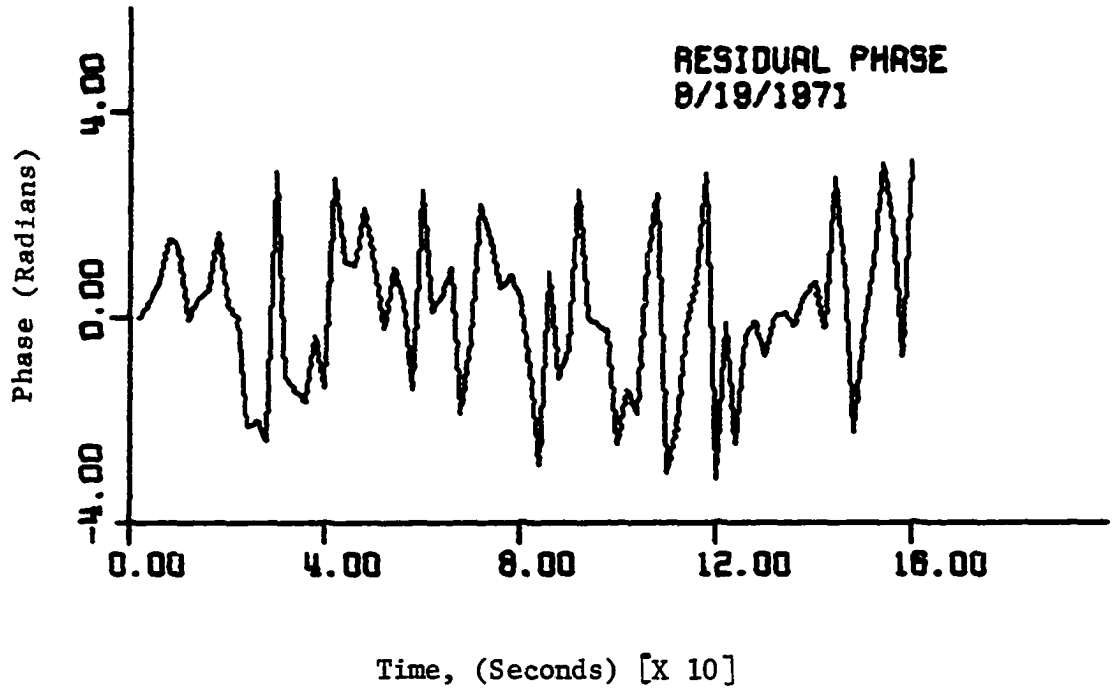


(a) Original phse data before 2 i π ambiguities were removed.

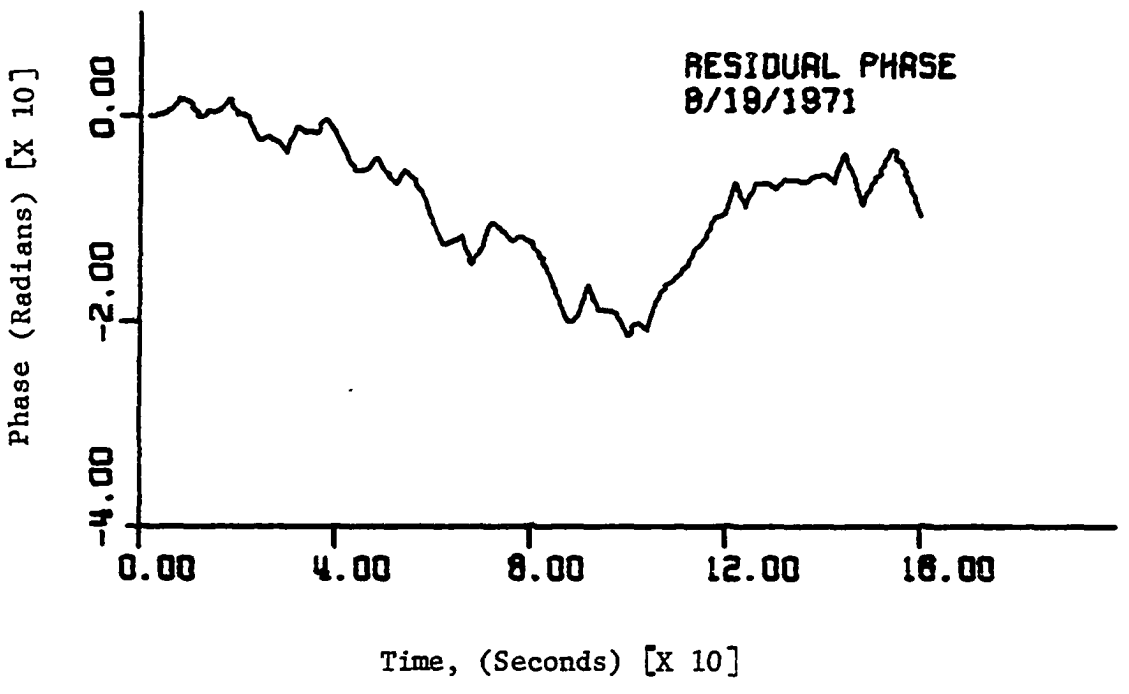


(b) Phase data after 2 i π ambiguities were removed

Figure 4.a.3. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI and 2 Programs. Data collected on August 10, 1971

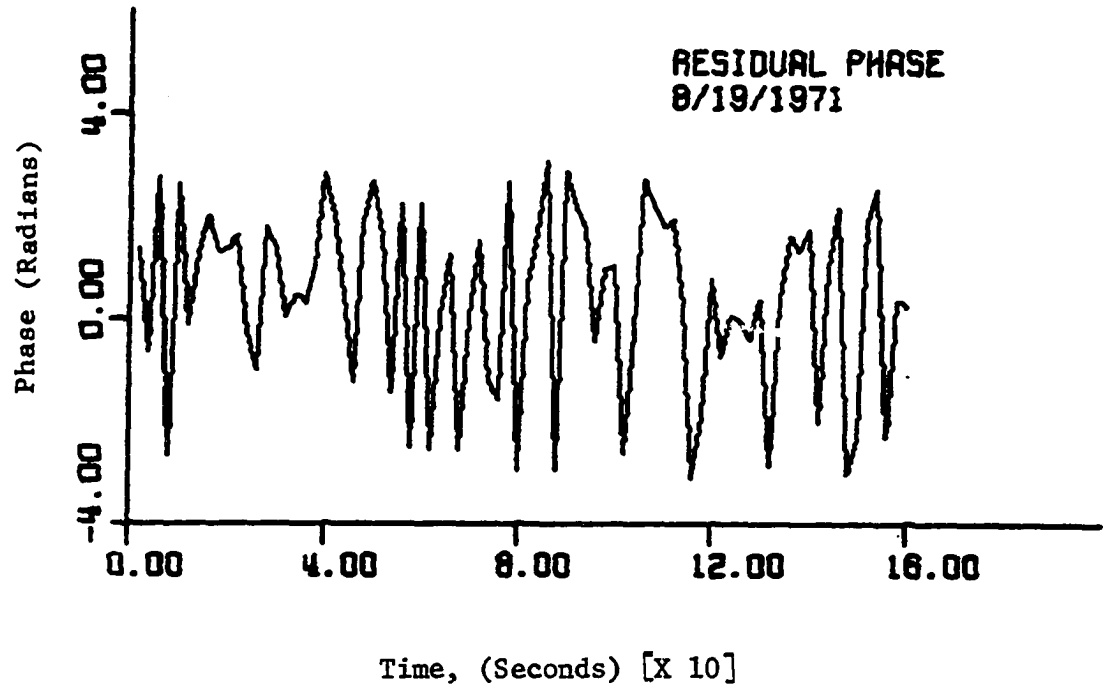


(a) Original phase data before 2 i π ambiguities were removed.

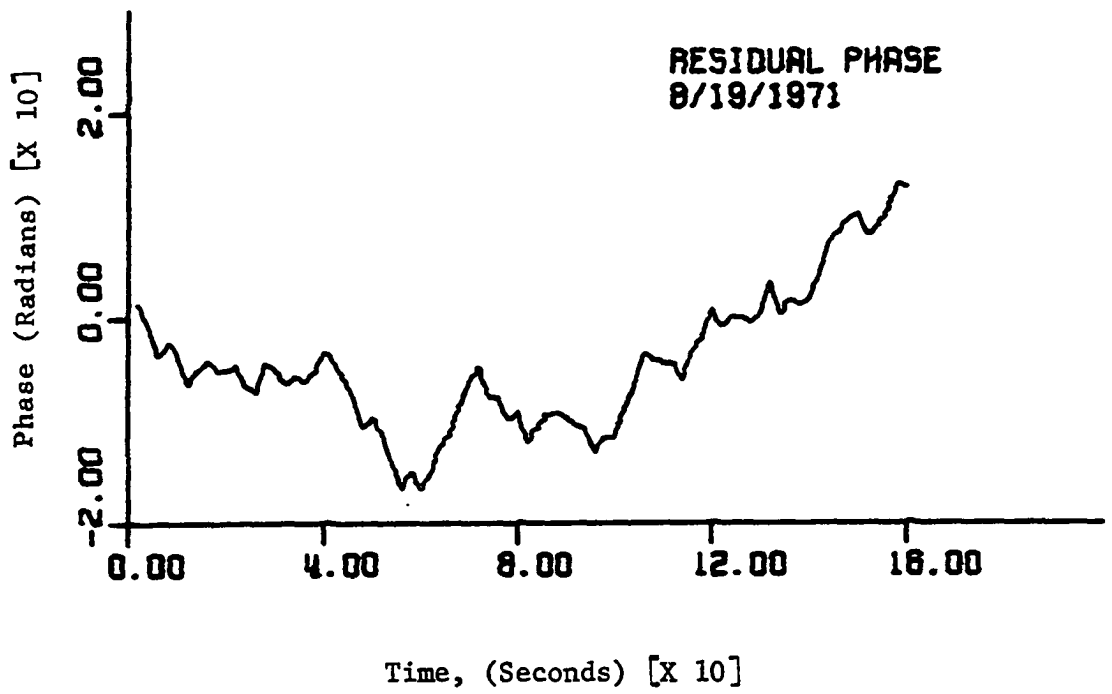


(b) Phase data after 2 i π ambiguities were removed.

Figure 4.a.4. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI 1 and 2 programs. Data collected on August 19, 1971

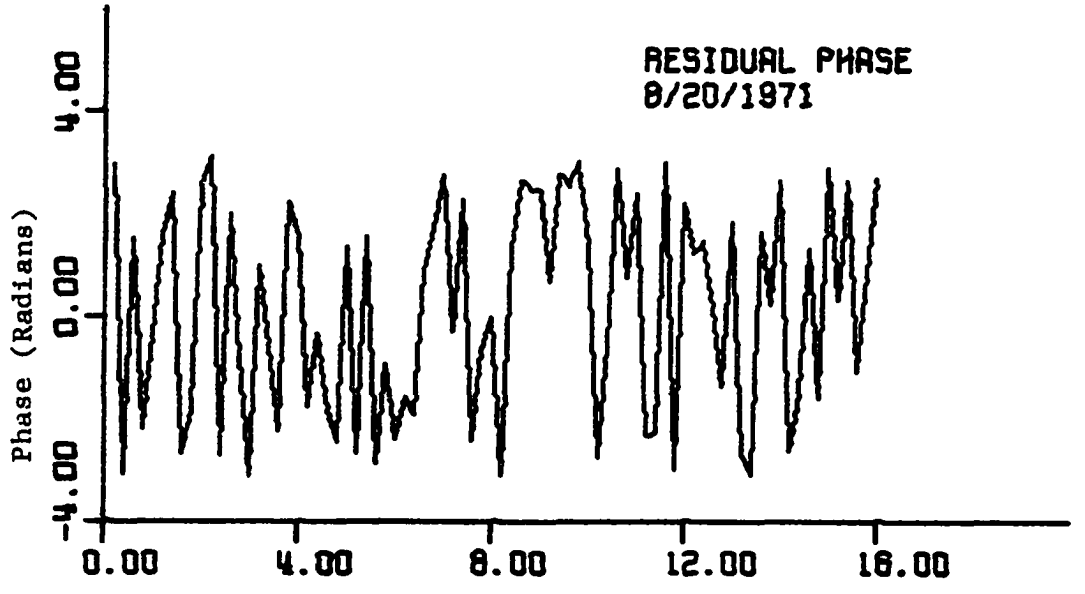


(a) Original phase data before 2 i π ambiguities were removed.

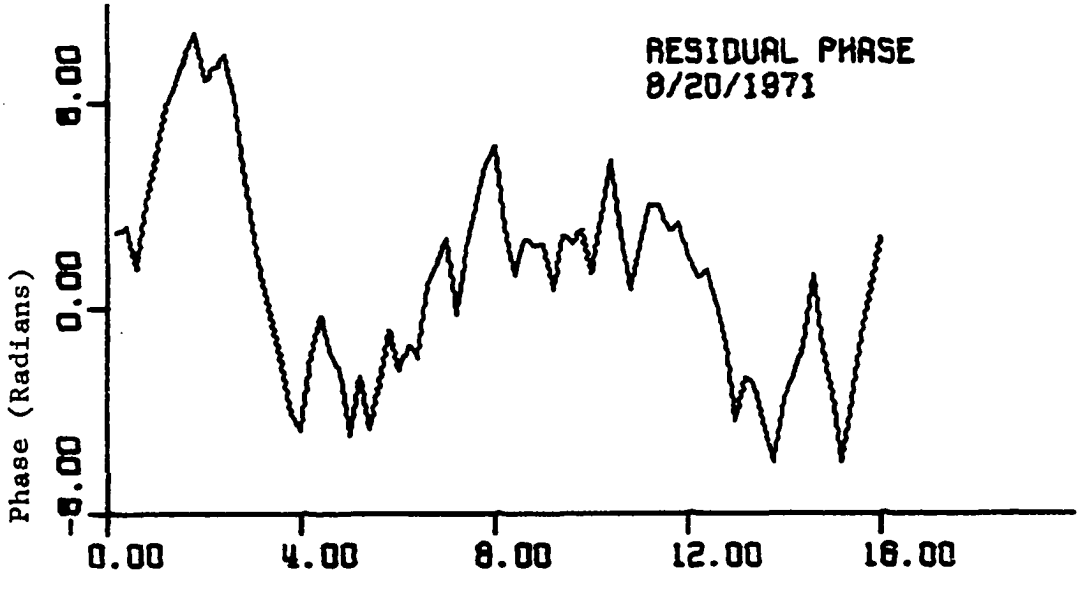


(b) Phase data after 2 i π ambiguities were removed.

Figure 4.a.5. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI 1 and 2 programs. Data collected August 19, 1971

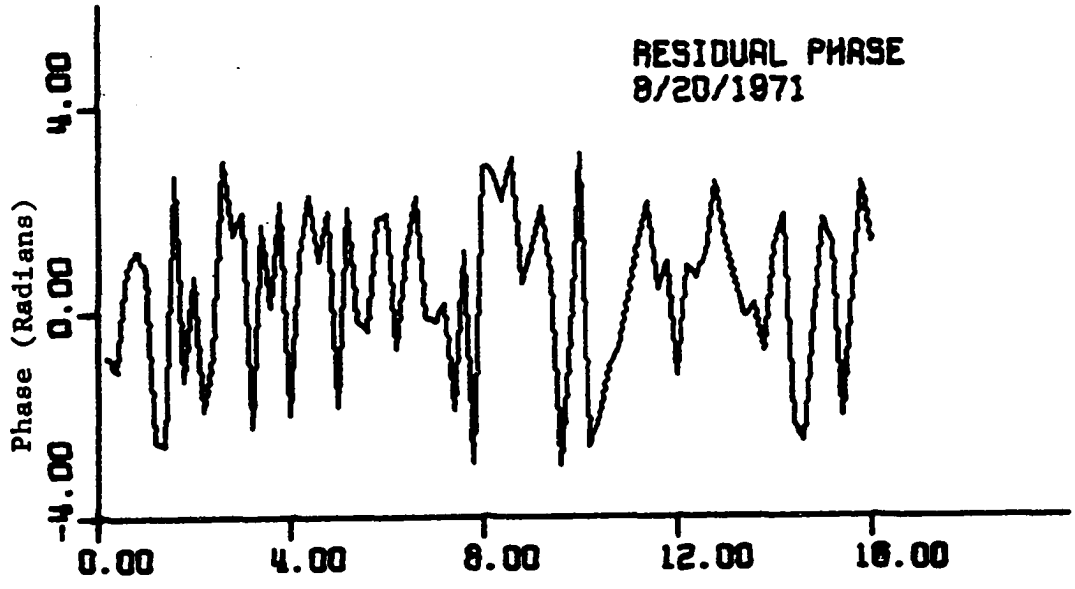


(a) Original phase data before 2 i π ambiguities were removed.

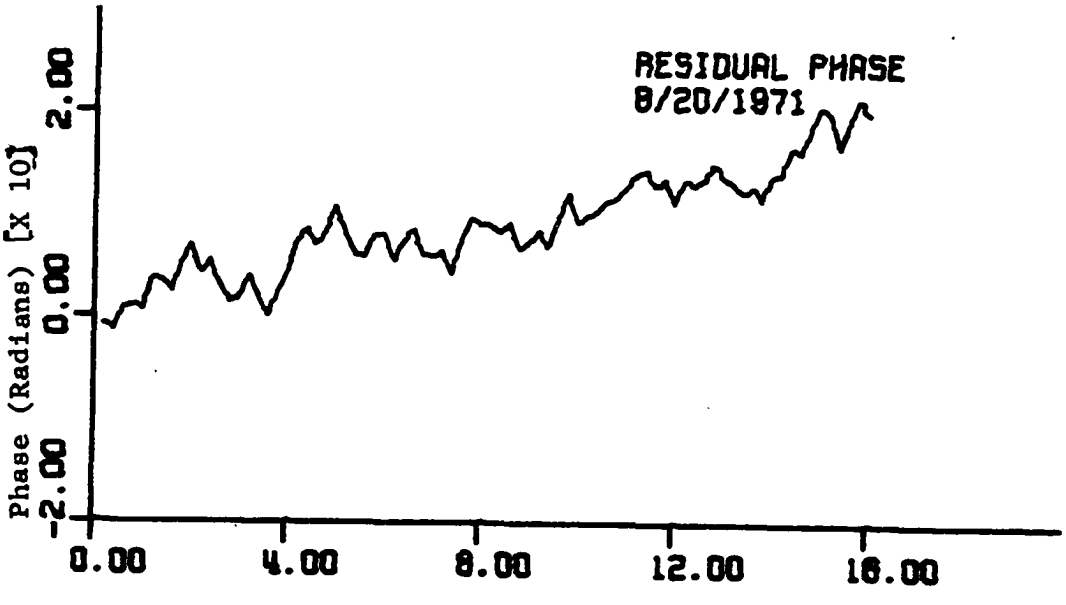


(b) Phase data after 2 i π ambiguities were removed.

Figure 4.a.6. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI 1 and 2 programs. Data collected on August 20, 1971

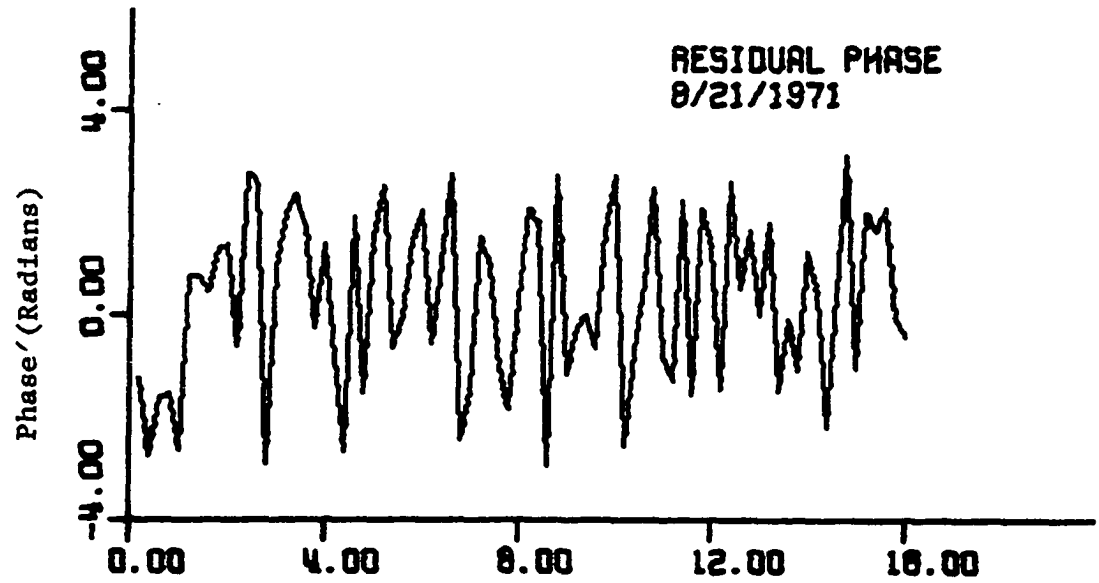


(a) Original phase data before 2 i π ambiguities were removed

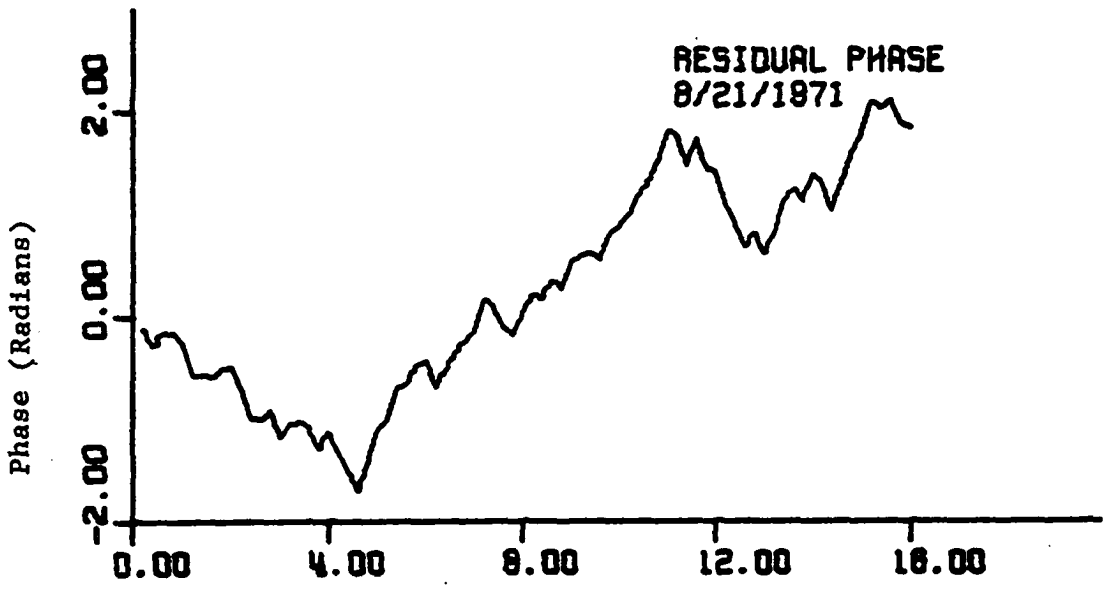


(b) Phase data after 2 i π ambiguities were removed.

