JPL PUBLICATION 80-84

An Analysis of Source Structure Effects in Radio Interferometry Measurements

J. B. Thomas

December 15, 1980

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL PUBLICATION 80-84

An Analysis of Source Structure Effects in Radio Interferometry Measurements

J. B. Thomas

December 15, 1980

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

CONTENTS

Figures

Figures (continued)

ABSTRACT

As the accuracy of the technique of long-baseline radio interferometry improves, the errors contributed by the extended structure of natural sources will become increasingly important. To begin a study of such structure effects, this report presents a theoretical framework, proposes an effective-position approach to structure corrections based on brightness distribution measurements, and analyzes examples of analytical and measured brightness distributions. Other topics include the effect of the frequency dependence of a brightness distribution on bandwidth synthesis (BWS) delay, the determination of the absolute location of a measured brightnese distribution, and structure effects in dual-frequency calibration of charged-particle delays. For the 10 measured distributions analyzed in this report, it is found that the structure effect in BWS delay at X-band (3.6 cm) can reach 30 cm but typically falls in the range of 0-5 cm. The largest BWS delay effects occur at points in the u-v plane with very small fringe visibility. To provide a possible method for reducing and estimating structure effects in BWS delay, a limit approach is investigated through a general analysis and through a study of particular analytical sources. A trial limit equation that is dependent on visibility is successfully tested against the 10 measured brightness distributions (seven sources). If the validity of this particular equation for an upper limit can be established for nearly all sources, the structure effect in BWS delay could be greatly reduced without supplementary measurements of brightness distributions, perhaps to the 2 cm level (1σ) on intercontinental baselines and to less than a centimeter on regional baselines.

I. INTRODUCTION

As the accuracy of the technique of long-baseline radio interferometry improves, the effects of source structure will become more important in measurements of baseline, source positions, and other quantities. A crude method for demonstrating the magnitude of structure effects is to note that changes in source position due to structure are expected to be of the order of the fringe spacing, which is given in radians by the ratio λ/B where λ is the observing wavelength and B is the baseline length. Since the partial of delay with respect to source position is of the order of B in meters/radian, the corresponding effect on delay will be of the order of $(\lambda/B)B = \lambda$. Thus, for observations at X-band ($\lambda = 3.6$ cm) some sources will contribute unacceptable structure errors if centimeter-level accuracy is desired.

In addition to the geophysical applications of VLBI, differential VLBI measurements can be used to measure the separation between close pairs of radio sources (natural-natural and natural-spacecraft). Source structure effects will be even more significant in this application since error sources other than source structure are expected to contribute a total error of only 10^{-3} to 10^{-5} arcseconds in position on intercontinental baselines. This potential accuracy is to be compared with the estimated structure error (λ/B), which is equal to about 10^{-3} arcseconds at X-band with $B = 6000$ km. Thus for differential measurements to reach their full potential, source structure effects must be treated very carefully.

To begin the investigation of source structure effects, this report presents a theoretical framework, proposes an effective-position approach to structure corrections based on brightness distribution measurements, analyzes a few examples of analytical and measured brightness distributions, and develops a limit formula for reducing and estimating structure effects in bandwidth synthesis (BWS) delay. The report is organized in the following way. The next section, Section II, rederives the cross-correlation function by beginning with a description of the noise wave emitted by a natural source, and tracing the signal through the instrumentation and cross-correlation process. The derivation is repeated in detail to make this report self-contained and to emphasize the relation of the parameters of the brightness distribution to the parameters of the cross-correlation function. That relation is needed to demonstrate clearly how the absolute location of an independently measured brightness distribution can be determined through standard VLBI measurements. Section III reformulates the brightness transform to cast it in a form more suitable for the analysis of the effect of structure on VLBI observables. Section IV discusses the determination of structure effects in VLBI observables (BWS delay, delay rate, phase delay) and introduces the concept of effective position. Sections V and VI consider the structure effects associated with bandwidth synthesis (BWS) delay and with delay rate, respectively. Section VII combines the results of Sections V and VI to derive the effective position associated with BWS delay and delay rate. Section VIII considers the structure effects associated with the phase-delay observable. Section IX shows how to correct delay and delay rate for structure effects and in the process demonstrates a technique for determining the absolute location for the origin of the coordinates with which the brightness distribution is specified. Section X considers differential measurements of the relative positions of two natural sources and notes that such measurements might be used as a direct and accurate test of the procedures used to correct structure effects. The subsequent four sections present