Figure 4.a.7. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI 1 and 2 programs. Data collected August 20, 1971.

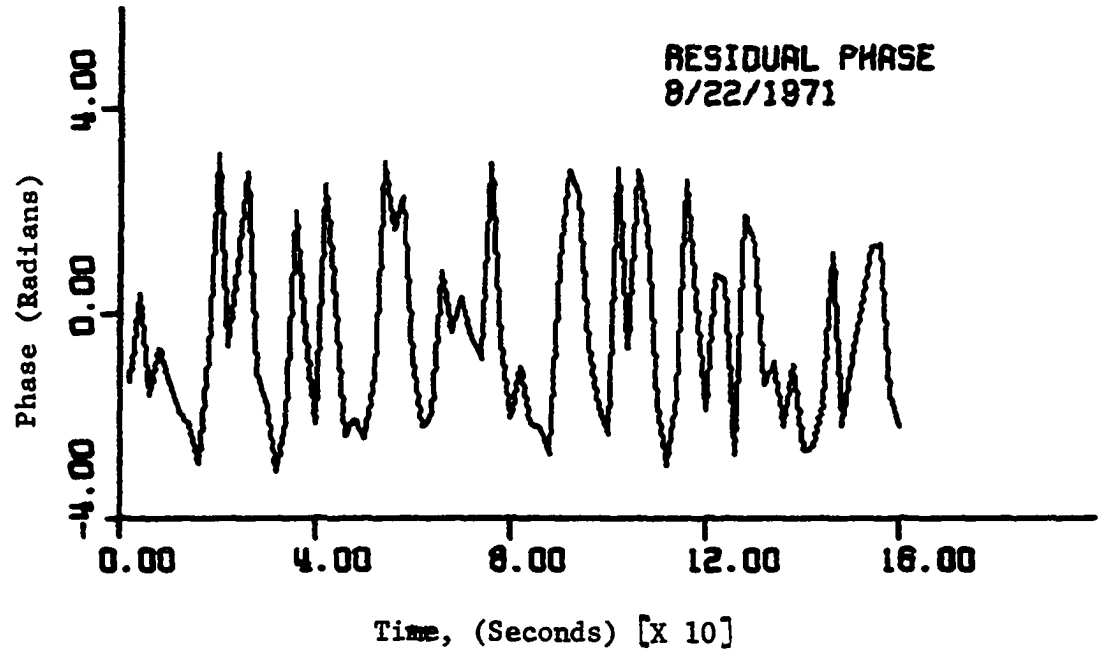


(a) Original phase data before 2 i π ambiguities were removed

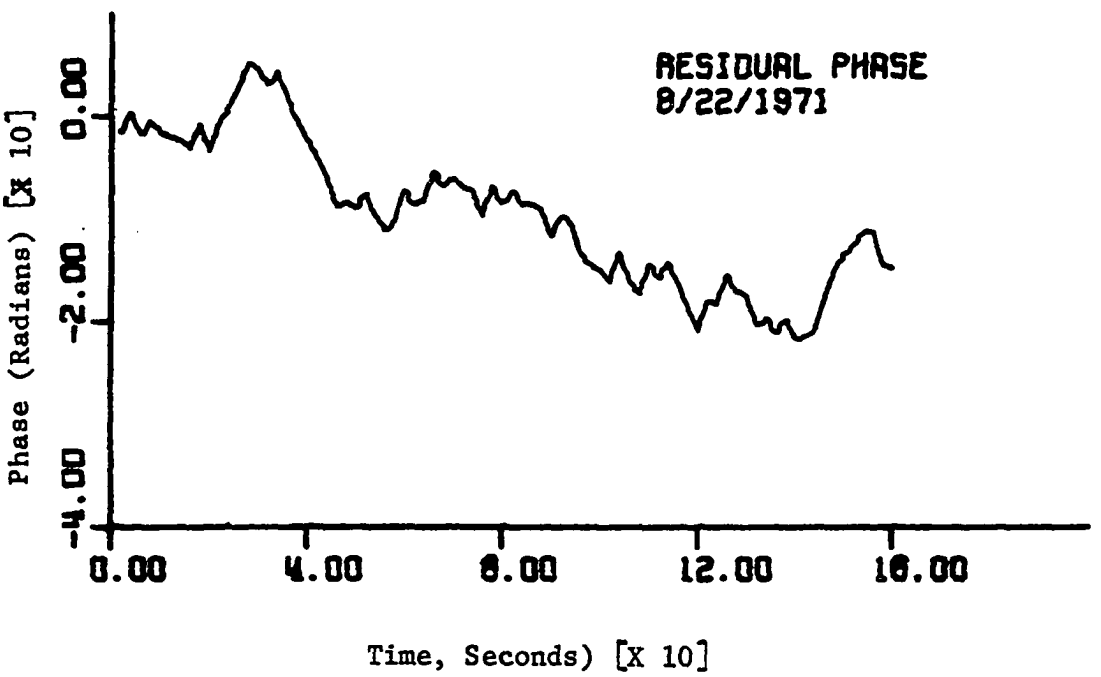


(b) Phase data after 2 i π ambiguities were removed

Figure 4.a.8. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI 1 and 2 programs. Data collected on August 21, 1971

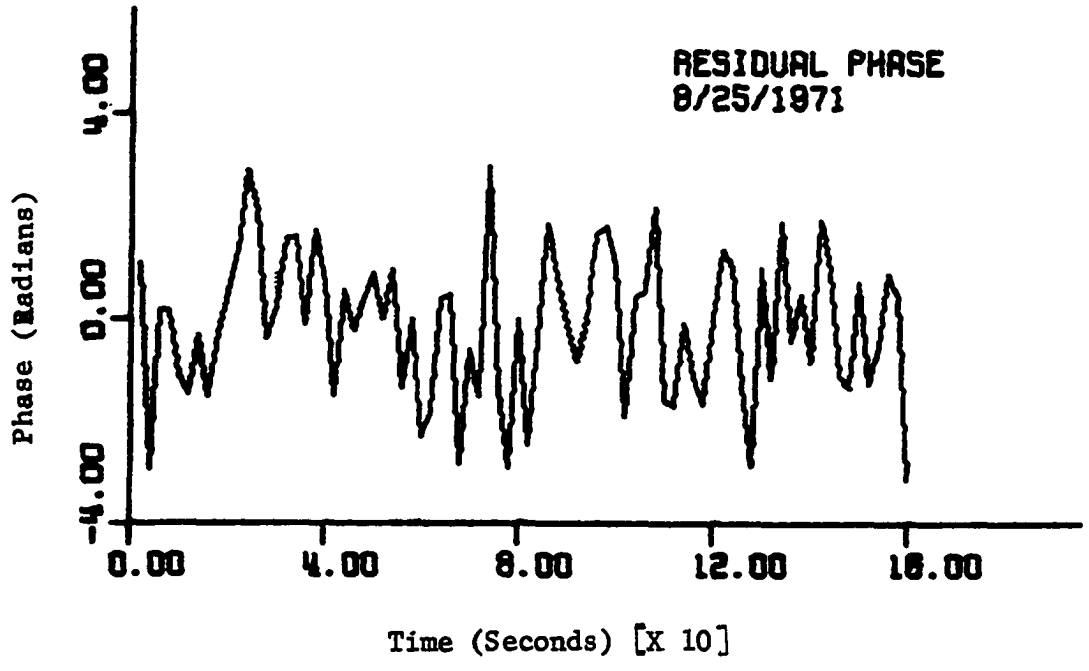


(a) Original phase data before 2 i π ambiguities were removed

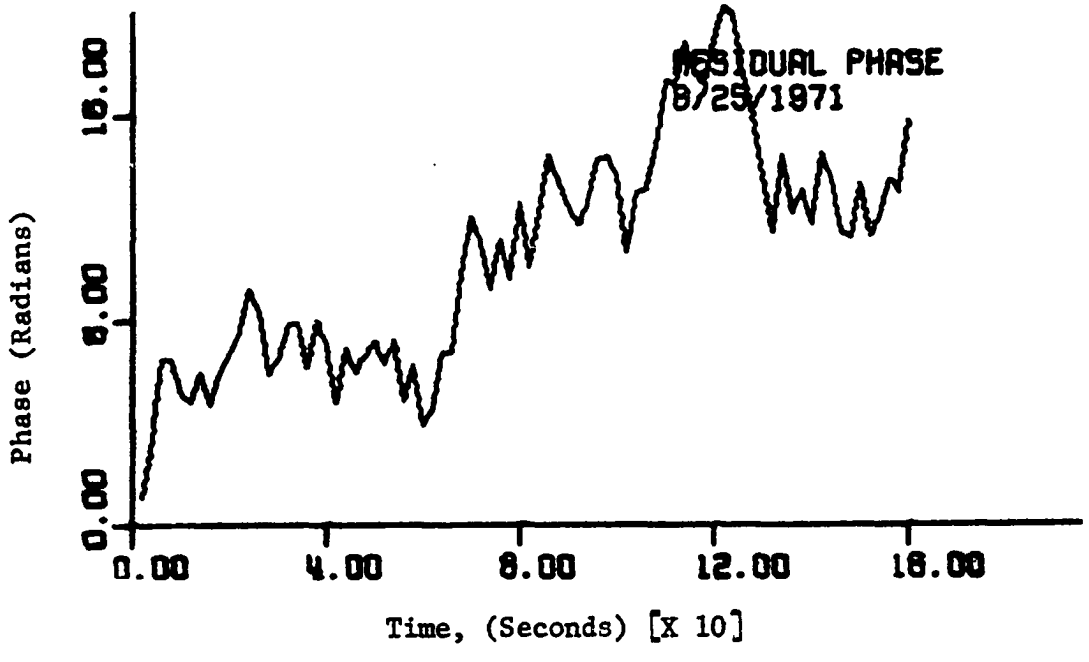


(b) Phase data after 2 i π ambiguities were removed

Figure 4.a.9. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI 1 and 2 programs. Data collected on August 22, 1971.



(a) Original phase data before 2 i π ambiguities were removed



(b) Phase data after 2 i π ambiguities were removed

Figure 4.a.10. Phase data on 3C 144 (Crab Nebula) after processing by the standard VLBI 1 and 2 programs. Data collected on August 25, 1971

than the other days. The linear trend on these records is evident but the noise fluctuations are small. The calculated fringe amplitude for the August 7th observation was much higher than those obtained from the other observations. The fringe amplitude on August 19th was the second highest.

Except for the two days mentioned above, the ionosphere was quite disturbed in August 1971. The fringe amplitudes obtained during these disturbed conditions of the ionosphere were below the 3σ threshold level which is the measure of the minimum detectable fringe amplitude.

To remove this random variation from the phase data we modeled the ionospheric phase noise as a periodic random process described by three parameters that can be estimated through the use of the autocorrelation of the phase data (see Chapter III). These three parameters are, ω, the frequency of the process; β, the parameter in the argument of the exponential function in the covariance function describing the periodic process, and the variance of the process, i.e., the covariance at zero lag.

Once preliminary estimations of these parameters are determined, a search for optimum values can be made by trying different values of the parameters until the model removes the systematic noise in the phase data. This process may take excess computer time; therefore, it is not suggested unless needed and was not applied to our data.

The estimated values of the ionospheric phase noise model can be used for the following observations if we have a reason to believe that the ionospheric conditions are not changing during these experiments.

A nice feature of the method is that once we obtain parameters for the ionospheric noise model, we let the Kalman filter run between the two

adjacent observations. When the second observation starts we can update the model by changing a few parameters in the computer program.

The details of how the ionospheric phase noise model parameters are obtained are given in Chapter III. We discussed the phase noise due to the relative phase drift of the two local oscillators and assumed that this drift was linear with a random slope. The estimation of the drift rate can also be obtained from a regression analysis on the corrected phase data. The estimated value of the drift rate was compared with the one obtained by comparing the atomic frequency standards used in the experiments.

The statistical phase noise is assumed to be white and its variance may be calculated by using the method given in Chapter II. We estimated the variance of this noise from the regression analysis on the corrected phase data.

Having obtained the parameters for the phase noise components we removed these componenets from the corrected phase data by using the Kalman filter algorithm given in Appendix B. The details of the usage of the technique are given in the previous chapter.

The residual phase data after noise removal are shown in Figures 4.a.11 - 4.a.20. Each plot represents phase data from a different 160 sec long observation.

From Figures 4.a.11 - 4.a.20 we see that the linear trends and slowly varying ionospheric phase are removed from the phase data. The filtered phase data for each observation are much like random white noise processes. The data shown in Figures 4.a.13 and 4.a.14 have nonzero means. This either

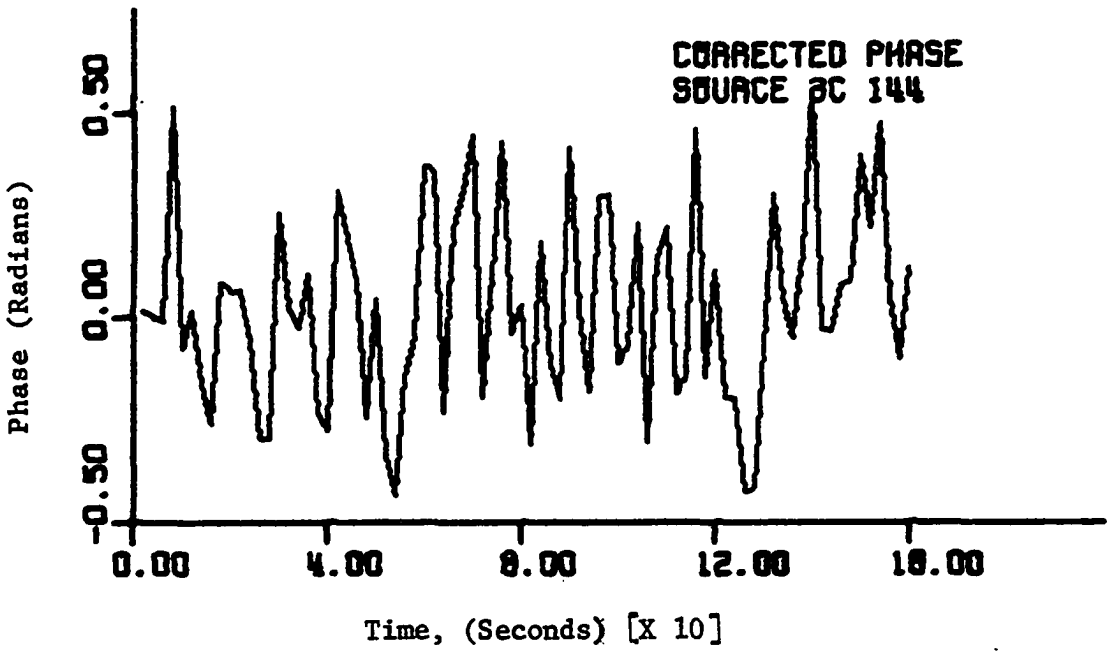


Figure 4.a.11. Phase data corrected for clock drift and ionospheric noise. Data collected on August 6, 1971.

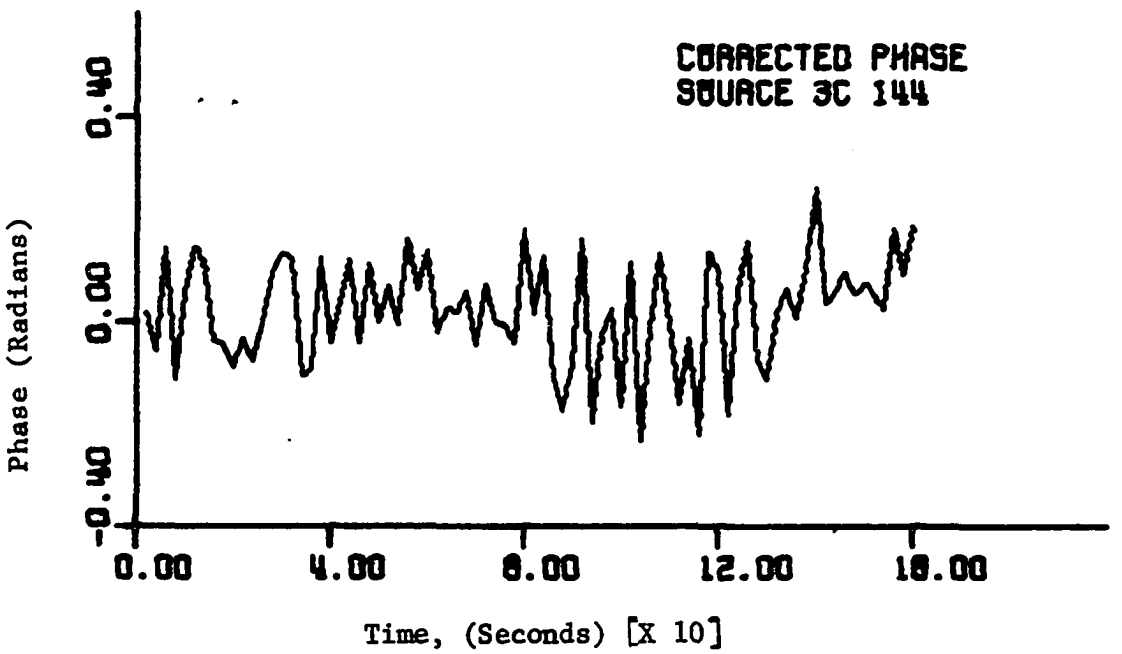


Figure 4.a.12. Phase data corrected for clock drift and ionospheric noise. Data collected on August 7, 1971

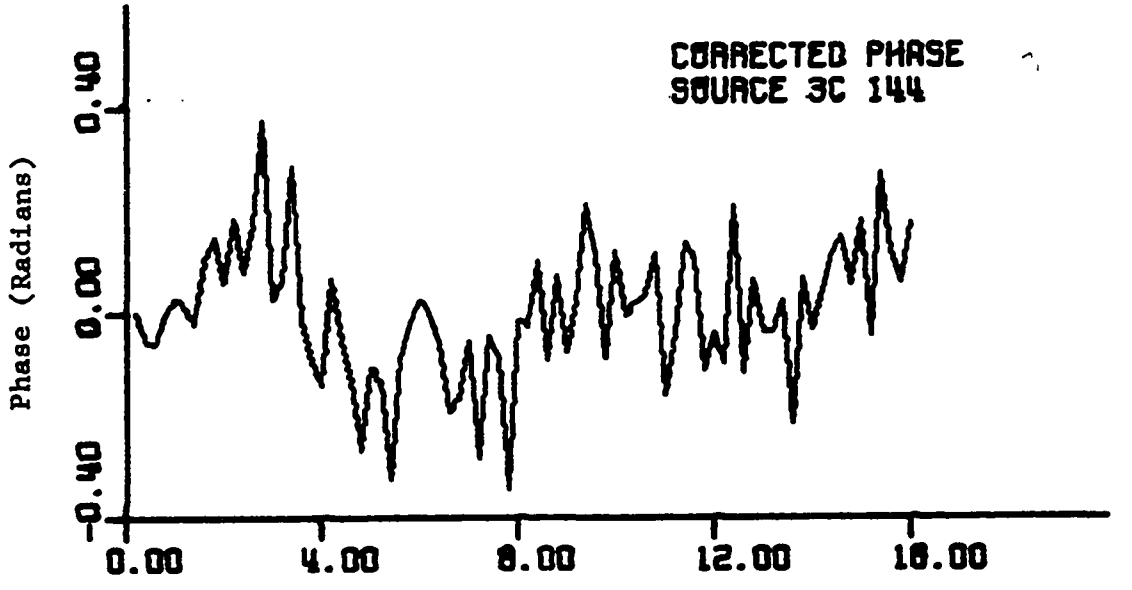


Figure 4.a.13. Phase data corrected for clock drift and ionospheric noise. Data collected on August 19, 1977.

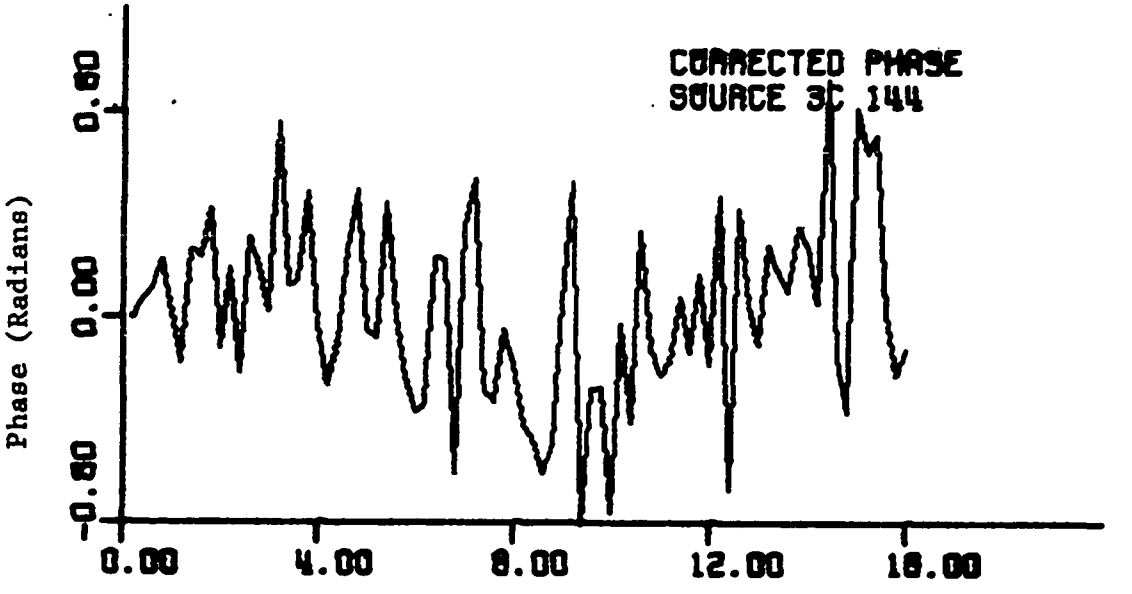


Figure 4.a.14. Phase data corrected for clock drift and ionospheric noise. Data collected on August 20, 1977

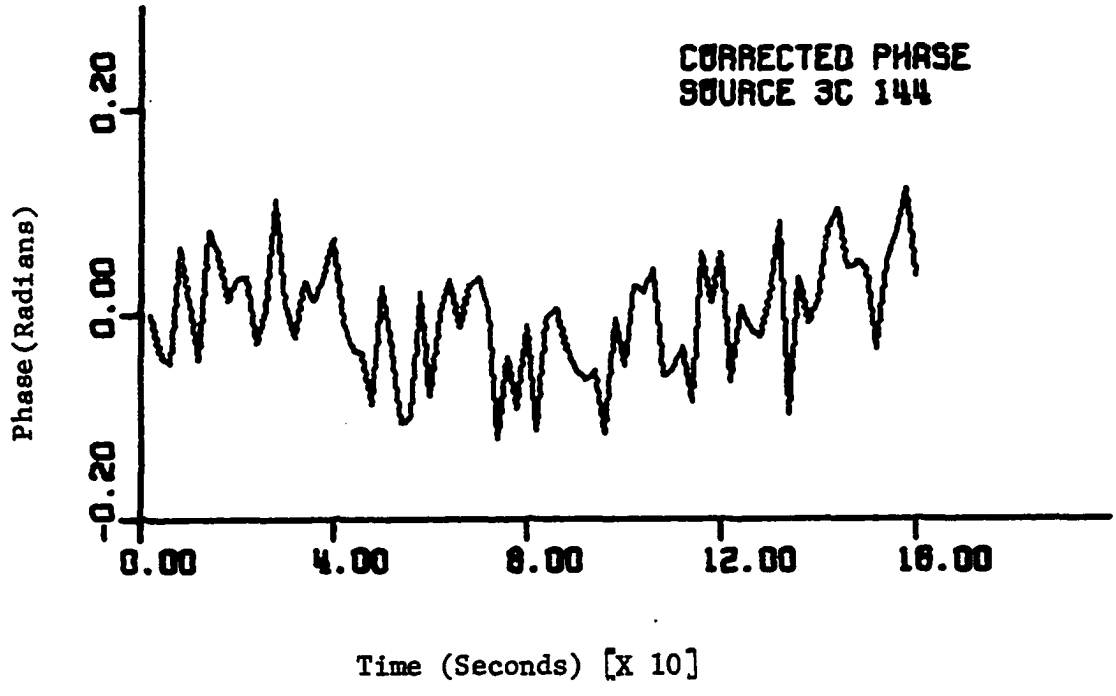


Figure 4.a.15. Phase data corrected for clock drift and ionospheric noise. Data collected on August 20, 1971

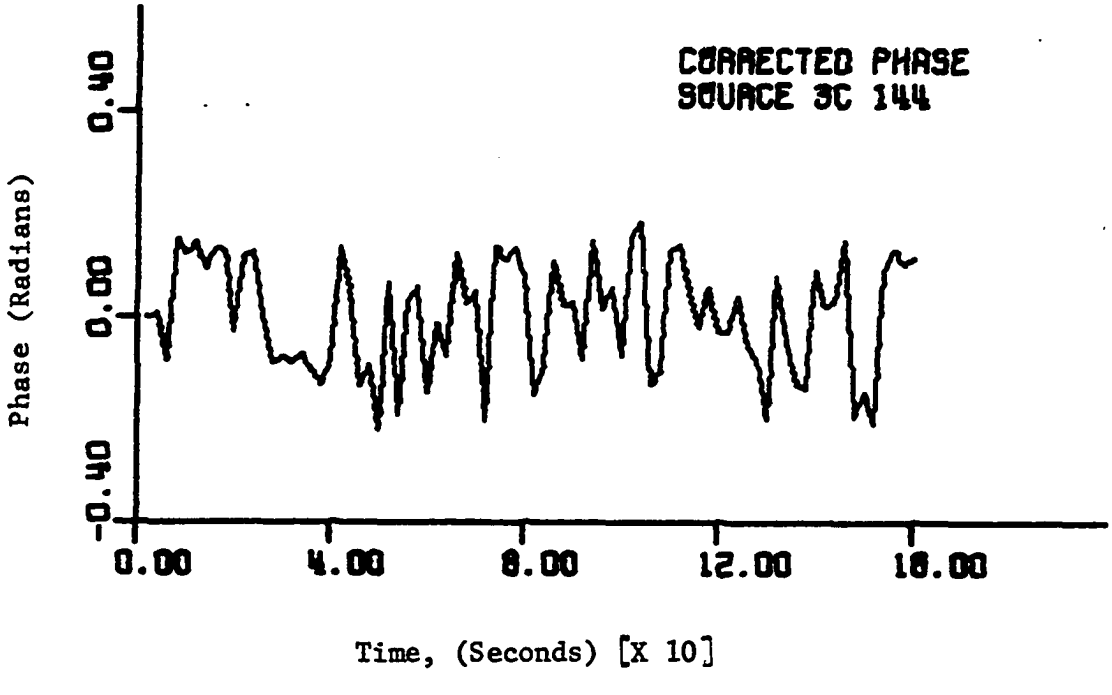


Figure 4.a.16. Phase data corrected for clock drift and ionospheric noise. Data collected on August 20, 1977

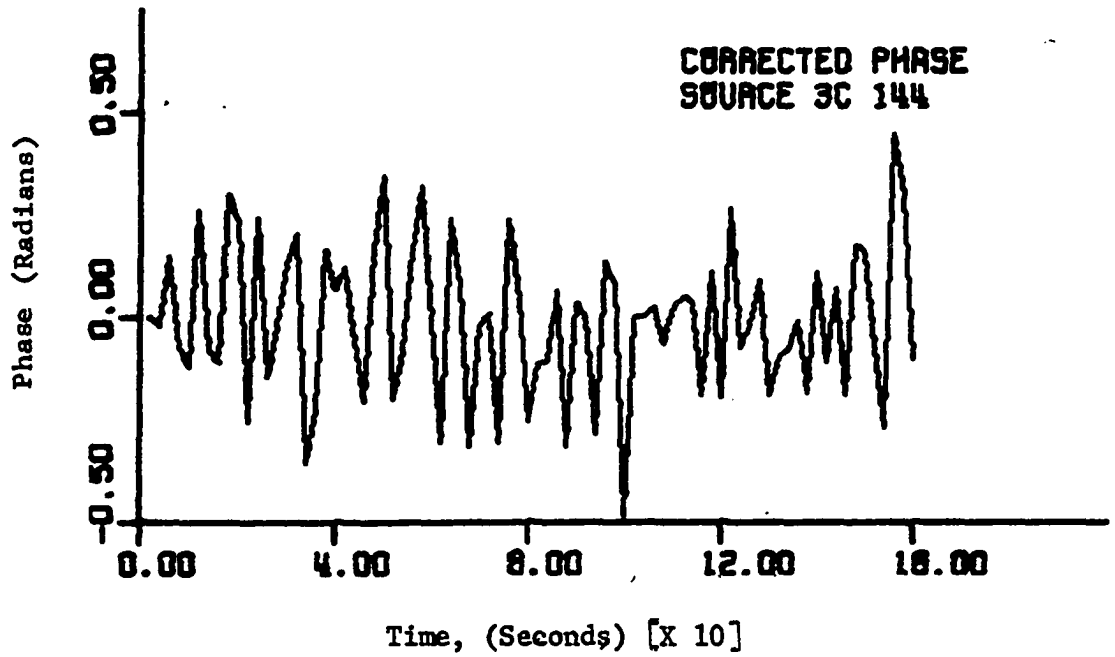


Figure 4.a.17. Phase data corrected for clock drift and ionospheric noise. Data collected on August 20, 1971

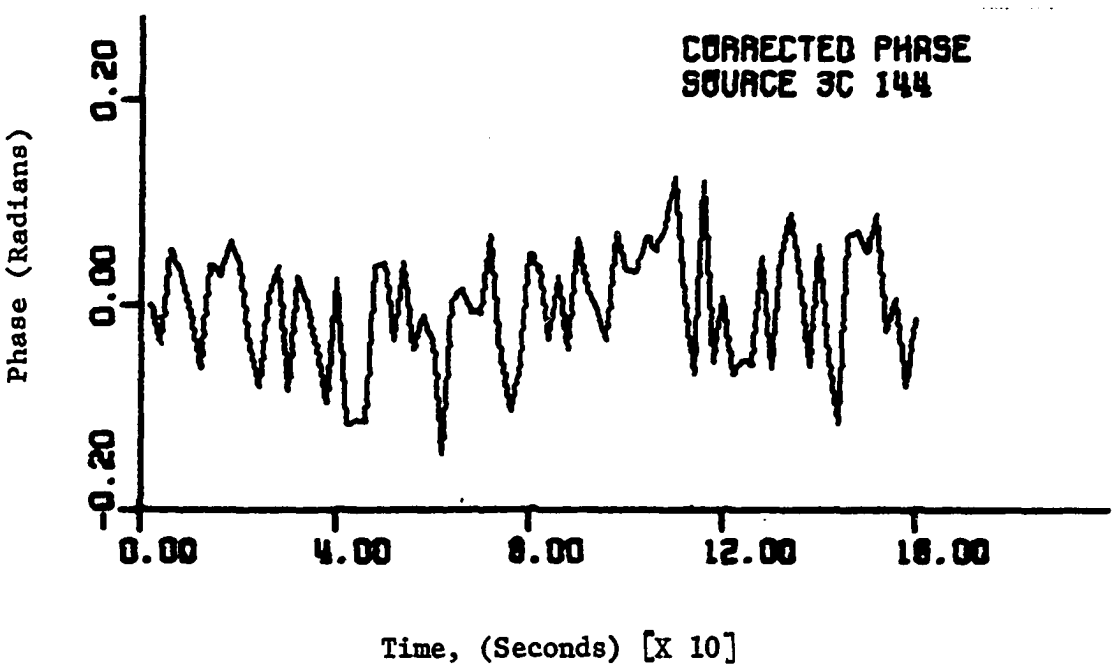


Figure 4.a.18. Phase data corrected for clock drift and ionospheric noise. Data collected on August 20, 1971

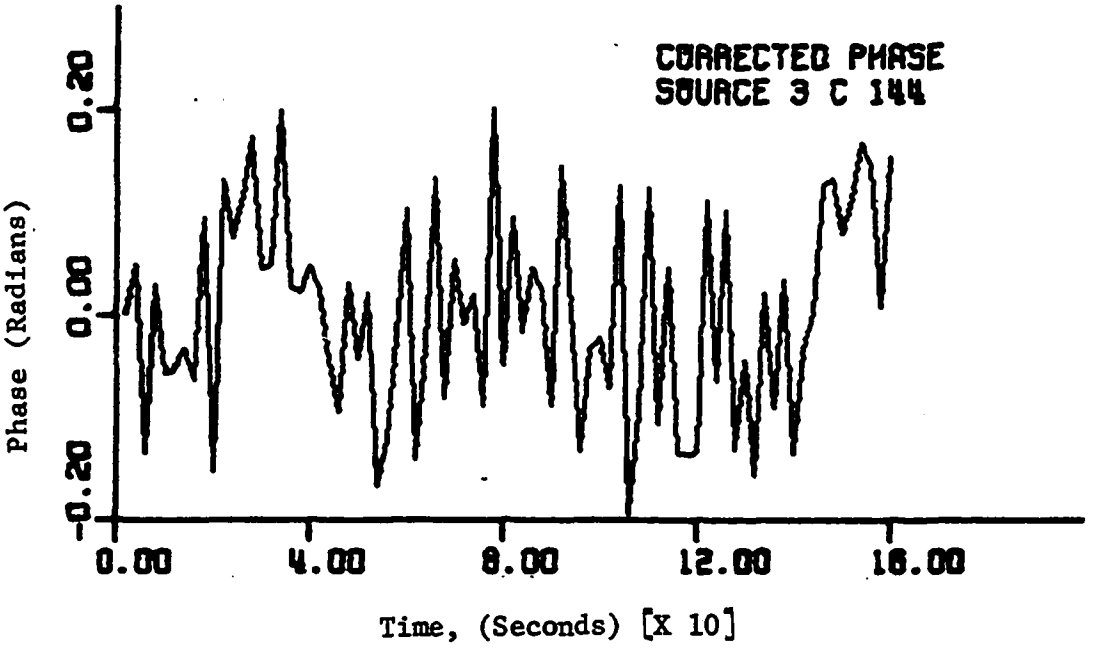


Figure 4.a.19.. Phase data corrected for clock drift and ionospheric noise. Data collected on August 22, 1971

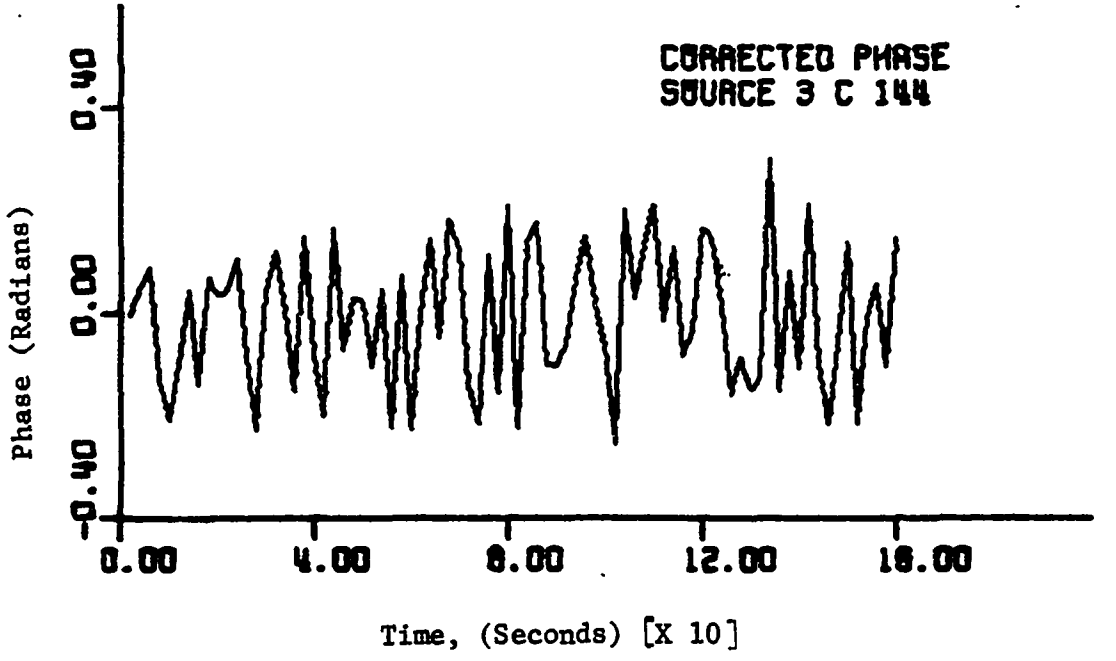


Figure 4.a.20. Phase data corrected for clock drift and ionospheric noise. Data collected on August 29, 1971

means the filter did not remove the ionospheric phase noise very well or the complex source visibility function has a nonzero phase for the effective baseline corresponding to these two observations.

This point is very important. Any low-pass filter removes the slowly varying phase noise components and reduces the variance of the phase noise data. A low-pass filter removes the phase component due to the angular structure of a radio source as well as it removes the other slow varying phase noises.

The present technique, when used with care, does not remove the slowly varying source phase component. Thus, it provides information about the complex fringe visibility phase while it reduces the phase noise. This feature is a unique characteristic of the Kalman filter. Using the filtered phase data and amplitude data, a new set of complex cross-correlation numbers were obtained for each observation. Then these new correlation numbers were averaged coherently to obtain a new coherent estimate of the normalized cross-correlation which is related to the complex fringe amplitude. This process was repeated for each observation. In Chapter II, we discussed the incoherent averaging technique. The fringe amplitude is overestimated by a factor of $1+\frac{1}{2}(\mathrm{SNR})^{-2}$, where SNR is the signal-to-noise ratio of the interferometer. As mentioned previously, the coherent averaging underestimates the fringe amplitude by a factor of $\exp(-\sigma_\phi^2)$, where σ_ϕ^2 is the variance of the phase noise process. Thus the technique presented in this thesis lies between these two methods.

By the use of the bias recovery method given in the previous chapter the bias may be recovered at the expense of an increase in the uncertainty in the final estimation of the fringe amplitude. This increase in the un-

certainty results from the fact that with a limited number of phase noise samples, the variance of the phase noise process cannot be estimated exactly.

The techniques to recover the bias due to coherent averaging, and to remove the phase noises due to the local oscillators and the ionosphere were applied to the data from observations of 3C 144, the Crab Nebula.

The ionosphere and oscillator phase noise models were obtained by a regression analysis on the phase data corrected for 2π ambiguities.

The model parameters obtained from the regression analysis are given in Table 4.a.1. These parameters were used for the noise removal method.

Table 4.a.1. Ionosphere and oscillator phase noise model parameters.

Observation number	ω (rad/sec)	β	$\varphi_{0_2}^2$ (rad^2)	σ_{stat}^2 (rad)^2	$X_5(0)$ (rad/sec)
1	0.19634	0.352	2.959	1.379	-0.00374
2	0.11635	0.075	7.515	1.251	-0.15004
3	0.31416	0.122	8.628	1.540	0.05014
4	0.10833	0.021	32.427	1.499	-0.06617
5	0.10134	0.054	32.344	2.185	0.08593
6	0.10472	0.084	12.969	1.922	-0.01959
7	0.26180	0.015	4.034	1.196	0.11127
8	0.10134	0.054	26.509	2.054	0.19483
9	0.22439	0.081	11.851	1.508	-0.14017
10	0.10134	0.027	7.043	1.4349	0.08213

Observations 3 and 7 had the largest values of ω, indicating that the scale sizes of the ionospheric blobs were larger than the blobs on the other days. The largest β values were observed during the first and third observations. The large β values show that during these two observations the correlation distances of the blobs were short compared to the correlation distances for the other observations. As seen from the data plots in Figures 4.a.2, 4.a.7, 4.a.8 and 4.a.9, the clock drifts had larger values than on the other days.

Comparing the 4th and 5th rows on Table 4.a.1, which were two observations made on the same day with seven-minute time intervals, we see that all corresponding parameters are close to each other except for σ^2_{stat} and $X_5(0)$. The large variances of the ionospheric phase, φ_0^2, on these two observations is a measure of the noisiness of the data shown in Figures 4.a.4 and 4.a.5.

In our analysis, we assumed that the frequency offset was removed from the data. Any frequency offset error adds a term $\Delta(\Delta\omega')t$ to the residual fringe phase as seen from Eq. (1.D.1) in Chapter I, where $\Delta\omega'$ is the frequency offset introduced in the beginning of each VLBI experiment, and $\Delta(\Delta\omega')$ is the error in $\Delta\omega'$. Therefore, the discrepancy between linear trend slopes, $X_5(0)$, could have been caused by an incorrect frequency offset value during preprocessing of the data.