examples of brightness distributions to illustrate the general analysis. Specifically, Sections XI, XII and XIII treat the analytical examples of a symmetric distribution, a double-point source and an N-point source (with emphasis on triple point), respectively, while Section XIV presents results for actual natural sources. Finally, Section XV derives and tests a limit formula for structure effects in BWS delay.

II. THE CROSS-CORRELATION FUNCTION

This section presents a derivation of the cross-correlation function for an extended natural source that is completely incoherent. The derivation begins with a description of the noise wave emitted by the source and traces the signal through the instrumentation and cross-correlation process. Considerable detail is included so that the parameters of the brightness distribution are precisely defined in relation to the cross-correlation function. Such care is required to show that it is possible to use the standard multiparameter fits to delay to determine the absolute location of the unknown origin of an independently measured brightness distribution. As the following derivation shows, a distinction must be made between the actual effective position, the origin of structure coordinates and an assigned reference position (see Section IV).

The radio noise generated by a very distant extended natural source can be expressed as a superposition of plane waves integrated over possible wave directions and frequencies:

$$E(\vec{x},t) = \int_{\hat{k}} \int_{0}^{\infty} A(\hat{k},\omega) e^{i\omega(t-\hat{k}\cdot\vec{x}/c)} d\omega d\Omega + \text{c.c.} \qquad (1)$$

where $E(\vec{x},t)$ is the electric field at point $\vec{x}$ and time t, $A(\hat{k},\omega)$ is the random Fourier amplitude at frequency ω for the wave portion received from direction $\hat{k}$, and the letters c.c. denote complex conjugate. The wave is assumed to be linearly polarized for simplicity. All quantities are measured with respect to a quasi-inertial geocentric frame with axes defined by true equatorial coordinates of date. The wave direction $\hat{k}$, which includes annual aberration, must be expressed as a function of two parameters. In terms of right ascension and declination, the wave direction is given

$$\hat{k} = -(\cos\delta\cos\alpha,\ \cos\delta\sin\alpha,\ \sin\delta) \qquad (2)$$

where α,δ are apparent right ascension and declination relative to true equatorial coordinates of date. The quantity $d\Omega$ represents a differential solid angle for integration over the possible wave directions and is given by $\cos\delta d\alpha d\delta$ in the case of right ascension and declination. The term $\hat{k}$ in the argument of the Fourier amplitude stands for the two direction parameters.

The electric field detected at antenna j is given by

$$E_j(t) = E(\vec{x}_j(t), t) \tag{3}$$

$$= \int_{\hat{k}} \int_0^\infty A(\hat{k},\omega) e^{i\omega[t-\hat{k}\cdot\vec{x}_j(t)/c]} d\omega d\Omega + c.c. \tag{4}$$

where $\vec{x}_j(t)$ is the location of antenna j at time t. The voltage signal recorded at antenna j is given by the expression (Ref. 1)

$$V_j(t) = \int_{\hat{k}} \int_0^\infty A(\hat{k},\omega) G_j \, e^{i\omega[t-\hat{k}\cdot\vec{x}_j(t)/c)]-\psi_j(t)} d\omega d\Omega + c.c. + n_j(t) \tag{5}$$

where G_j is the effective system bandpass filter and n_j is random instrumental noise. The phase ψ_j represents all instrumental phase effects and includes the heterodyne phase, cable delays and bandpass phase shifts. For simplicity, we will neglect transmission media effects since they will only add more terms to the phase and will be of no significance in the derivation of the cross-correlation function. We have assumed that the antenna pattern is large compared to the source size and can therefore be neglected.