The ionospheric phase noise model parameters ω, β and φ_0^2 approximately have the same values in the two observations. This means that the ionospheric disturbance is not changing its character so rapidly.

By the use of the phase noise removal method given in the previous chapter, the clock and ionospheric phase noises were removed.

Using these data we obtained a coherent average of the complex correlation function. Table 4.a.2 gives the amplitude of the coherent average along with the original coherent average amplitudes as given by the VLBI II program output.

Table 4.a.2. The original and corrected coherent averages of the complex cross-correlation functions.

Observation number	Original coherent average X 10 E-4	Phase noise Corrected coherent average X 10 E-4
1	5.8	15.6
2	16.9	41.5
3	9.4	18.0
4	4.6	14.8
5	11.9	31.4
6	3.8	15.4
7	7.2	25.9
8	4.2	22.2
9	5.3	16.7
10	5.7	15.6

From the review of Table 4.a.2 we note that when the phase noise due to the ionosphere and the local oscillators are removed the coherent averages of the cross-correlation functions are increased. The average for the second column is 7.48×10^{-4} and for the third column is 21.71×10^{-4}. Thus,

the increase in the average value is 2.9 times. Assuming a Gaussian distribution for the normalized cross-correlation functions, the standard error for both columns is given as

$$\Delta\delta = 2 \times 2.14 \times 10^{-4}$$
$$= 4.28 \times 10^{-4}$$

where $\Delta\delta$ is the standard error for 95% confidence interval. The number 1.92×10^{-4} is the standard deviation of the coherent estimator for the normalized cross-correlation function which was calculated during preprocessing of the data.

Using the phase data samples for each observation, the variance of the phase noise was computed before and after the phase noise removal process.

By using the estimated variance of the phase noise process for each observation, the bias removal technique described in Chapter III was applied to the data given in Table 4.a.2. The bias corrected data are given in Table 4.a.3.

We notice from Table 4.a.3 that when we apply the bias correction method to the raw data we get correlation values of almost three times the original values.

The application of the phase noise removal technique increases the original correlation values almost three times. For the latter case, the standard errors are much smaller than that of the data with phase noises not removed.

Table 4.a.3. The bias removed coherent average data for the original and phase noise removed data.

Observation number	Bias corrected coherent average X 10 E-4	Standard error X 10 E-4	Bias and phase noise corrected coherent average X 10 E-4	Standard error X 10 E-4
1	15.3	9.0	16.0	3.2
2	58.1	24.8	41.7	2.9
3	22.6	9.5	18.2	3.0
4	14.7	10.9	14.8	2.9
5	35.4	15.0	31.4	2.9
6	11.9	10.2	15.5	2.9
7	16.9	8.3	26.4	3.0
8	17.6	14.8	22.3	2.9
9	17.2	11.5	16.8	2.9
10	16.0	32.2	15.8	3.0

The 5th column shows the standard error associated with each observation in column 4. The standard errors are very similar to each other even though the ionospheric phase noise is not the same for every observation. This can be explained by noticing that column 4 was obtained by applying the bias removal technique to the 3rd column of Table 4.9.2.

After phase-noise removal the fluctuations in the phase data were very small, therefore, the variance of each phase data, $\hat{\sigma}^2_{\phi}$, was close to zero. In the calculation of the fifth column these estimated small variances were used. The variance of the new fringe estimator was given in Chapter III as

$$\sigma^2_{\hat{Z}} = [(\pi^2 \exp \hat{\sigma}^2_{\phi})/8BT] + \hat{\sigma}^4_{\phi} \hat{Z}^2/(n-1).$$

For small values of $\hat{\sigma}^2_{\phi}$, as seen from the above formula, $\sigma^2_{\hat{Z}}$ does not change considerably because the deviation of function $\exp\hat{\sigma}^2_{\phi}$ is small about $\hat{\sigma}^2_{\phi} = 0$.

Because of the same reason the 3rd column of Table 4.a.2 and the 4th column of Table 4.a.3 are very close to each other. Since the latter was obtained by using the equation

$$\hat{Z} = \hat{R} \exp(\hat{\sigma}^2_{\phi}/2).$$

We expect that the numbers given in the 4th column must have less spread than those given in the 2nd column, i.e., the standard deviation for the phase noise corrected data should be smaller than the standard deviation of the data not corrected for the ionospheric and clock noise effects. These statistics are given in Table 4.a,4. Since this table was obtained from 10 observations, we would expect that the standard error of the overall 10 observations would be smaller than those given in Table 4.a.3, column 5. Comparing the two tables we notice that this is not true, because the phase data used did not represent the true phase noise exactly.

The entries in column 4 of Table 4.a.3 disagree with each other more than the standard error in column 5 predicts. The standard error is a measure of the phase noise whereas column 4 contains amplitude information which was not filtered.

Table 4.a.4. Average and standard deviations of the phase noise corrected and the raw data for 10 observations.

Statistic	Corrected for phase noise	Not corrected for phase noise
Average (X 10 E-4)	22.12	22.64
Standard deviation (X 10 E-4)	8.86	14.06

In the case of a smaller sampling period we would expect to use more complicated phase noise models, i.e., a noise model described by more than two differential equations.

The Kalman filter algorithm involves matrix inversions whenever we have a measurement. The matrix inversion takes an excessive amount of computer time if the dimension of the matrix is big. Therefore, the complicated phase noise models must be avoided whenever possible.

The proposed method for reducing the phase noise can also be used at high frequencies where the earth's atmosphere is a major phase noise source. The atmosphere has a nonumiform refractive index as does the ionosphere. The source of the nonuniform refractive index in the atmosphere is often attributed to the water vapor content distribution of this medium. A typical atmospheric phase noise is shown in Figure 4.a.21. The curve was obtained by Basart el al. at centimeter wavelengths with a radio link connected interferometer (Basart et al., 1970). The observed phase noise due to the atmosphere suggests that we can also use the periodic random phase noise model for the atmosphere.

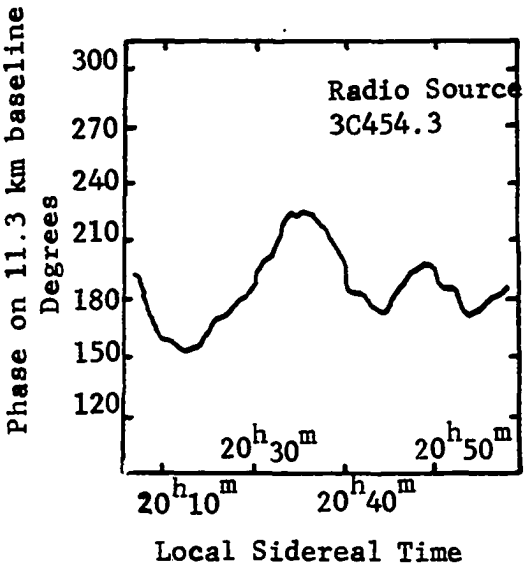


Figure 4.a.1. A typical atmospheric phase noise curve (After Basart et al.)

At high frequencies the errors in the frequency standards used for VLBI cause more problems than at low frequencies (Moran, 1976). Therefore, our simple clock phase noise model may not apply to the high frequency case. However, an appropriate model can be found by observing the differential phase behavior of two atomic standards operating at high frequencies.

One point must be made clear here. That is, we always assumed that the errors on the baseline parameters, the source position and the frequency offset were negligible. However, in practice this is not the case. Especially, at high frequencies a small error on any of the above parameters can cause a linear phase drift. By using different baseline and source parameters the linear phase drift may be eliminated.

B. Half-power Size of the Pulsar NP 0532 of the Crab Nebula

Assuming a one dimensional Gaussian source brightness distribution model, visibility of the source is given by

$$\gamma(u) = \exp\left[-\frac{\pi^2\theta^2u^2}{4\ell n\,2}\right] \tag{4.B.1}$$

where $\gamma(u)$ is the visibility amplitude of the source, θ is the half-power size of the source, and u is the length of the projected baseline on the u - v plane.

For the half-power size of the source we obtain from Eq. (4.B.1)

$$\theta = \left[-\frac{4\ \ell n\ 2}{\pi^2 u^2}\ \ell n\ \gamma(u)\right]^{\frac{1}{2}} \tag{4.B.2}$$

The uncertainty on the source size θ is given by

$$\frac{\Delta\theta}{\theta} = \frac{\Delta\gamma(u)}{\gamma(u)}\ [1/2\ \ell n\ \gamma(u)], \tag{4.B.3}$$

provided the u in Eq. (4.B.2) is known perfectly.

The fringe visibility of the source is defined as

$$\gamma(u) = \frac{S_c}{S} \tag{4.B.4}$$

where S_c is the correlated flux density of the source at the long baseline and S is the total flux density of the source. The unit for flux density is the Jansky ($= 10^{-26}\,W\,m^{-2}Hz^{-1}$).

The uncertainty in the fringe visibility is calculated by using the error-propagation formula as follows:

$$(\Delta\gamma)^2 \simeq \left(\frac{\partial\gamma}{\partial S_c}\right)^2(\Delta S_c)^2 + \left(\frac{\partial\gamma}{\gamma S}\right)^2(\Delta S)^2 \tag{4.B.5}$$

From Eq. (4.B.4) we obtain

$$\frac{\partial \gamma}{\partial S_c} = \frac{1}{S} , \tag{4.B.6}$$

$$\frac{\partial \gamma}{\partial S} = - \frac{S_c}{S^2} \tag{4.B.7}$$

Substituting Eqs. (4.B.6) and (4.B.7) into Eq. (4.B.5) yields

$$\begin{aligned} (\Delta\gamma)^2 &= (\frac{1}{S})^2 (\Delta S_c)^2 + (- \frac{S_c}{S^2})^2 (\Delta S)^2 \\ &= \frac{1}{S^2} (\Delta S_c)^2 + \frac{S_c^2}{S^2} (\frac{\Delta S}{S})^2 \end{aligned} \tag{4.B.8}$$

From Eq. (4.B.8) we have

$$\Delta\gamma = [\frac{(\Delta S_c)^2}{S^2} + \frac{S_c^2}{S^2} (\frac{\Delta S}{S})^2]^{\frac{1}{2}} \tag{4.B.9}$$

From calculated normalized cross-correlation we can find the correlated source flux density by the relation

$$S_c = \delta (T_{sys\,1} T_{sys\,2} / e_1 e_2)^{\frac{1}{2}}$$

where S_c is the correlated flux density, δ is the normalized cross-correlation coefficient, $T_{sys\,1}$ and $T_{sys\,2}$ are the total system temperatures at the two stations, and e_1 and e_2 are the antenna gains in kelvins per Jansky corrected for the appropriate source hour angle at the two stations.

The value of $[(T_{sys\,1} T_{sys\,2})/e_1 e_2]$ was approximately 3.45×10^4 Jan. for the system used during the experiments in 1971 August.

The corrected value of the correlated flux density for the compact component in the source 3C 144 was 76.4 Jan. If we take 800 Jan. for the total flux density of this compact component (Cronyn, 1970, p. 153) the

visibility is computed as 0.095. Using a baseline length of 8.24×10^4 in wavelengths, we calculated the half-power size of the source as

$$\theta = 2''.00 \pm 0''.24$$

To compare this result with other published results we have to correct θ for the interplanetary scintillation effects. The half-power size for a point source due to the interplanetary medium is given by (Cronyn, 1970, p. 7)

$$\theta_{\frac{1}{2}} = 6.49 \times 10^{-2} P^{-2} \lambda^2 \text{ arc-sec}, \qquad (4.B.10)$$

where λ is the wavelength, P is the closest approach to the sun given by

$$P = (V \tan \epsilon)/\Omega , \qquad (4.B.11)$$

where V is the velocity of the blobs in the interplanetary medium and Ω is the angular rotational velocity of the sun, and ϵ is the elongation angle of the source.

Typical values for Ω and V are 400 rad/sec and 400 km/sec, respectively. The average elongation angle of the source was $65^{\circ}.4$ during the August 1971 observations. By using these values in Eq. (4.B.11) we obtain $P \simeq 2.184$ A. . Substituting this value of P and $\lambda = 11.4$ m into Eq. (4.B.10) yields

$$\theta_{\frac{1}{2}} = 6\ 49 \times 10^{-2} \times (2.184)^{-2} \times (11.4)^2 \text{ arc-sec}$$

$$\simeq 1.77 \text{ arc-sec.}$$

Since $\theta_{\frac{1}{2}}$ and θ are independent, the size of the pulsar corrected for the interplanetary scintillations is given by

$$\theta_c = [\theta^2 - \theta_{\frac{1}{2}}^2]^{\frac{1}{2}}. \qquad (4.B.12)$$

Substituting $\theta = 2''.00$ and $\theta_{\frac{1}{2}} = 1''.77$ into Eq. (4.B.12) yields

$$\theta_c = [4 - (1.77)^2]^{\frac{1}{2}} = 0''.93.$$

Assuming no errors in $\theta_{\frac{1}{2}}$ and using the error propagation formula we find that the error in θ_c is $\Delta\theta_c = \pm 0''.26$. Thus, we have

$$\theta_c = 0''.93 \pm 0''.26.$$

In 1974 Mutel et al. (1974) measured the Gaussian half-power size of the same source (the pulsar NP 0532) by using a very long baseline interferometer operating at 26.3 MHz. Their baseline was 187,000 λ and the interferometer consisted of Arecibo Observatory's 1000-foot (305 m) spherical reflector and the University of Florida's 640-element dipole array. The Gaussian half-power size for the source they measured was $1''.3(+0''.23, -0''.13)$. This result and our result agree with each other, although the interferometer used in the August 1971 observations had much less sensitivity than the interferometer used by Mutel et al.

Scattering by irregularities in the interstellar medium causes the apparent size of a radio source to increase with wavelength. Figure 4.b.1 shows the sizes of the pulsar NP 0532 at various frequencies. The solid line in the figure was obtained by a least-squares fit to the plotted points. The slope of the line is 2.05 ± 0.25. Thus the measurements of

the pulsar NP 0532 are in a good agreement with λ^2 scattering law.

As a result we conclude that it is possible to recover the bias due to coherent averaging when the ionospheric scintillation effects are present. Reducing the fringe phase noise increases the coherent averaging time, which increases the sensitivity of the interferometer.

However, one must be careful in phase noise modeling. The Kalman filter is quite sensitive to modeling errors. When the noise process is not modeled properly it is quite possible to add more noise to the phase data instead of removing the noise.

Further study may be done to implement the proposed technique in real time. That is, an on line computer would estimate the noise model parameters and subtract the predicted ionospheric phase noise from the phase data.

For complicated noise models the use of the computer in real time becomes very difficult owing to the matrix inversion processes in the Kalman filter algorithm.

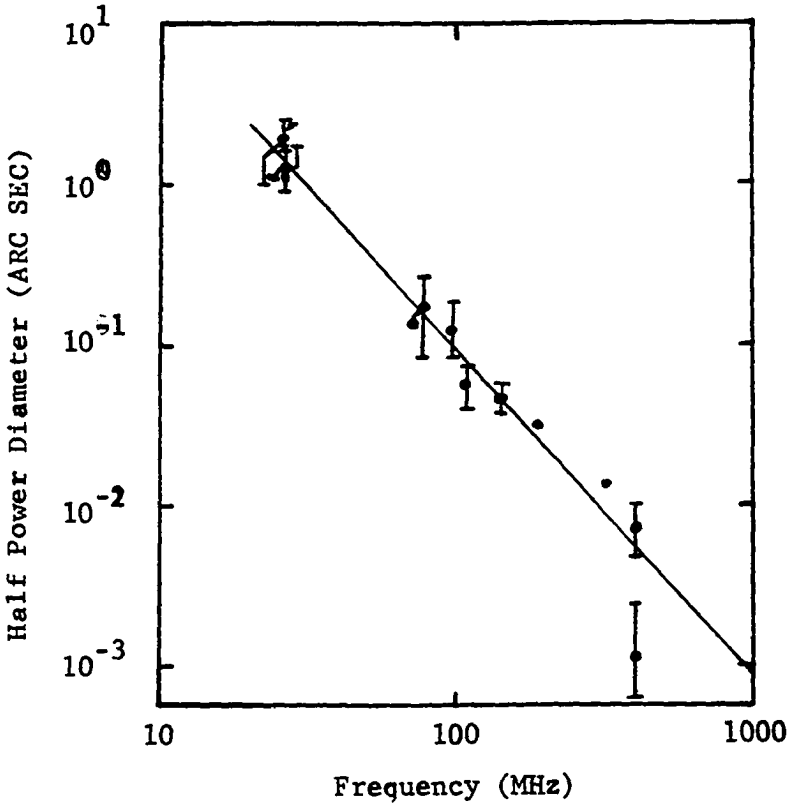


Figure 4.b.1. Gaussian half-power diameter versus frequency for NP 0532 (After Mutel et al., 1974)

V. A SIMULATION OF KALMAN FILTER AS IT APPLIES TO RADIO INTERFEROMETRY

In this chapter, a computer simulation of a Kalman filter application will be discussed. We will assume that the signal from a radio source is a Gaussian random process, $w(t)$, with a mean of zero and a variance of σ_s^2 which may be measured. Assuming a local oscillator operating at the angular frequency ω_0 and having no phase noise, we can model one end of an interferometer as shown in Figure 5.1.

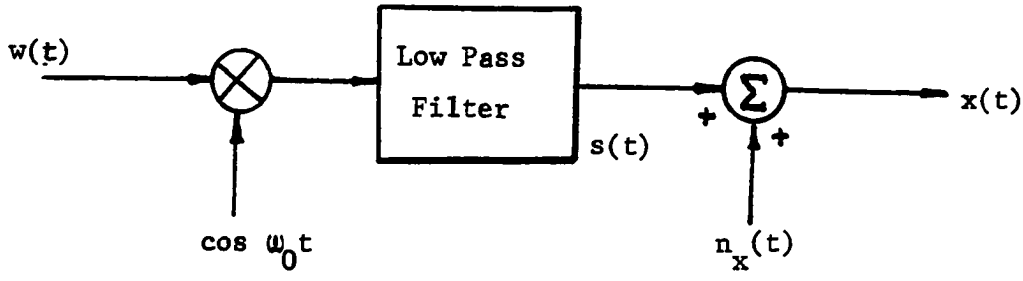


Figure 5.1. Receiver model

From the figure we write

$$x(t) = s(t) + n_x(t). \tag{5.1}$$

If we assume that at the other end of the interferometer we have a similar system, we write

$$y(t) = s(t) + n_y(t) \tag{5.2}$$

where $n_x(t)$ and $n_y(t)$ are normal white noise processes at the two ends of the interferometer, respectively. They represent receiver and other meas-

urement noises. The noise powers $\sigma_{n_x}^2$ and $\sigma_{n_y}^2$ can be measured when the source is not in the antenna beams.

We want to estimate the correlated power $E[s^2(t)]$ within the band determined by the low pass filters, where $E[\cdot]$ denotes the expectation operator. If $w(t)$, $n_x(t)$ and $n_y(t)$ are ergodic processes, we estimate $E[s^2(t)]$ by

$$\hat{R} = \frac{1}{T}\int_0^T x(t)y(t)dt \tag{5.3}$$

where T is the observation time of signals x(t) and y(t).

If we do not have the exact geometric delay τ_g, we introduce a time shift in one of the signals x(t) or y(t). Thus,

$$\hat{R}(\tau) = \frac{1}{T}\int_0^T x(t)y(t-\tau)dt. \tag{5.4}$$

Substituting Eqs. (5.1) and (5.2) into Eq. (5.4) yields

$$\hat{R}(\tau) = \frac{1}{T}\int_0^T [s(t)+n_x(t)][s(t-\tau)+n_y(t-\tau)]dt$$

$$= \overline{[s(t)+n_x(t)][s(t-\tau)+n_y(t-\tau)]} \tag{5.5}$$

where the overbar denotes the time average. Eq. (5.5) can be written as

$$\hat{R}(\tau) = \overline{s(t)s(t-\tau)} + \overline{s(t)n_y(t-\tau)}$$

$$+ \overline{n_x(t)s(t-\tau)} + \overline{n_x(t)n_y(t-\tau)} \tag{5.6}$$

As $T\to\infty$ the last three terms in Eq. (5.6) go to zero if s(t), $n_x(t)$ and $n_y(t)$ are statistically independent of each other.

Thus,

$$\lim_{T \to \infty} \hat{R}(\tau) = E[s(t)s(t-\tau)] \tag{5.7}$$

For a finite value of T, the standard error for $\hat{R}(\tau)$ is given by

$$\Delta\hat{R}(\tau) = \left(\frac{\hat{\sigma}_x^2 \hat{\sigma}_y^2}{2BT}\right)^{\frac{1}{2}} \tag{5.8}$$

where B is the noise equivalent bandwidth of the low-pass filter shown in Figure 5.1, $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$ are the estimated variances of the processes x(t) and y(t). Intuitively, if we filter x(t) and y(t), then cross-correlate them this would result in a smaller uncertainty since the variances of the filtered signals are smaller than before filtering. It is well-known that the cross-correlation technique is the optimum method to get an estimate for the signal power $E[s^2(t)]$ if we have two noisy measurements of the signal. If

$$T \to \infty \ , \tag{5.9}$$

the measured cross-correlation function of x(t) and y(t) will be the true auto-correlation function of the signal s(t) as stated by Eq. (5.7). Therefore, if Eq. (5.9) holds, then no matter what we do to the signals x(t) and y(t) before cross-correlating them, there will be no difference, i.e. we can not improve the estimate by prefiltering. However, in reality Eq. (5.9) never holds, that is to say, T has a finite value.

We can perform filtering on x(t) and y(t) either in the frequency domain or in the time domain. If we want to use the digital computer, working in the time domain will save computing time. The filtering can be

done recursively when the signal and noise processes are modeled properly.

Let u(t) be the signal defined by

$$u(t) = w(t) \cos \omega_0 t \tag{5.10}$$

In the beginning of this chapter we assumed that w(t), the radio source signal, was a white noise process having a flat spectrum shown in Figure 5.2

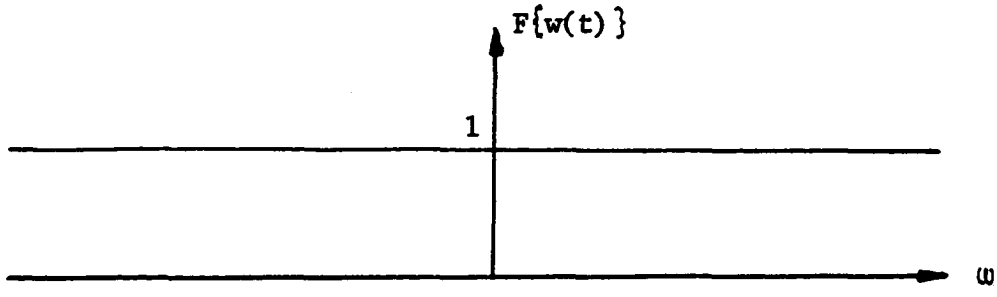


Figure 5.2. Spectrum of a white noise process.

where

$$W(\omega) = F\{w(t)\} = \int_{-\infty}^{\infty} w(t)\ e^{-j\omega t} dt \tag{5.11}$$

The Fourier transform of $\cos \omega_0 t$ is known to be

$$F\{\cos \omega_0 t\} = \frac{1}{2}\ \delta(\omega - \omega_0) + \frac{1}{2}\ \delta(\omega - \omega_0) \tag{5.12}$$

where δ is the Dirac delta function.

Then using the convolution theorem we find

$$F\{u(t)\} = U(\omega) = W(\omega) \circledast F\{\cos \omega_0 t\}$$

$$= W(\omega) \circledast [\frac{1}{2}\ \delta(\omega - \omega_0) + \frac{1}{2}\ \delta(\omega - \omega_0)]$$

where $\circledast$ denotes the convolution. From the definition of convolution we write

$$U(\omega) = \frac{1}{2}\int_{-\infty}^{\infty} W(\omega-\omega')\delta(\omega'-\omega_0)d\omega'$$

$$+\frac{1}{2}\int_{-\infty}^{\infty} W(\omega-\omega')\delta(\omega'+\omega_0)d\omega' \qquad (5.13)$$

$$=\frac{1}{2}W(\omega-\omega_0)+\frac{1}{2}W(\omega+\omega_0)$$

$W(\omega)$ was defined as a constant function of frequency, therefore, $U(\omega)$ is a constant function of frequency. Thus, u(t) is a white noise process. Now, let us consider the subsystem shown in Figure 5.3.

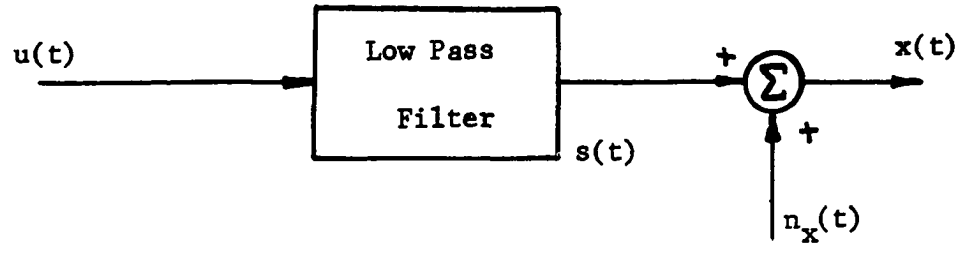


Figure 5.3. A subsystem of the receiver system

We can model the low pass filter in Figure 5.3 as a single negative feedback circuit as shown in Figure 5.4.

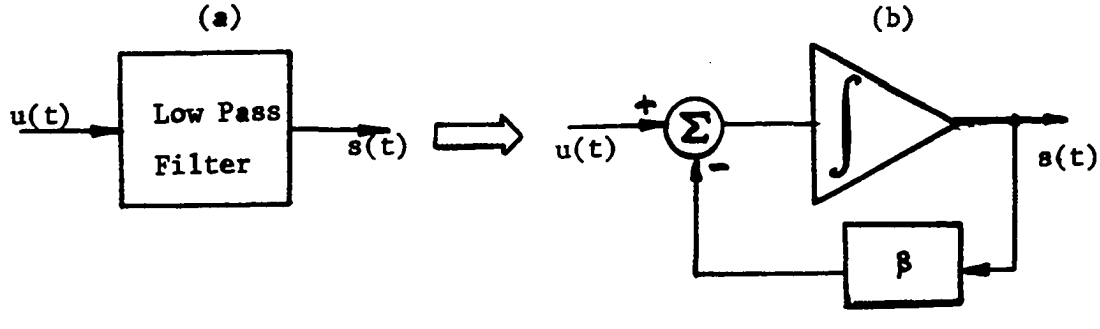


Figure 5.4. A low-pass filter model

From Figure 5.4.b we write

$$\frac{ds(t)}{dt} = -\beta\, s(t) + u(t) \tag{5.14}$$

where β is the feedback coefficient and is related to the bandwidth and the gain characteristics of the low pass filter. To relate β and the bandwidth, take the Fourier transform of both sides of Eq. (5.14). This gives

$$j\omega S(\omega) = -\beta\, S(\omega) + U(\omega) \tag{5.15}$$

where

$$S(\omega) = \int_{-\infty}^{\infty} s(t)\, e^{-j\omega t} dt \tag{5.16}$$

and

$$U(\omega) = \int_{-\infty}^{\infty} u(t)\, e^{-j\omega t} dt \tag{5.17}$$

To get Eq. (5.14) we assumed that

$$s(0) = 0. \tag{5.18}$$

From Eq. (5.15) the transfer function of the system is found to be

$$F(\omega) = \frac{S(\omega)}{U(\omega)} = \frac{1}{\beta + j\omega} \tag{5.19}$$

If B denotes the half bandwidth, by definition in hertz

$$\frac{1}{\beta^2 + (2\pi B)^2} = \frac{1}{2\beta^2} \tag{5.20}$$

Solving Eq. (5.20) for β yields

$$\beta = 2\pi B. \tag{5.21}$$

Eq. (5.14) defines a first order Markov process when the input function u(t) is a white noise stochastic process (Gelb, 1974), p. 42). If s(t) is sampled with the period Δt, the discrete time series corresponding to s(t) is given by (Gelb, 1974, p. 82)

$$s_k = e^{-\beta \Delta t} s_{k-1} + u_{k-1} \tag{5.22}$$

where s_k is the value of s(t) at time $k\Delta t$, $k = 0, 1, 2, \ldots$ and u_{k-1} is the value of u(t) at time $(k-1)\Delta t$.

The measured signal at the filter output was given by Eq. (5.1). The sampled version of this equation is simply

$$x_k = s_k + n_{xk} \tag{5.23}$$

We can represent the system defined by Eq. (5.22) and Eq. (5.23) in in a block diagram form as shown in Figure 5.5.

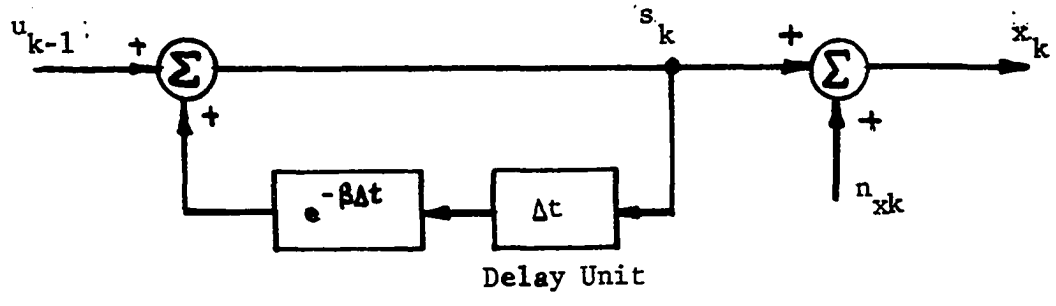


Figure 5.5. Signal and measurement models

Having defined the signal and measurement models, we want to use a Kalman filter to reduce the effects of the noise process n_{xk} on the signal

x_k. After filtering we have an estimate of s_k which will be denoted by $\hat{s}_k$. Now, we rewrite the signal and measurement model equations and state our assumptions:

$$\text{the signal model:} \quad s_k = e^{-\beta\Delta t} s_{k-1} + u_{k-1} \tag{5.24}$$

$$\text{the measurement model:} \quad x_k = s_k + n_{xk} \tag{5.25}$$

$$\text{Assumptions:} \quad E[s_k n_{xk}] = 0 \tag{5.26}$$

$$E[u_k^2] = q \tag{5.27}$$

$$E[u_k] = E[n_{xk}] = 0 \tag{5.28}$$

$$E[n_{xk}^2] = r_x \tag{5.29}$$

where q and r_x are the covariances of u_k and n_{xk}, respectively. Eq. (5.26) states that the signal and measurement noise processes are statistically independent. We have a similar set of equations for the second site of the interferometer:

$$s_k = e^{-\beta\Delta t} s_{k-1} + u_{k-1}, \tag{5.30}$$

$$y_k = s_k + n_{yk} , \tag{5.31}$$

$$E[s_k y_k] = 0 , \tag{5.32}$$

$$E[u_k^2] = q , \tag{5.33}$$

$$E[u_k] = E[n_{yk}] = 0 , \tag{5.34}$$

$$E[n_{yk}^2] = r_y . \tag{5.35}$$

We assume that the noise processes at the two ends of the interferometer are independent, i.e.

$$E[n_{xk} n_{yk}] = 0. \tag{5.36}$$

To see if the prefiltering improves the correlation between the noisy signals x(t) and y(t) we did a computer simulation as follows: Using a Fortran random number generating subroutine we obtained a normal discrete white noise process u_k with zero mean and variance 1. Eq. (5.22) was used to obtain the Markov signal s_k, and in this process we chose the value of $\beta\Delta t$ as 0.69. This corresponds to a sampling rate of $1/2B_m$, where B_m is the frequency at which the gain of the IF filter drops to approximately 0.05 times the gain at zero frequency.

By adding two independent normal processes with zero means and with variances 16 and 25, respectively, we obtained two noisy signals x(t) and y(t).

A recursive Kalman filter was applied to both of the signals x(t) and y(t). The filtered estimates of s(t) were then cross-correlated. The process is shown in a block diagram form in Figure 5.6.

The cross-correlation functions between various signals shown in Figure 5.6 were computed. The figures given in the following pages show normalized cross-correlation functions along with computed standard errors and T statistics given by

$$T = \hat{\delta}_{xy}\sqrt{N-1} \tag{5.37}$$

where $\hat{\delta}_{xy}$ is the computed normalized cross-correlation coefficient and N

is the number of samples used. For a 95% significance level T should be larger than 1.96 for $\hat{\delta}$ to be significant.

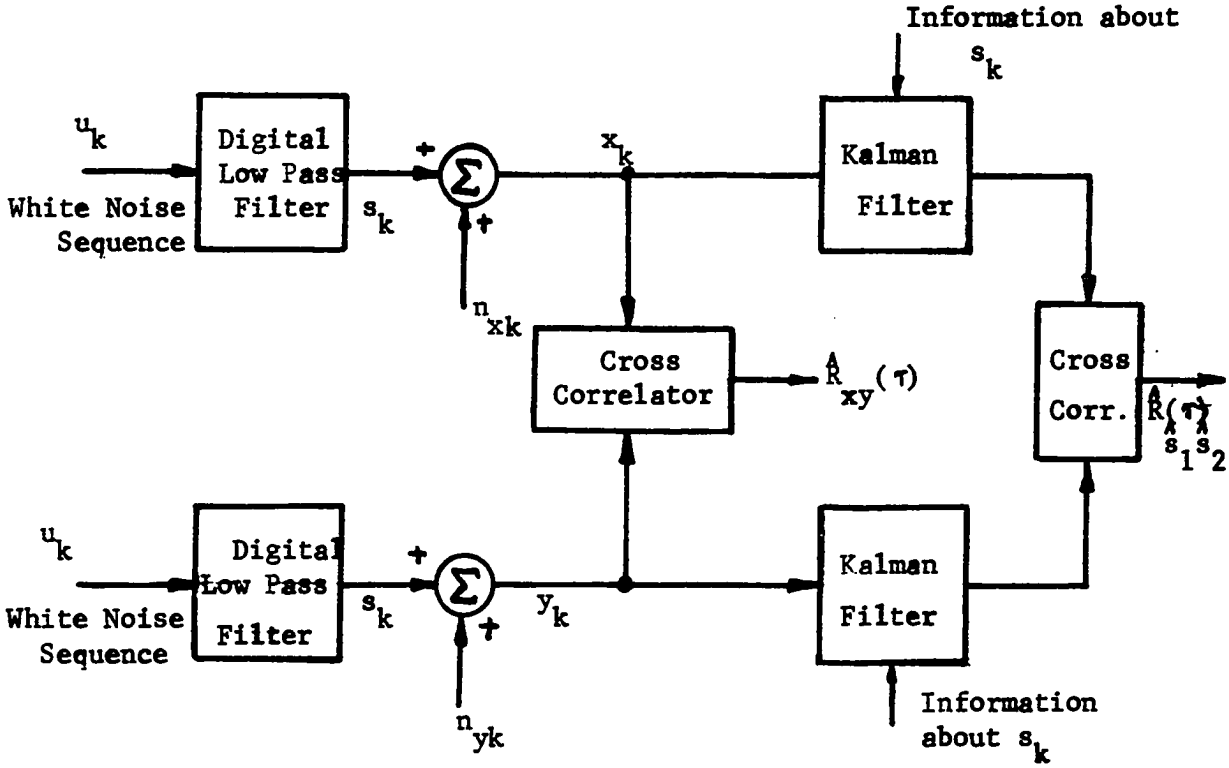


Figure 5.6. The block diagram of the prefiltering process.

Lag	Corr	S.E.	T	
-8	-0.018	0.032	-0.564	I
-7	-0.011	0.032	-0.358	I
-6	0.000	0.032	0.011	I
-5	0.061	0.032	1.917	I
-4	0.040	0.032	1.263	I
-3	-0.043	0.032	-1.357	I
-2	0.005	0.032	0.159	I
-1	0.012	0.032	0.381	I
0	0.127	0.032	4.016	I
1	0.035	0.032	1.107	I
2	0.033	0.032	1.027	I
3	0.024	0.032	0.770	I
4	0.026	0.032	0.824	I
5	-0.034	0.032	-1.074	I
6	-0.037	0.032	-1.173	I
7	0.002	0.032	0.078	I
8	0.006	0.032	0.173	I

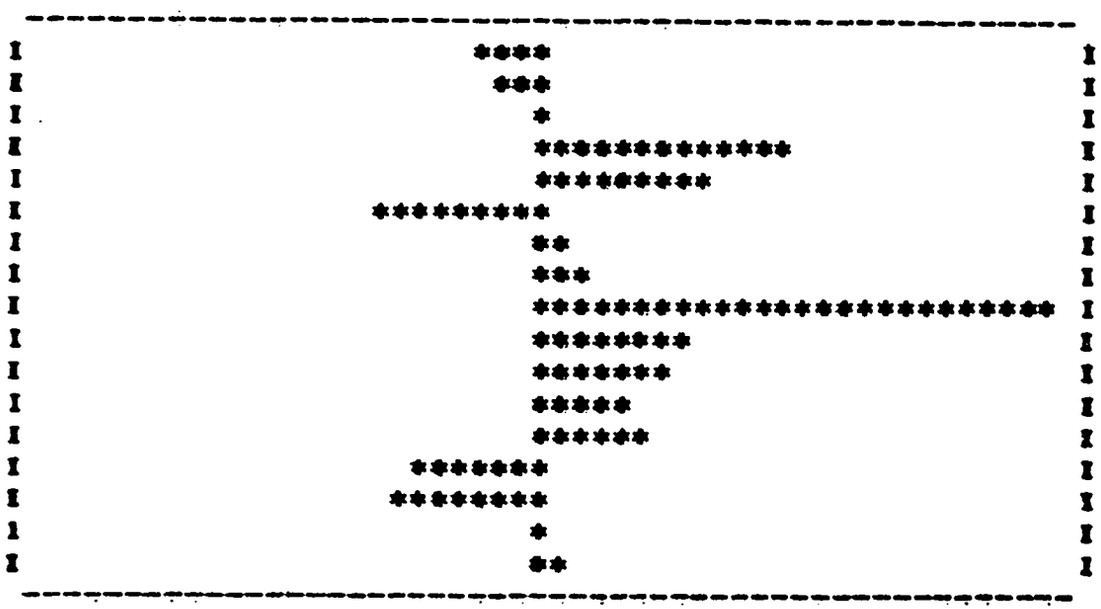


Figure 5.7. Normalized cross-correlation function of x(t) and y(t).

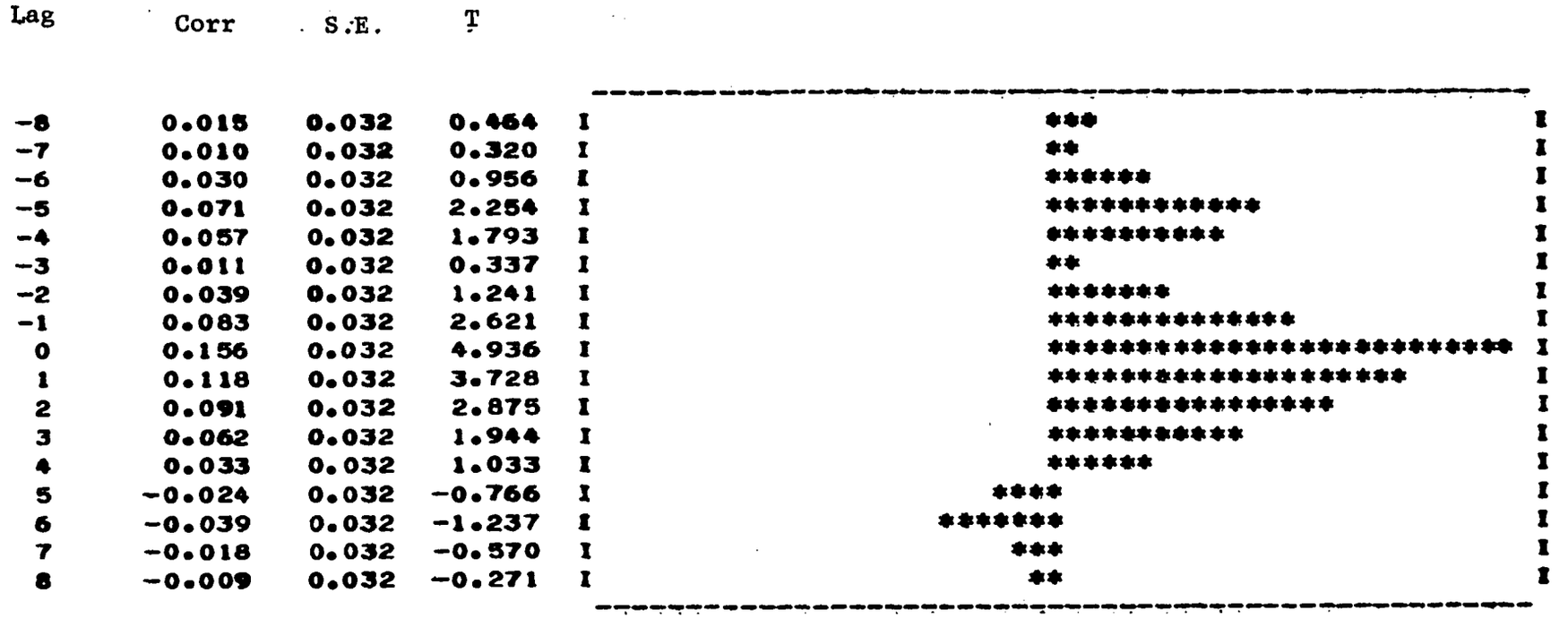


Figure 5.8. Normalized cross-correlation function of $\hat{s}_1(t)$ and $\hat{s}_2(t)$.

Lag	Corr	S.E.	T	
-8	0.022	0.032	0.702	I
-7	-0.002	0.032	-0.066	I
-6	-0.014	0.032	-0.448	I
-5	0.039	0.032	1.218	I
-4	0.017	0.032	0.538	I
-3	-0.005	0.032	-0.145	I
-2	0.000	0.032	0.001	I
-1	-0.008	0.032	-0.248	I
0	0.052	0.032	1.636	I
1	0.021	0.032	0.675	I
2	-0.040	0.032	-1.249	I
3	-0.027	0.032	-0.856	I
4	-0.024	0.032	-0.768	I
5	-0.033	0.032	-1.036	I
6	0.006	0.032	0.203	I
7	-0.003	0.032	-0.103	I
8	0.028	0.032	0.882	I

Figure 5.9. Normalized cross-correlation function of x(t) and y(t) with s(t) = 0

Lag	Corr	S.E.	T
-8	-0.027	0.032	-0.858
-7	-0.009	0.032	-0.294
-6	0.016	0.032	0.495
-5	0.053	0.032	1.678
-4	0.048	0.032	1.513
-3	-0.047	0.032	-1.475
-2	-0.015	0.032	-0.481
-1	-0.011	0.032	-0.349
0	0.057	0.032	1.818
1	0.003	0.032	0.108
2	0.022	0.032	0.696
3	0.032	0.032	0.999
4	0.032	0.032	1.006
5	-0.024	0.032	-0.756
6	-0.043	0.032	-1.349
7	0.006	0.032	0.179
8	-0.002	0.032	-0.078

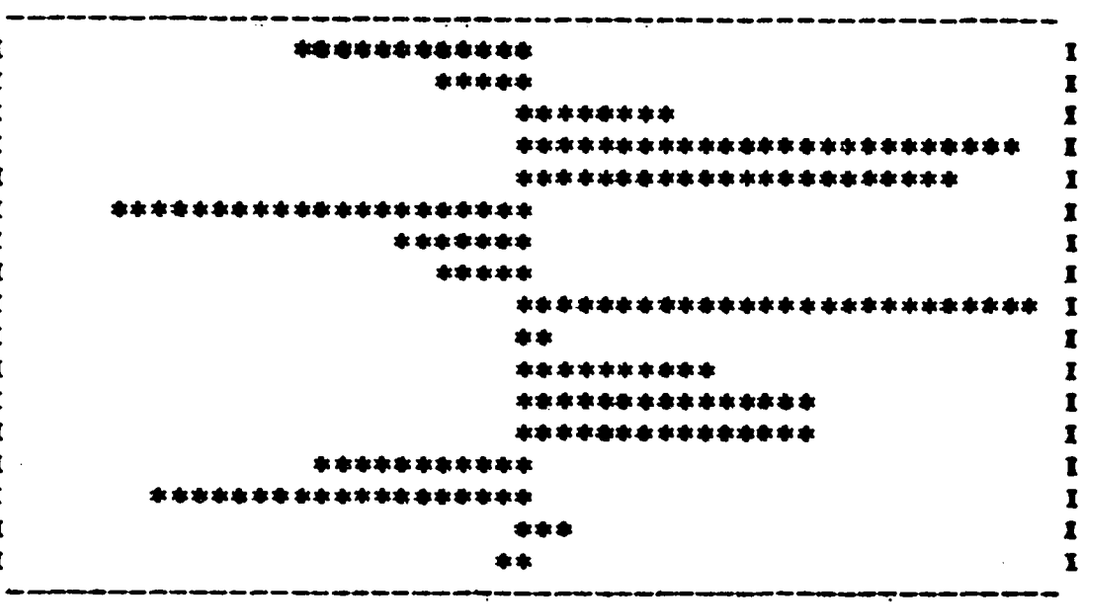


Figure 5.10. Normalized cross-correlation function of $n_x(t)$ and $n_y(t)$.

Figures 5.7 and 5.8 were obtained by cross-correlating the signals x(t) and y(t) before and after filtering, respectively. The noisy signals x(t) and y(t) both contain a first order Markov process s(t) defined by the difference equation

$$s_k = 0.5\, s_{k-1} + u_{k-1} \tag{5.38}$$

where U_{k-1} is the $(k-1)^{th}$ sample from the white normal process $u(t) \sim N(0,1)$.

To compare the standard correlation and prefiltering techniques we define the signal-to-noise ratio of the correlation process. The normalized cross-correlation function $\delta(h)$ was obtained according to the formula

$$\delta(h) = \Big[\sum_{i=1}^{n-h}(x_i - \bar{x})(y_i - \bar{y})\Big] \Big/ \Big[\Big[\sum_{i=1}^{n}(x_i - \bar{x})^2\Big]\Big[\sum_{i=1}^{n}(y_i - \bar{y})^2\Big]\Big]^{\frac{1}{2}} \tag{5.39}$$

where $\bar{x}$ and $\bar{y}$ are the computed average values of the random processes x(t) and y(t), respectively, h is the time lag, and n is the total number of samples.

We define the cross-covariance between x(t) and y(t) for lag h as

$$\mathrm{Cov}(X,Y) = \frac{1}{n-h}\sum_{i=1}^{n-h}(x_i - \bar{x})(y_i - \bar{y}) \tag{5.40}$$

and we have

$$\hat{\sigma}_X^2\,\hat{\sigma}_Y^2 = \frac{1}{n}\Big[\Big[\sum_{i=1}^{n}(\bar{x} - x_i)^2\Big]\Big[\sum_{i=1}^{n}(\bar{y} - y_i)^2\Big]\Big]^{\frac{1}{2}} \tag{5.41}$$

where $\hat{\sigma}_X^2$ and $\hat{\sigma}_Y^2$ are estimated variances of x(t) and y(t), respectively. Now, we define signal-to-noise ratio as

$$SNR = \text{cov}(X,Y)/\hat{\sigma}_X \hat{\sigma}_Y \tag{5.42}$$

Substituting Eqs. (5.40) and (5.41) into Eq. (5.42) yields

$$SNR = \frac{1}{n-h}\left[\sum_{i=1}^{n-h}(x_i - \bar{x})(y_i - \bar{y})\right]/$$

$$\frac{1}{n}\left[\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right]\right]^{\frac{1}{2}} = \frac{n}{n-h}\,\delta(h) \tag{5.43}$$

For $h=0$, Eq. (5.43) reduces to

$$SNR = \delta(0) \tag{5.44}$$

where $\delta(0)$ is the normalized cross-correlation function for zero lag. Thus, the normalized cross-correlation function equals to the SNR of the cross-correlation process.

Now, we can compare the prefiltering and the standard correlation techniques by comparing the normalized cross-correlation functions for lag zero. As seen from Figure 5.7, for zero lag, the normalized correlation function is 0.127 while it is 0.156 in Figure 5.8. The signal-to-noise ratio improvement is 22%. The number of data samples was 1000.

As the number of data samples increase we would expect less SNR improvement and for $n \to \infty$ there would be no SNR improvement as discussed in the beginning of this chapter. For the sample size 150, and using the same random processes we found that the SNR improvement was 120% which is very significant improvement.

The prefiltering method was tested by cross-correlating the filter outputs while no signal was present, i.e., $s(t) = 0$. The result is shown

in Figure 5.9. As seen from the figure, the correlation function is not significant for any lag. The noise processes used for this test were the same as those used for Figures 5.7 and 5.8. Finally, we conclude that, for small number of data samples, the prefiltering method increases the signal-to-noise ratio for the detection of the first order Markov signals. When the two noise processes (no signal present) the filters do not force the two noise processes to be alike and hence increase the cross-correlation because of it.

In the above simulation we assumed that the two noise processes $n_x(t)$ and $n_y(t)$ are independent. Figure 5.10 shows the normalized cross-correlation function of the two noise processes. Clearly we see from the figure that there is no significant correlation between the two noise processes.

In our analysis we assumed the local oscillators used at the two ends of the interferometer system did not have any phase noises so that after multiplying the signals from the antennas by the local oscillator signals we obtained a white noise process. Moreover we assumed that the noise signals introduced by the receivers were white.

In practice none of the assumptions we stated above holds exactly, First, there is no oscillator which has a perfect phase behavior, i.e., the phase of the oscillator is constant. Second, the noise at each receiver has a component filtered through the first stages of the receiver. Thus, the noise is not exactly white.

Violation of each assumption clearly degrades the performance of the prefiltering method.

Further work may be done to analyze the proposed scheme as it applies to interferometry. A quantitative analysis of the suggested method has to be carried out before it can be applied to a real system.

We used a recursive filter in time domain rather than a filter in the frequency domain, because the former was much simpler to implement. The characteristics of the filter may be changed by simply changing a few parameters in the program. On the other hand, the above approach is difficult to implement when the IF filter bandwidths are large, because the sampling rate increases as the bandwidth of the signal becomes large. High sampling rates create problems for the computer since during a sampling period the computer has to go through all calculations.

Instead of using discrete filter, a continuous filter may be used in real time since there is no data rate problem for this filter. A difficult problem may be that the solution of a nonlinear Riccati equation must be supplied to the filter during signal processing. The parameters of the Riccati equation are determined by the signal and noise processes.

For the signal and measurement models represented by Eqs. (5.14) and (5.1), the continuous Kalman filter equations are simple, however. It can be shown that an optimum estimate of the signal s(t) is given by (Gelb, 1974, p. 123)

$$\frac{d\hat{s}(t)}{dt} = -\beta\hat{s}(t) + k(t)[x(t) - \hat{s}(t)], \quad \hat{s}(0) = \hat{s}_0 \tag{5.45}$$

where

$$k(t) = p(t)/R, \tag{5.46}$$

$$R = E[\epsilon_x^2(t)] , \tag{5.47}$$

$$p(t) = E\{[\hat{s}(t) - s(t)]^2\} , \tag{5.48}$$

$$R = E[n_x^2(t)] . \tag{5.49}$$

p(t) satisfies the nonlinear Riccati equation

$$\frac{dp(t)}{dt} = -2\beta p(t) + Q - k^2(t)R, \quad p(0) = p_0 \tag{5.50}$$

where $Q = E[u^2(t)]$. It can be shown that the steady-state value of p(t) is given by (Gelb, 1974, p. 147)

$$p_\infty = \beta R[1 + (1 + Q/\beta^2 R)^{\frac{1}{2}}] . \tag{5.51}$$

The corresponding estimation equation is

$$\frac{d\hat{s}(t)}{dt} = -\beta\hat{s}(t) + k_\infty[x(t) - \hat{s}(t)] \tag{5.52}$$

Taking the Fourier transform of both sides of Eq. (5.52) yields

$$j\omega\hat{S}(\omega) = -\beta\hat{S}(\omega) + k_\infty[X(\omega) - \hat{S}(\omega)] . \tag{5.53}$$

The above equation can be written as

$$\frac{\hat{S}(\omega)}{X(\omega)} = \frac{k_\infty}{j\omega + (\beta + k_\infty)} . \tag{5.54}$$

This is the transfer function of the corresponding optimal Wiener filter.

The bandwidth of the filter can be obtained from Eq. (5.5.4) by squaring the magnitude and equating it to one-half of the zero-frequency power:

$$\frac{k_\infty^2}{\omega_B^2 + (\beta + k^2)} = \frac{k_\infty^2}{2(\beta + k_\infty)^2} , \tag{5.55}$$

where ω_B is the angular frequency at which the power is half the zero-power. Solving Eq. (5.55) for ω_B yields

$$\omega_B = \beta + k_\infty \ . \tag{5.56}$$

From Eqs. (5.46) and (5.51) we have

$$k_\infty = \beta[1 + (1 + Q/\beta^2 R)^{\frac{1}{2}}] \ . \tag{5.57}$$

Substituting Eq. (5.57) into Eq. (5.56) yields

$$\omega_B = \beta\{1 + [1 + (1 \pm \frac{Q}{\beta^2 R})^{\frac{1}{2}}]\} \ . \tag{5.58}$$

Thus, the bandwidth B is

$$B = \frac{\omega_B}{2\pi} = \frac{\beta}{2\pi} \{1 + [1 + (1 + \frac{Q}{\beta^2 R})^{\frac{1}{2}}]\} \ .$$

VI. BIBLIOGRAPHY

Al'pert, L. Y. 1973. Radio wave propagation and the ionosphere. Vol. 1. New York: Consultants Bureau.

Basart, J. P., G. K. Miley, and B. G. Clark. 1970. Phase measurements with an interferometer baseline of 11.3 km. IEEE Trans. on Antennas and Propagation. Ap-18, No. 3: 375-379.

Booker, H. G. and R. C. Clemmow. 1950. The concept of an angular spectrum of plane waves, and its relation to that of polar diagram and aperture distribution. Proc. IEE (London) 97: 11-17.

Burke, B. F. 1969. Long-baseline interferometry. Physics Today. 22: 54-63.

Carter, W. H. and L. E. Somers. 1976. Coherence theory of a radio telescope. IEEE Trans. on Antennas and Propagation. AP-24, No. 6: 815-819.

Clark, B. G., K. I. Kellermann, C. C. Bare, M. H. Cohen, and D. L. Jauncey, 1968. High-resolution observations of small-diameter radio sources at 18-centimeter wavelength. Astrophysical Journal. 153: 705-714.

Clark, B. G. 1973. The NRAO tape-recorder interferometer system. Proc. IEEE 61, No. 9: 1242-1248.

Clark, T. A. and W. C. Ericson. 1973. Long wavelength. Proc. IEEE. 61, No. 9: 1230-1233.

Cohen, M. H. 1973. Introduction to very-long-baseline interferometry. Proc. IEEE. 61, No. 9: 1192-1197.

Counselman, C. C. III. 1973. Very-long-baseline interferometry techniques applied to problems of geodesy, geophysics, planetary science, astronomy and general relativity. Proc. IEEE. 61, No. 9: 1225-1230.

Cronyn, W. M. 1970. Radio scattering in the interplanetary medium. Ph.D. dissertation, University of Maryland.

Cronyn, W. M. 1972. Interferometer visibility scintillation. Astrophysical Journal. 174: 181-200.

Davenport, W. B. Jr. and W. L. Root. 1958. An Introduction to the theory of random signals and noise. New York: McGraw-Hill Book Co.

Davies, Kenneth. 1969. Ionospheric radio waves. Waltham, Massachusetts: Blaiddell Publishing Co.

Fomalont, E. B. 1973. Earth-rotation aperture synthesis. Proc. IEEE. 61, No. 9: 1211-1218.

Fomalont, E. B. and C. H. Wright. 1974. Interferometry and aperture synthesis. Pages 256-290 in G. L. Verschuur and K. I. Kellermann, eds. Galactic and extragalactic radio astronomy. New York: Springer Verlag.

Fuller, W. A. 1976. Introduction to statistical time series. New York: John Wiley & Sons, Inc.

Gelb, Arthur, ed. 1974. Applied optimal estimation and design. Cambridge, Mass.: M.I.T. Press.

Hagfors, T. 1976. The ionosphere. Pages 119-134 in M. L. Meeks, ed. Methods of experimental Physics. Vol. 12, part B. New York: Academic Press, Inc.

Hays, W. L. and R. L. Winkler. 1970. Statistics: probability, inference, and decision. Vol. 1. New York: Holt, Rinehart and Winston, Inc.

Jenkins, G. M. and D. G. Watts. 1968. Spectral analysis and its application. San Francisco: Holden-Day.

Ko, H. C. 1967. Coherence theory of radio-astronomical measurements. IEEE Trans. on Antennas and Propagation. AP-15, No. 1: 10-20.

Kraus, J. D. 1966. Radio astronomy. New York: McGraw-Hill Book Co.

Lawrence, R. S., G. G. Little, and H. J. A. Chievers. 1963. A survey of ionospheric effects upon earth-space radio propagation. Proc. IEEE. 52, No. L: 4-28.

Mathur, N. C., M. D. Grossi, and M. R. Pearlman. 1970. Atmospheric effects in very-long-baseline interferometry. Radio Sci. 5, No. 10: 1253-1261.

Meeks, M. L., ed. 1976. Methods of experimental physics. Vol. 12, part C. New York: Academic Press, Inc.

Meyer, S. L. 1975. Data analysis for scientists and engineers. New York: John Wiley & Sons, Inc.

Moran, J. M. 1973. Spectral-line analysis of very-lone-baseline interferometric data. Proc. IEEE. 61, No. 9: 1236-1242.

Moran, J. M. 1976. Very-long-baseline interferometric observations and data reduction. Pages 228-256 in M. L. Meeks, ed. Methods of experimental physics. Vol. 12, part C. New York: Academic Press, Inc.

Mutel, R. L. 1975. Theory and observations of interferometer visibility scintillations with application to the interplanetary medium. Ph.D. dissertation, The University of Colorado.

Mutel, R. L., J. J. Broderic, T. D. Carr, M. Lynch, W. W. Warnock, and W. K. Klemperer. 1974. VLBI observations of the Crab nebula and the wavelength dependence of the interstellar scattering. Astrophysical Journal. No. 193: 279-282.

Papoulis, Athanasios. 1965. Probability, random variables, and stochastic processes. Tokyo, Japan: Kogakusha Co., Ltd.

Oppenheim, A. V. and R. W. Schafer. 1975. Digital signal processing. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Rogers, A. E. E. 1970. Very-long-baseline interferometry with large effective bandwidth for phase-delay measurements. Radio Sci. 5, No. 10: 1239-1247.

Rogers, A. E. E. 1976. Theory of two-element interferometers. Pages 139-153 in M. L. Meeks, ed. Methods of experimental physics. Vol. 12, part C. New York: Academic Press, Inc.

Rönnäng, Bernt. 1971. Report on the VLBI techniques, and data processing of very-long-baseline-interferometry. Gothenburg, Sweden: Chalmers University of Technology.

Sage, A. P. and Melsa, J. L. 1971. Estimation theory with applications to communications and control. New York: McGraw-Hill Book Co.

Salzberg, I. M. 1973. Tracking the apollo lunar rover with interferometry techniques. Proc. IEEE. 61, No. 9: 1230-1233.

Schultz, D. G. and J. L. Melsa. 1967. State functions and linear control systems. New York: McGraw-Hill Book Co.

Schwartz, M., and L. Shaw. 1975. Signal Processing, discrete spectral analysis, detection and estimation. New York: McGraw-Hill Book Co.

Shapiro, I. I. 1976. Estimation of astrometric and geodetic parameters. Pages 261-266 in M. L. Meeks, ed. Methods of experimental Physics. New York: Academic Press, Inc.

Stannard, K. M., G. A. Dulk, and B. Rayhrer. 1970. Very-long-baseline interferometry of decametric radiation from Jupiter. Radio Sci. 5, No. 10: 1271-1280.

Whitney, A. R. 1974. Precision geodesy and astrometry via very-long-baseline interferometry. Ph.D. dissertation, Massachusetts Institute of Technology.

Whitney, A. R., A. E. E. Rogers, H. F. Hinteregger, C. A. Knight, J. I. Levine, S. Lippincott, T. A. Clark, I. I. Shapiro, and D. S. Robertson. 1976. A very-long-baseline interferometer for geodetic applications. Radio Sci. 11, No. 5: 421-432.

VII. ACKNOWLEDGMENTS

First of all I would like to express my gratitude to my major professor Dr. J. P. Basart for his valuable guidance throughout this work and providing the VLBI data.

I would like to thank Dr. W. A. Fuller for extended discussions of the statistical aspects of this work.

My thanks extends to Dr. W. Q. Meeker, Jr. for providing help in the use of his TSERIES program, and my committee members for their valuable times.

I also acknowledge the Turkish Government and the Electrical Engineering Department of I.S.U. for providing financial support.

I thank Betty Carter for her patience in typing the manuscript.

Finally, I extend my deepest appreciation to my wife Serife for her patience and understanding during my graduate work at I.S.U.

VIII. APPENDIX A. STATISTICAL ESTIMATION THEORY

Stochastic processes: Let (Σ,P) be a probability space and let R be a set of real numbers. Then a stochastic process is a real valued function $X(t,\eta)$ defined on $R \times \Sigma$ such that for each $t \in R$, $X(t,\eta)$ is a random variable on (Σ,P), where Σ is a set of possible outcomes for $X(t,\eta)$, η is an elementary event, and P is the set $[0,1]$.

For a fixed η, $X(t,\eta)$ is called a realization of the stochastic process. If t takes continuous values, the process is called a continuous stochastic process, and if t takes discrete values, then the process is called a discrete stochastic process or time series.

The collection of all possible realizations is called the ensemble of realizations. Hereafter we will omit η in $X(t,\eta)$, and represent a realization from a stochastic process by X(t).

In practice, we do not have the ensemble of a stochastic process, but we have a single realization from the ensemble. Therefore, the discussion will be directed toward obtaining the basic properties of a stochastic process from a single realization of this random process. We will not attempt to treat the subject in detail since it is not the purpose of this thesis. But we will give some results which have been used in the early chapters.

A stochastic process is defined completely if the joint distribution function

$$F(X_{t_1},X_{t_2},\ldots,X_t) = P\{\eta: X_{t_1} \leq x_{t_1}, \ldots, X(t_N) \leq x_{t_N}\} \qquad (8.1)$$

is known. When the equality below

$$F(X_{t_1},X_{t_2},\ldots,X_{t_N}) = F(X_{t_1+h},X_{t_2+h},\ldots,X_{t_N+h}) \tag{8.2}$$

holds then the process is said to be stationary. Eq. (8.2) simply means that the joint distribution function depends on only the distance between elements.

For a stationary stochastic process the covariance of X_t and X_{t+h} is only a function of h, the distance between the two elements. Thus,

$$\mathrm{Cov}\{X_t,X_{t+h}\} = E[(X_t - \mu_X)(X_{t+h} - \mu_X)] = \Gamma(h) \tag{8.3}$$

for $t\in R$. Where $\mu_X = E[X(t)]$. $E[\cdot]$ denotes ensemble expected value operation. For a stationary process the expected values of the random variables $X_{t_1},X_{t_2},\ldots,X_{t_N}$ are all the same. Also the variance defined by

$$\mathrm{Var}[X_t] = E[(X_t - \mu_X)^2] \tag{8.4}$$

is constant for $t\in R$. Where $\mu_X = E[X(t)]$.

To deduce the properties of a process we can use the one realization if the process is ergodic. An ergodic process is necessarily stationary, but the inverse is not true. If time averages of a realization equal the ensemble averages of the process from which the particular realization came, then, the process is said to be ergodic. The subject about how a process is found to be ergodic will not be treated here. For a continuous ergodic process we have the following properties:

$$\text{a)}\quad E[X_t] = \lim_{T\to\infty} \frac{1}{T}\int_0^T X(t)dt \tag{8.5}$$

$$\text{b)}\quad \Gamma(h) = E[X_t X_{t+h}] = \lim_{T\to\infty} \frac{1}{T}\int_0^T X(t)X(t+h)dt \tag{8.6}$$

where we assumed that

$$E[X(t)] = 0 \tag{8.7}$$

One of the most important time series is the white-noise process. For such a process

$$\begin{aligned} \Gamma(h) = E[w_t w_{t+h}] &= \sigma^2 \qquad h = 0 \\ &= 0 \qquad \text{otherwise} \end{aligned} \tag{8.8}$$

We assume that $E[w(t)] = 0$. Eq. (8.7) implies that the elements of a white-noise process, w(t), are uncorrelated. Such a process never exists in nature. It is used for mathematical convenience.

A stochastic process usually takes a name according to the type of its joint normal process. The short hand notation for a process is

$$w(t) \sim N(0,\sigma^2) \tag{8.9}$$

meaning w(t) is a zero-mean process with σ^2 variance. The N in Eq. (8.8) denotes a normal or Gaussian process.

Vector stochastic processes: Let $X_1(t), X_2(t), \ldots, X_n(t)$ be a stationary and ergodic time series with means $\mu_1, \mu_2, \ldots, \mu_n$, respectively. Now, consider a column vector:

$$\underline{X}(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_n(t) \end{bmatrix} \tag{8.10}$$

where the underbar denotes a vector. $\underline{X}(t)$ is called a vector stochastic process. The expected value of this vector process is again a vector process and given by

$$\underline{\mu} = E[\underline{X}(t)] = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu_n \end{bmatrix} \tag{8.11}$$

Assuming that $\underline{\mu} = \underline{0}$, we define the covariance matrix of $\underline{X}(t)$ by

$$\underset{\sim}{\Gamma}(h) = E[\underline{X}(t)\underline{X}^T(t+h)] = \{a_{ij}\} \quad i,j = 1, 2, \ldots, n \tag{8.12}$$

where

$$a_{ij} = E[X_i(t)X_j(t+h)] \tag{8.13}$$

The T in Eq. (8.12) denotes the transpose of a matrix and the tilde under $\Gamma(h)$ implies that $\underset{\sim}{\Gamma}(h)$ is a matrix.

Statistical estimation: A detailed discussion of statistical estimation theory will not be given here. The subject is wide and treated by many authors (Fuller, 1976; Jenkins and Watts, 1968). We will only give the results which have been referred to in the previous chapters.

When a single realization from a stochastic process is given, we are usually interested in finding the average properties of the process, e.g., mean, autocorrelation function, spectral density of the process. The definitions of these characteristics are given below:

Mean: The mean or expected value of a stationary process is given by

$$\mu_X = E[X(t)] = \int_{-\infty}^{\infty} X(t)p_{X(t)}(X(t))dX(t) \tag{8.14}$$

where $p_{X(t)}X(t))$ is the probability density function defined by

$$P(-\infty \leq X(t) \leq x(t)) = \int_{-\infty}^{x(t)} p_{X(t)}(X(t))dX(t) \tag{8.15}$$

Autocorrelation function: The autocorrelation function of a stochastic process is given by

$$R(h) = E[X(t)X(t+h)] \tag{8.16}$$

where h is a lag value having the unity of the variable t. When $\mu_X = 0$, then, the autocorrelation function equals the autocovariance function defined by Eq. (8.3).

Cross-correlation function: Let X(t) and Y(t) be two stationary processes with means μ_X and μ_Y, respectively. Then the function

$$R_{XY}(h) = E[X(t)Y(t+h)] \tag{8.17}$$

is called the cross-correlation function of the random processes X(t) and Y(t). The cross-covariance function of X(t) and Y(t) is defined by

$$\Gamma_{XY}(h) = E\{[X(t) - \mu_X][Y(t+h) - \mu_Y]\} \tag{8.18}$$

As pointed out earlier, for an ergodic process the average properties of the process can be deduced from a single realization. Before we continue, it is appropriate to give some definitions. The operation used for obtaining a parameter of a stochastic process from a single realization is called an estimator. For example, m_X obtained from

$$m_X = \frac{1}{N} \sum_{k=1}^{N} x(k) \tag{8.19}$$

is an estimator. Where x(k) is the value of the realization at time t_k. N is the number of samples from a realization.

In general an estimator is defined by

$$\hat{\alpha} = f(x_1, x_2, \ldots, x_n) \tag{8.20}$$

where $\hat{\alpha}$ is the estimate of a parameter of a process, $x_1, x_2, \ldots, x_n$ are the samples from a single realization of the process. The estimate, $\hat{\alpha}$, is also a random variable, having a probability distribution function $p_{\hat{\alpha}}(\hat{\alpha})$. In statistical estimation theory, we seek a function f which has the following properties:

a) The bias defined by

$$B \triangleq \alpha - E[\hat{\alpha}] \tag{8.21}$$

must be zero. Where α is the true value of the parameter to be estimated.

b) The variance

$$\mathrm{Var}[\hat{\alpha}] = E\{[\hat{\alpha} - E(\hat{\alpha})]^2\} = \sigma^2_{\hat{\alpha}} \tag{8.22}$$

must be as small as possible.

c) The estimator must be consistent, i.e., as the number of samples used for estimation increases, then, the variance must go to zero.

If an estimator has zero bias, it is called an unbiased estimator.

In practice, it is often difficult to obtain the probability distributions of the estimators except a few. Now, we will give the estimators

for some parameters of a stochastic process.

Mean: The mean value or expected value of a random process can be found by using

$$\hat{\mu}_X = \frac{1}{N} \sum_{k=1}^{N} x(k) \tag{8.23}$$

where $x(k) = x(t_k)$, N is the number of samples. If X(t), the process from which $\{x(k), k = 1, 2, \ldots, N\}$ comes, is Gaussian, then $\hat{\mu}_X$ is Gaussian. This is not difficult to show. $\hat{\mu}_X$ is simply the weighted sums of normal random variables x(1), x(2), ..., x(N), thus the sum itself is a normal random variable. The bias for Eq. (8.23) is zero, and the variande of $\hat{\mu}_X$ given by

$$\text{Var}\{\hat{\mu}_X\} = \frac{1}{N} \sigma_X^2 \tag{8.24}$$

where σ_X^2 is the variance of the process X(t) (Oppenheim and Schafer, 1975, p. 537). As seen from Eq. (8.24) as N goes to infinite the variance becomes zero, hence, the estimator is a consistent estimator.

Estimation of autocovariance function: Assuming that $E[X] = 0$, we have the estimator

$$\hat{\Gamma}_X(h) = \frac{1}{N - |h|} \sum_{k=1}^{N-|h|} x(k)x(k+h) \tag{8.25}$$

with

$$\text{bias} = \Gamma_X(h) \frac{1}{N} \tag{8.26}$$

For large h, the bias becomes large, and it is a function of $\Gamma_X(h)$, Since we do not know $\Gamma_X(h)$, there is no way of estimating the bias. All we can do is to avoid the large h's. It can be shown that the estimator

$$\hat{\Gamma}_X(h) = \frac{1}{N} \sum_{k=1}^{N-h} x(k)x(k+h) \tag{8.27}$$

has a smaller variance compared to the one given by Eq. (24) (Fuller, 1976, p. 236).

The variance of the estimator given by Eq. (B.24) is given as

$$\text{Var}[\hat{\Gamma}_X(h)] \simeq \frac{N}{[N - |h|]^2} \sum_{r=-\infty}^{\infty} [\Gamma_X^2(h) + \Gamma_X(r+h)\Gamma_X(r-h)] \tag{8.28}$$

$$\lim_{N \to \infty} \text{Var}[\hat{\Gamma}_X(h)] = 0 \tag{8.29}$$

(Oppenheim and Schafer, 1975, p. 540).

Estimation of cross-covariance: The estimator for the cross-covariance function of the processes X(t) and Y(t) is given by

$$\hat{\Gamma}_{XY}(h) = \frac{1}{N} \sum_{k=1}^{N-h} x(k)y(k+h) \tag{8.30}$$

with

$$\text{bias} = E[\hat{\Gamma}_{XY}(h) - \Gamma_{XY}(h)]$$

$$= -\frac{h}{N} \Gamma_{XY}(h)$$

$$\text{Var}\{\hat{\Gamma}_{XY}(h)\} = \frac{1}{N} \sum_{r=-\infty}^{\infty} \Gamma_X(r)\Gamma_Y(r)$$

$$+ \frac{1}{N} \sum_{r=-\infty}^{\infty} \Gamma_{XY}(r+h)\Gamma_{YX}(r-h) \tag{8.31}$$

where $\Gamma_{YX}(h) = E[Y(t)X(t+h)]$, provided that $E[X(t)] = E[Y(t)] = 0$ (Fuller, 1976, p. 263).

A. Maximum Likelihood Estimators

A class of estimators which is often used is the maximum likelihood estimates. A maximum likelihood estimate is based on the following principle. Suppose that a random variable X has a distribution that depends only on some population parameter θ. And suppose that a sample of N independent observations of the random variable X is available. Let $L(X_1, X_2, \ldots, X_N|\theta)$ be the probability density function of this particular sample result given θ. Then, according to the maximum likelihood principle we choose the value of θ for which $L(x_1, x_2, \ldots, x_N|\theta)$ takes its largest value. Clearly, this means that we choose a parameter value which would have made the sample actually obtained have the highest probability.

The maximum likelihood estimator can only be used when the function $L(x_1, x_2, \ldots, x_n|\theta)$ is known. In some cases this function can be found or assumed. As an example, let us suppose that we want to estimate the mean value of a population having a Gaussian probability density function. Assuming that we have a set of samples $[X_1, X_2, \ldots, X_N]$ from the population. Then the density of each X_i is

$$p(X_i|\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(X_i - \mu)^2/2\sigma^2]$$

where μ is the population mean to be estimated, and σ^2 is the variance of the population.

We write the function L as

$$L(X_1, X_2, \ldots, X_N|\mu) = (\frac{1}{\sqrt{2\pi\sigma^2}} \exp[-(X_1 - \mu)^2/2\sigma^2])$$

$$\ldots\ldots\ldots\ldots\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\exp\left[-(X_N-\mu)^2/2\sigma^2\right])$$

$$=\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N\exp\left[-\sum_{i=1}^{N}(X_i-\mu)^2/2\sigma^2\right]$$

It can be shown be shown that the value of μ which maximizes the above function is given by (Hays and Winkler, 1970, p. 321)

$$\hat{\mu}=\frac{1}{N}\sum_{i=1}^{N}X_i\,.$$

IX. APPENDIX B. KALMAN FILTERING

Let us assume that we want to estimate a signal vector $\underline{X}(t)$ from the measurement data corrupted by noise.

Let

$$\underline{X}^T(t) = [x_1(t)\ x_2(t)\ \ldots,\ x_n(t)], \tag{9.1}$$

$$\underline{Z}^T(t) = [Z_1(t)\ Z_2(t)\ \ldots,\ Z_m(t)] \tag{9.2}$$

where $\underline{x}(t)$ is the signal vector, $\underline{Z}(t)$ is the measurement vector, T denotes the transpose operator, m,n are integers. Let $\underline{n}(t)$ be the noise vector of length m, then we write

$$\underline{Z}(t) = \underset{\sim}{S}\,\underline{x}(t) + \underline{n}(t) \tag{9.3}$$

where $\underset{\sim}{S}$ is a $m \times n$ matrix.

Let $\underline{x}(t)$ be modeled as the state vector of a linear dynamic system defined by (Gelb, 1974, p. 52).

$$\dot{\underline{X}}(t) = \underset{\sim}{F}(t)\underline{x}(t) + \underset{\sim}{G}(t)\underline{w}(t) \tag{9.4}$$

where $\underset{\sim}{F}(t)$ = $n \times n$ system matrix,

$\dot{\underline{X}}(t)$ = derivative of the state vector $\underline{X}(t)$ with respect to time,

$\underset{\sim}{G}(t)$ = $n \times n$ system input matrix,

$\underline{w}(t)$ = white input noise vector of length n.

We make the following assumptions:

$$\underline{w}(t) \sim N(\underline{0},\ \underline{Q}) \tag{9.5}$$

$$\underline{n}(t) \sim N(\underline{0},\ \underset{\sim}{R}) \tag{9.6}$$

where

$$\underset{\sim}{Q} = E[\underline{w}(t)\ \underline{w}^T(t)] \tag{9.7}$$

$$\underset{\sim}{R} = E[\underline{n}(t)\ \underline{n}^T(t)] \tag{9.8}$$

where $E[\cdot]$ denotes the expected value operator, $\underline{0}$ is the zero vector.

Suppose we are given $\underline{x}(0)$, $\underset{\sim}{S}$, $\underset{\sim}{F}(t)$, $\underset{\sim}{G}(t)$, $\underset{\sim}{Q}$, and $\underset{\sim}{R}$. Where $\underline{x}(0)$ is the initial condition of the vector $\underline{X}(t)$, i.e.,

$$\underline{X}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \\ \vdots \\ x_n(0) \end{bmatrix} \tag{9.9}$$

Now we ask the question: What is the best way to get an optimum estimate of the signal vector $\underline{X}(t)$ at time $t>0$. By optimum estimate, we mean that the estimate $\hat{\underline{X}}(t)$ must satisfy the following conditions:

$$E\{[\hat{\underline{X}}(t) - \underline{X}(t)][\hat{\underline{X}}(t) - \underline{X}(t)]^T\} = \text{minimum} \tag{9.10}$$

$$E[\hat{\underline{X}}(t) - \underline{X}(t)] = 0 \tag{9.11}$$

By using the so-called continuous Kalman filter, the optimum estimate $\hat{\underline{X}}(t)$ can be found (Sage and Melsa, 1971, p. 283). The discrete version of this estimation procedure is called the discrete Kalman filter. Here, we will only give the algorithm of the discrete Kalman filter. An excellent review of the topic is given by Gelb (1974).

To use the discrete Kalman filter algorithm, we write discrete versions of Eqs. (9.3) - (9.8).

$$\underline{Z}_k = \underset{\sim}{S}_k \underline{x}_k + \underline{n}_k \tag{9.12}$$

$$\underline{X}_k = \underset{\sim}{H}_{k-1} \underline{X}_{k-1} + \underline{w}_{k-1} \tag{9.13}$$

$$\underline{w}_k \sim N(\underline{0}, \underset{\sim}{Q}_k) \tag{9.14}$$

$$\underline{n}_k \sim N(\underline{0}, \underset{\sim}{R}_k) \tag{9.15}$$

where

$$\underline{w}_k = \int_{t_k}^{t_{k+1}} \underset{\sim}{H}(t_{k+1}, \tau) \underset{\sim}{G}(\tau) \underline{w}(\tau) d\tau \tag{9.16}$$

$$\underset{\sim}{H}_k = \underset{\sim}{H}(t_{k+1}, t_k) \tag{9.17}$$

where $\underset{\sim}{H}(t,t_0)$ is called the state transition matrix of the system and defined by

$$\underset{\sim}{H}(t - t_0) = \exp[\underset{\sim}{F}(t)(t - t_0)] \ , \tag{9.18}$$

It is clear that $\underset{\sim}{F}$ must be a square matrix since $\underset{\sim}{F}^2$, $\underset{\sim}{F}^3$, ... can only be defined for square matrices. Eq. (9.18) can be expressed as

$$\underset{\sim}{H}(t,t_0) = \underset{\sim}{I} + (t - t_0) \frac{\underset{\sim}{F}}{1!} + (t - t_0)^2 \frac{\underset{\sim}{F}^2}{2!} + \ldots \tag{9.19}$$

where $\underset{\sim}{I}$ is the unity matrix with the dimensions of matrix $\underset{\sim}{F}$. $\underset{\sim}{R}_k$ and $\underset{\sim}{Q}_k$ are given by

$$\underset{\sim}{R}_k = E[\underline{n}_k \underline{n}_k^T] \ , \tag{9.20}$$

$$\underset{\sim}{Q}_k = E[\underline{w}_k \underline{w}_k^T] \tag{9.21}$$

Next we define the error covariance matrix $\underset{\sim}{P}_k$ as

$$\underset{\sim}{P}(+) = E[\underline{\hat{X}}(+) - \underline{X}_k)(\underline{\hat{X}}_k(+) - \underline{X}_k)^T] \tag{9.22}$$

where $\underset{\sim}{P}_k$ is a n × n matrix, $\hat{\underline{X}}_k(+)$ is the best estimate of the signal vector $\underline{X}_k$ at time t_k, k = 0, 1, 2, ...

If $\underset{\sim}{H}_k$, $\underset{\sim}{S}_k$, $\underset{\sim}{Q}_k$, $\underset{\sim}{R}_k$, $\underset{\sim}{P}_k(-)$, $\hat{\underline{X}}_k(-)$, and $\underline{Z}_k$ are given, $\hat{\underline{X}}_k(+)$ can be found by a recursive algorithm. $\hat{\underline{X}}_k(-)$ is the best predicted value of the signal vector at time t_k, and $\underset{\sim}{P}_k(-)$ is defined by

$$\underset{\sim}{P}_k(-) = E[(\hat{\underline{X}}_k(-) - \underline{X}_k)(\underline{X}_k(-) - \underline{X}_k)^T] \tag{9.23}$$

To obtain the optimum estimate of $\underline{X}_k$ we follow the following steps;

(1) Compute the Kalman gain matrix $\underset{\sim}{K}_k$ by using the equation

$$\underset{\sim}{K}_k = \underset{\sim}{P}_k(-)\underset{\sim}{S}_k^T[\underset{\sim}{S}_k \underset{\sim}{P}_k(-)\underset{\sim}{S}_k^T + \underset{\sim}{R}_k]^{-1} \tag{9.24}$$

(2) Compute the updated state vector $\underline{X}_k(+)$ from

$$\hat{\underline{X}}_k(+) = \hat{\underline{X}}_k(-) + \underset{\sim}{K}_k[\underline{Z}_k - \underset{\sim}{S}_k \hat{\underline{X}}_k(-)] \tag{9.25}$$

(3) Compute the updated error covariance matrix $\underset{\sim}{P}_k(+)$ from

$$\underset{\sim}{P}_k(+) = [\underset{\sim}{I} - \underset{\sim}{K}_k \underset{\sim}{S}_k] \underset{\sim}{P}_k(-) \tag{9.26}$$

(4) Extrapolate the state vector $\underline{X}_k(+)$by using the equation

$$\hat{\underline{X}}_{k+1}(-) = \underset{\sim}{H}_k \hat{\underline{X}}_k(+) \tag{9.27}$$

(5) Extrapolate the error covariance matrix $\underset{\sim}{P}_k(+)$ by the equation

$$\underset{\sim}{P}_{k+1}(-) = \underset{\sim}{H}_k \underset{\sim}{P}_k(+) \underset{\sim}{H}_k^T + \underset{\sim}{Q}_k \tag{9.28}$$

We notice that the 4^{th} and 5^{th} steps provide the initial state vector and error covariance matrix values for the next step.

To start the estimation process we should obtain the initial values $\hat{\underline{x}}_0(-)$ and $\underset{\sim}{P}_0(-)$.

X. APPENDIX C. CALCULATION OF STATE TRANSITION MATRIX

Let $\underset{\sim}{F}(t)$ be the system matrix of a linear dynamic system. Then, the state transition matrix of the system is defined by (Schultz and Melsa, 1967, p. 121)

$$\underset{\sim}{H}(t, t_0) = \exp[\underset{\sim}{F}(t)(t - t_0)] \tag{10.1}$$

where

$H(t,t_0)$ = the state transition matrix,

t_0 = the time origin.

An expansion of Eq. (10.1) was given by Eq. (9.19) in Appendix B as

$$\begin{aligned} \underset{\sim}{H}(t,t_0) &= \underset{\sim}{I} + \underset{\sim}{F}(t)(t - t_0) + \frac{\underset{\sim}{F}^2(t)}{2!}(t - t_0)^2 + \\ &= \sum_{i=0}^{\infty} \frac{\underset{\sim}{F}^i(t)}{i!}(t - t_0)^i \end{aligned} \tag{10.2}$$

where $\underset{\sim}{I}$ is the identity matrix.

The system matrix of the system defined by Eq. (3.C.9) was given by Eq. (3.C.14) as

$$\underset{\sim}{F}(t) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -\alpha^2 & -2\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{10.3}$$

where α and β were defined in Chapter III. Substituting Eq. (10.3) into Eq. (10.2) yields

$$\underset{\sim}{H}(t,t) = \underset{\sim}{I} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -\alpha^2 & -2\beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (t - t_0)$$

$$+ \frac{1}{2!} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -\alpha^2 & -2\beta & 0 & 0 \\ 0 & +2\alpha^2\beta & -\alpha^2+4\beta^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (t - t_0)^2$$

$$+ \frac{1}{3!} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & ab & a+b^2 & 0 & 0 \\ 0 & a(a+b^2) & 2ab+b^3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} (t - t_0)^3$$

$$+ \ldots\ldots\ldots \qquad (10.4)$$

where $a = -\alpha^2$, (10.5)

$b = -2\beta$, (10.6)

and for simplicity we set t_0 to zero. Then, taking the first four terms in Eq. (4) we find

$$\underset{\sim}{H}(t,0) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & h_{22}(t) & h_{23}(t) & 0 & 0 \\ 0 & h_{32}(t) & h_{33}(t) & 0 & 0 \\ 0 & 0 & 0 & 1 & t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{10.7}$$

where

$$h_{22}(t) = 1 + \frac{at^2}{2} + \frac{abt^3}{6} ,$$

$$h_{23}(t) = t + \frac{bt^2}{2} + \frac{(a+b^2)t^3}{6} ,$$

$$h_{32}(t) = at + ab\,\frac{t^2}{2} + \frac{a(a+b^2)t^3}{6} ,$$

$$h_{33}(t) = 1 + bt + (a+b^2)\,\frac{t^3}{2} + \frac{2ab+b^3}{6}\,t^3 ,$$

a and b were defined by Eqs. (10.5) and (10.6).

The discrete version of the state transition matrix is given by (Gelb, 1974, p. 66).

$$\underset{\sim}{H}_k = \underset{\sim}{H}(\Delta t) \tag{10.8}$$

where Δt is the sampling period of the signal in a discrete time-invariant system. Substituting Δt for t in Eq. (10.7)

$$\underset{\sim}{H}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & h_{22}(\Delta t) & h_{23}(\Delta t) & 0 & 0 \\ 0 & h_{32}(\Delta t) & h_{33}(\Delta t) & 0 & 0 \\ 0 & 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{10.9}$$

where $h_{22}(\Delta t) = 1 + \frac{a}{2}(\Delta t)^2 + \frac{ab}{6}(\Delta t)^3$,

$h_{23}(\Delta t) = \Delta t + \frac{b}{2}(\Delta t)^2 + \frac{a+b^2}{6}(\Delta t)^3$,

$h_{32}(\Delta t) = a(\Delta t) + \frac{ab}{2}(\Delta t)^2 + \frac{a(a+b^2)}{6}(\Delta t)^3$,

$h_{33}(\Delta t) = 1+b(\Delta t) + (a+b^2)\frac{(\Delta t)^2}{2} + \frac{2ab+b^3}{6}(\Delta t)^3$.

XI. APPENDIX D. LISTING OF COMPUTER PROGRAMS

```
C       ********************************
C       *                              *
C       *                              *
C       *        SUBROUTINE FOR        *
C       *                              *
C       *    REMOVING  ( 2 * PI )      *
C       *                              *
C       *    A M B I G U I T I E S     *
C       *                              *
C       *                              *
C       ********************************
C       X IS PRINCIPAL PART OF PHASE
C       U IS FRINGE AMPLITUDE
C       AAA=1.
C       XL=LABEL FOR HORIZANTAL AXIS IN  PHASE PLOT
C       YL=LABEL FOR VERTICAL AXIS IN PHASE PLOT
C       GLAB=LABEL FOR GRAPH
C       DLAB=LABEL FOR DATA
        SUBROUTINE CORREC(U,X,AAA,XL,YL,GLAB,DLAB)
C ROW PHASE DATA HAVE AMBIGUITIES OF 2*PI
C THIS PROGRAM CORRECTS ROW PHASE DATA
        REAL AVER
        REAL YF(80)
        INTEGER IX,IY
        REAL ALPHA,BETA
        REAL SUM1,SUM2
        REAL AAA
        REAL  U(80),V(80)
        REAL EST(80),E(80)
        REAL P(80)
        REAL CORR
        REAL AH(22),HH(22),AMIN,AMAX
        REAL X(80),Y(80)
        REAL Z(80)
        REAL TIME(80)
```

```
      REAL XL(5),YL(5),GLAB(5),DLAB(5),XLAB(5),YLAB(5),
     - DATLAB(5),GRLAB(5)
      REAL Z0,Z1,Z2,Z3,Z4,Z5,Z6,Z7,Z8,Z9,Z10,
     - Z11,Z12,Z13,Z14,Z15,Z16,Z17,Z18,Z19,Z20
      REAL W0,W1,W2,W3,W4,W5,W6,W7,W8,W9,W10,
     - W11,W12,W13,W14,W15,W16,W17,W18,W19,W20
      REAL W
      REAL R
      DO 90 I=1,80
90    V(I)=XK*U(I)
      THETA=0.
      AMIN=-6.
      AMAX=6.
      A=1.
      SIGMA=0.0
      WEIGHT=1.
      PI=3.1415927
      Y(1)=X(1)
      DO 80 I=1,79
      Z0=X(I+1)
      Z1=X(I+1)-360.
      Z2=X(I+1)+360.
      Z3=X(I+1)-2*360.
      Z4=X(I+1)+2*360.
      Z5=X(I+1)-3*360.
      Z6=X(I+1)+3*360.
      Z7=X(I+1)-4*360.
      Z8=X(I+1)+4*360.
      Z9=X(I+1)-5*360.
      Z10=X(I+1)+5*360.
      Z11=X(I+1)-6*360.
      Z12=X(I+1)+6*360.
      Z13=X(I+1)-7*360.
      Z14=X(I+1)+7*360.
      Z15=X(I+1)-8*360.
```

```
      Z16=X(I+1)+8*360.
      Z17=X(I+1)-9*360.
      Z18=X(I+1)+9*360.
      Z19=X(I+1)-10*360.
      Z20=X(I+1)+10*360.
      W0=ABS(Y(I)-Z0)
      W1=ABS(Y(I)-Z1)
      W2=ABS(Y(I)-Z2)
      W3=ABS(Y(I)-Z3)
      W4=ABS(Y(I)-Z4)
      W5=ABS(Y(I)-Z5)
      W6=ABS(Y(I)-Z6)
      W7=ABS(Y(I)-Z7)
      W8=ABS(Y(I)-Z8)
      W9=ABS(Y(I)-Z9)
      W10=ABS(Y(I)-Z10)
      W11=ABS(Y(I)-Z11)
      W12=ABS(Y(I)-Z12)
      W13=ABS(Y(I)-Z13)
      W14=ABS(Y(I)-Z14)
      W15=ABS(Y(I)-Z15)
      W16=ABS(Y(I)-Z16)
      W17=ABS(Y(I)-Z17)
      W18=ABS(Y(I)-Z18)
      W19=ABS(Y(I)-Z19)
      W20=ABS(Y(I)-Z20)
      W=MIN1(W0,W1,W2,W3,W4,W5,W6,W7,W9,W10,W11,W12,W13,
     - W14,W15,W16,W17,W18,W19,W20)
      IF(W.EQ.W0) GO TO 50
      IF(W.EQ.W1) GO TO 51
      IF(W.EQ.W2) GO TO 52
      IF(W.EQ.W3) GO TO 53
      IF(W.EQ.W4) GO TO 54
      IF(W.EQ.W5) GO TO 55
      IF(W.EQ.W6) GO TO 56
```

```
      IF(W.EQ.W7) GO TO 57
      IF(W.EQ.W8) GO TO 58
      IF(W.EQ.W9) GO TO 59
      IF(W.EQ.W10) GO TO 60
      IF(W.EQ.W11) GO TO 61
      IF(W.EQ.W12) GO TO 62
      IF(W.EQ.W13) GO TO 63
      IF(W.EQ.W14) GO TO 64
      IF(W.EQ.W15) GO TO 65
      IF(W.EQ.W16) GO TO 66
      IF(W.EQ.W17) GO TO 67
      IF(W.EQ.W18) GO TO 68
      IF(W.EQ.W19) GO TO 69
      IF(W.EQ.W20) GO TO 70
50    Y(I+1)=X(I+1)
      GO TO 80
51    Y(I+1)=X(I+1)-360.
      GO TO 80
52    Y(I+1)=X(I+1)+360.
      GO TO 80
53    Y(I+1)=X(I+1)-2*360.
      GO TO 80
54    Y(I+1)=X(I+1)+2*360.
      GO TO 80
55    Y(I+1)=X(I+1)-3*360.
      GO TO 80
56    Y(I+1)=X(I+1)+3*360.
      GO TO 80
57    Y(I+1)=X(I+1)-4*360.
      GO TO 80
58    Y(I+1)=X(I+1)+4*360.
      GO TO 80
59    Y(I+1)=X(I+1)-5*360.
      GO TO 80
60    Y(I+1)=X(I+1)+5*360.
```

```
      GO TO 80
61    Y(I+1)=X(I+1)-6*360.
      GO TO 80
62    Y(I+1)=X(I+1)+6*360.
      GO TO 80
63    Y(I+1)=X(I+1)-7*360.
      GO TO 80
64    Y(I+1)=X(I+I)+7*360.
      GO TO 80
65    Y(I+1)=X(I+1)-8*360.
      GO TO 80
66    Y(I+1)=X(I+1)+8*360.
      GO TO 80
67    Y(I+1)=X(I+1)-9*360.
      GO TO 80
68    Y(I+1)=X(I+1)+9*360.
      GO TO 80
69    Y(I+1)=X(I+1)-10*360.
      GO TO 80
70    Y(I+1)=X(I+1)+10*360.
80    CONTINUE
      TIME(I)=2.
      TIME(2)=4.
      DO 200 I=3,80
200   TIME(I)=TIME(I-1)+2.
      DO 999 I=1,80
999   Z(I)=X(I)*PI/180.
      DO 300 I=1,80
      P(I)=Y(I)*PI/180.
      WRITE(7,1)TIME(I),V(I),P(I)
1     FORMAT(F4.0,10X,F10.5,10X,F10.5)
300   CONTINUE
      N=80
      ISYM=3
      MODE=4
```

```
      XSIZE=5.
      YSIZE=2.5
      XSF=0.
      XMIN=1.
      YSF=0.
      YMIN=-700.
C PLOT RAW DATA VS TIME
      CALL GRAPH(N,TIME,Z,ISYM,MODE,XSIZE,YSIZE,XSF,XMIN,
     - YSF,YMIN,XL,YL,GLAB,DLAB)
C PLOT CORRECTED PHASE VS TIME
      CALL GRAPH(N,TIME,P,ISYM,MODE,XSIZE,YSIZE,XSF,XMIN,
     - YSF,YMIN,XL,YL,GLAB,DLAB)
      RETURN
      END
```

```
C      ************************
C      *                      *
C      *                      *
C      *    SUBROUTINE  FOR   *
C      *                      *
C      *  PHASE NOISE REMOVAL *
C      *                      *
C      *                      *
C      ************************
       SUBROUTINE FIL(TIME,V,Z,SLO,OMEGA,SIGMA,BETA,VARU,VARSL,XION,
C      THIS SUBROUTINE REMOVES CLOCK AND IONOSPHERIC PHASE NOISES FROM THE
C      INTERFEROMETER FRINGE PHASE
C      T IS AN ARRAY OF TIMES AT WHICH THE PHASE MEASUREMENTS ARE TAKEN
C      V IS AN ARRAY OF FRINGE AMPLITUDES
C      Z IS AN ARRAY OF PHASES
C      SLO IS THE SLOPE OF THE DIFFERENTIAL CLOCK PHASE DRIFT (RAD/SECOND)
C      OMEGA IS THE ANGULAR FREQUENCY OF THE IONOSPHERIC PHASE NOISE
C      PROCESS
C      SIGMA IS THE VARIANCE OF THE IONOSPHERIC PHASE NOISE PROCESS
C      BETA IS THE PARAMETER FOR THE IONOSPHERIC PHASE NOISE PROCESS
C      VARU IS THE VARIANCE OF THE IONOSPHERIC PHASE NOISE PROCESS
C      SIGMA IS THE VARIANCE OF THE STATISTICAL PHASE NOISE PROCESS
C      VARSLO IS THE MEAN SQUARE UNCERTAINITY IN SLO
C      XION IS THE INITIAL VALUE OF THE IONOSPHERIC PHASE NOISE PROCESS
C      XCLOCK IS THE INITIAL VALUE OF THE CLOCK PHASE DRIFT
C      REQUIRED SUBROUTINES: FKAL,GRAPH(ISU PLOTTING SUBROUTINE)
     - XCLOCK)
       DIMENSION V(80),Z(80),TIME(80)
       REAL DIF(80)
       REAL VAR
       REAL CORR
       REAL MEAS(80)
       DIMENSION S(5)
       DIMENSION X(5),Y(1),R(1,1),T3(1)
       DIMENSION H(5,5),G(5,5),        P(5,5),T1(5,5),T2(5,5),Q(5,5)
```

```
      DIMENSION HT(5,5)
      DIMENSION P1(5,5)
      DIMENSION P2(5,5)
      REAL SLO,OMEGA,BETA,VARU,VARSL,XION,XCLOCK
      REAL SIGMA
      REAL XX
      INTEGER IER
      NPTS=80
      MODE=4
      ISYM=7
      XSF=0.
      YSF=0.
      XSIZE=5.
      YSIZE=2.5
      T=2.
      IN=5
      IS=1
      IL=5
      N=5
      M1=1
      L=5
      IT=5
      IER=0
      ALPHA2=OMEGA*OMEGA+BETA*BETA
      A=-ALPHA2
      B=-2*BETA
      ALPHA=SQRT(ALPHA2)
      QX=VARU*T
      DO 10 I=1,5
      DO 20 J=1,5
      Q(I,J)=0.
20    CONTINUE
10    CONTINUE
      Q(2,2)=QX
      Q(2,3)=(ALPHA-2*BETA)*QX
```

```
      Q(3,2)=Q(2,3)
      ALPBET=(ALPHA-2*BETA)*(ALPHA-2*BETA)
      Q(3,3)=ALPBET*QX
      DO 30 I=1,5
      DO 40 J=1,5
      G(I,J)=0.
40    CONTINUE
30    CONTINUE
      G(1,1)=1.
      G(2,2)=1.
      G(3,3)=1.
      G(4,4)=1.
      G(5,5)=1.
      D2=T*T/2.
      D3=D2*T/3.
      DO 50 I=1,5
      DO 60 J=1,5
      H(I,J)=0.
60    CONTINUE
50    CONTINUE
      H(1,1)=1.
      H(2,2)=1.+A*D2+A*B*D3
      H(2,3)=T+B*D2+(A+B*B)*D3
      H(3,2)=A*T+A*B*D2+A*(A+B*B)*D3
      H(3,3)=1.+B*T+(A+B*B)*D2+D3*(2*A*B+B*B*B)
      H(4,4)=1.
      H(4,5)=T
      H(5,5)=1.
      S(1)=1.
      S(2)=1.
      S(3)=0.
      S(4)=1.
      S(5)=0.
C
C
```

```
      X(1)=0.
      X(2)=XION
      X(3)=0.
      X(4)=XCLOCK
      X(5)=SLO
      DO 70 I=1,5
      DO 80 J=1,5
      P(I,J)=0.
80    CONTINUE
70    CONTINUE
      P(1,1)=0.
      P(2,2)=SIGMA
      P(3,3)=10.
      P(4,4)=10.
      P(5,5)=VARSL
      WRITE(6,250)
250   FORMAT('1',T3,'CCODE',T9,'TIME',T25,'NOISY PHASE',T40,
     - 'TRUE PHASE',T56,'ION.PHASE',
     - T72,'CLOCK PHASE',T88,'MSE(TRUE)',T104,'MSE(ION.)',T120,
     - 'MSE(CLOCK)')
      DO 300 I=1,80
      Y(1)=Z(I)
      NTIM=TIME(I)
      CALL FKAL(S,Y,P,X,H,G,Q,SIGMA,NTIM,XX)
      DIF(I)=XX
300   CONTINUE
      RETURN
      END
```

```
C       ************************
C       *                      *
C       *                      *
C       *     SUBROUTINE FOR   *
C       *                      *
C       *     KALMAN  FILTER   *
C       *                      *
C       *                      *
C       ************************
        SUBROUTINE FKAL(S,Y,P,X,H,G,Q,SIGMA,NTIM,XX)
C SIGNAL MODEL: X(K)=H*X(K-1)+G*U(K-1)
C MEASUREMENT MODEL: Y(K)=S*X(K)+V(K)
C WHERE U IS THE INPUT NOISE PROCESS
C V IS THE MEASUREMENT NOISE PROCESS
C COMPUTES ESTIMATED STATE VECTOR AND ERROR COVARIANCE MATRIX
C COMPUTES PROJECTED STATE VECTOR AND COVARIANCE MATRIX OF THE ERROR
C S: INPUT , MEASUREMENT ARRAY
C Y : MEASUREMENT , INPUT
C P: ERROR COVARIANCE MATRIX,INPUT
C X: STATE VECTOR,AT THE FIRST CALL HAS TO BE SPECIFIED
C H: STATE TRANSITION MATRIX,INPUT
C G: UNITY MATRIX
C Q: COVARIANCE MATRIX OF THE INPUT NOISE
C SIGMA: COVARIANCE OF THE MEASUREMENT NOISE
C  REQUIRED SUBROUTINES:ADD55,TRANS,M51,M55
        DIMENSION S(5),Y(1),P(5,5),X(5),G(5,5),P1(5),H(5,5),Q(5,5)
        DIMENSION P4(5,5),P5(5,5),P6(5,5)
        DIMENSION XP(5),HT(585),P9(5,5),P10(5,5),P11(5,5)
        REAL XX
        REAL P2,P3,K(5),P7,P8
        IER=0
C COMPUTE 1=P*S(TR)
        CALL M51(P,S,P1)
C COMPUTE P2=S*P*S(TR)
        P2=0.
```

```
      DO 10 I=1,5
10    P2=P2+S(I)*P1(I)
C COMPUTE P3=S*P*S(TR)+SIGMA
      P3=P2+SIGMA
C COMPUTE THE GAIN
      DO 20 I=1,5
20    K(I)=P1(I)/P3
C COMPUTE P4=K*H
      DO 40 I=1,5
      DO 30 J=1,5
      P4(I,J)=K(I)*S(J)
30    CONTINUE
40    CONTINUE
C COMPUTE P5=I-K*S
      DO 50 I=1,5
      DO 50 J=1,5
50    P5(I,J)=G(I,J)-P4(I,J)
C UPDATE THE COVARIANCE
      CALL M55(P5,P,P6)
      DO 60 I=1,5
      DO 60 J=1,5
60    P(I,J)=P6(I,J)
C UPDATE THE STATE
      P7=0.
      DO 70 I=1,5
70    P7=P7+S(I)*X(I)
      P8=Y(1)-P7
      DO 80 I=1,5
80    X(I)=X(I)+K(I)*P8
      XX=Y(1)-X(2)-X(4)
C
C PROJECT X AHEAD
      CALL  MS1(H,X,XP)
      DO 90 I=1,5
90    X(I)=XP(I)
```

```
C PROJECT THE COVARIANCE AHEAD
        CALL TRANS(H,HT)
        CALL M55(P,HT,P9)
        CALL M55(H,P9,P10)
        CALL ADD55(P10,Q,P11)
        DO 100 I=1,5
        DO 100 J=1,5
100     P(I,J)=P11(I,J)
        RETURN
        END
       SUBROUTINE ADD55(A,B,C)
C       THIS SUBROUTINE ADDS TWO 55 MATRICES
C       A IS A 55 INPUT MATRIX
C       B IS A 55 INPUT MATRIX
C       C IS A 55 OUTPUT MATRIX
       DIMENSION A(5,5),B(5,5),C(5,5)
       DO 1 I=1,5
       DO 2 J=1,5
     2 C(I,J)=A(I,J)+B(I,J)
     1 CONTINUE
       RETURN
       END
       SUBROUTINE TRANS(A,B)
C       THIS SUBROUTINE CALCULATES TRANSPOSE OF A 55 MATRIX
C       A IS A 55 INPUT MATRIX
C       B IS A 55 OUTPUT MATRIX
       DIMENSION A(5,5),B(5,5)
       DO 1 I=1,5
       DO 2 J=1,5
     2 B(J,I)=A(I,J)
     1 CONTINUE
       RETURN
       END
       SUBROUTINE M51(A,V1,V2)
C   SUBROUTINE M51 MULTIPLIES A 55 MATRIX BY AN ARRAY OF SIZE 5
```

```
C   A IS A 55 INPUT MATRIX
C   V1 IS AN INPUT ARRAY OF LENGTH 5
C   V2 IS AN OUTPUT ARRAY OF LENGTH 5
      DIMENSION A(5,5),V1(5),V2(5),V3(5)
      DO 1 I=1,5
      V3(I)=0.
      DO 2 J=1,5
      V3(I)=V3(I)+A(I,J)*V1(J)
    2 CONTINUE
    1 CONTINUE
      DO 3 I=1,5
    3 V2(I)=V3(I)
      RETURN
      END
      SUBROUTINE M55(A,B,C)
C      THIS SUBROUTINE MULTIPLIES TWO 55 MATRICES
C      A IS 55 INPUT MATRIX
C      B IS 55 INPUT MATRIX
C      C IS 55 OUTPUT MATRIX
      DIMENSION A(5,5),B(5,5),C(5,5),D(5,5)
      DO 1 I=1,5
      DO 2 J=1,5
      D(I,J)=0.
      DO 3 K=1,5
      D(I,J)=D(I,J)+A(I,K)*B(K,J)
    3 CONTINUE
    2 CONTINUE
    1 CONTINUE
      DO 4 I=1,5
      DO 4 J=1,5
    4 C(I,J)=D(I,J)
      RETURN
      END
```