In the following, we will neglect two-level sampling and treat only the analog signals. As is well-known, the digital result is the same as the analog result in the case of weak signals, except for a loss of $2/\pi$ in SNR. Further, we will neglect clock synchronization errors and will not treat delay offsetting of both signals. These considerations modify the phase only slightly, can be readily analyzed and corrected with adequate a priori information, and have no ultimate effect on the brightness transform derivation.

In the cross-correlation procedure, the signals from the two antennas are offset by a model delay and multiplied together, which gives

$$V_i(t)V_j(t+\tau_m) = \int_{\hat{k}'} \int_0^\infty \int_{\hat{k}} \int_0^\infty A(\hat{k},\omega)A^*(\hat{k}',\omega')G_iG_j \, e^{i\phi} \, d\omega d\Omega d\omega' d\Omega' \tag{6}$$

$$+ \int_{\hat{k}'} \int_0^\infty \int_{\hat{k}} \int_0^\infty A(\hat{k},\omega)A(\hat{k}',\omega')G_iG_j \, e^{i\phi_c} d\omega d\Omega d\omega' d\Omega'$$

$$+ c.\,c. + \text{noise terms}$$

where

$$\phi = (\omega - \omega')t + \omega'\hat{k}'\cdot\vec{x}_j(t + \tau_m)/c - \omega\hat{k}\cdot\vec{x}_i(t)/c + \psi_j(t + \tau_m) - \psi_i(t) - \omega'\tau_m \qquad (7)$$

and ϕ_c is a similar expression that will not be needed. To obtain the mean value for the voltage product, one can compute the ensemble average:

$$\langle V_i(t)V_j(t + \tau_m)\rangle =$$

$$\int_{\hat{k}'}\int_0^\infty\int_{\hat{k}}\int_0^\infty \langle A(\hat{k},\omega)A^*(\hat{k}',\omega')\rangle G_iG_j e^{i\phi} d\omega d\Omega d\omega' d\Omega' \qquad (8)$$

$$+ \int_{\hat{k}'}\int_0^\infty\int_{\hat{k}}\int_0^\infty \langle A(\hat{k},\omega)A(k',\omega')\rangle G_iG_j e^{i\phi_c} d\omega d\Omega d\omega' d\Omega'$$

$$+ \quad c.c.$$

Since the instrumental noise is uncorrelated between antennas and is uncorrelated with the radio noise, all terms involving the additive instrumental noise have averaged to zero. We will assume that the source is completely incoherent so that the source signal from one part of the source ($\hat{k}$) is completely uncorrelated with the signal from another part of the source ($\hat{k}'$). Further, we will assume that a signal component emitted at one frequency (ω) is completely uncorrelated with a signal component emitted at another frequency (ω'). These assumptions can be represented mathematically by a condition on the ensemble average of the component product:

$$\langle A(\hat{k},\omega)A^*(k',\omega')\rangle = \frac{1}{2} D(\hat{k},\omega)\delta(\hat{k} - \hat{k}')\delta(\omega - \omega') \qquad (9)$$

where D is the brightness distribution at frequency ω and $\delta(z)$ represents a Dirac delta function. (The delta function $\delta(\hat{k} - \hat{k}')$ must be parameterized in a manner consistent with the differential solid angle $d\Omega$, but this detail will be transparent to us).

Since the signal is real, we know

$$A(\hat{k},\omega) = A^*(\hat{k}, -\omega) \tag{10}$$

so that

$$\langle A(\hat{k},\omega)\, A(\hat{k}',\omega')\rangle = \tfrac{1}{2} D(\hat{k},\omega)\, \delta(\hat{k} - \hat{k}')\, \delta(\omega + \omega') \tag{11}$$

$$= 0 \text{ for } \omega,\ \omega' > 0 \tag{12}$$

When these conditions are substituted in Eq. (8) we obtain

$$\langle V_i(t) V_j(t + \tau_m)\rangle = \tfrac{1}{2} \int_{\hat{k}} \int_0^{\infty} D(\hat{k},\omega)\, G_i G_j\, e^{i\phi_f}\, d\omega d\Omega + c.\ c. \tag{13}$$

where the phase is given by

$$\phi_f = \omega \hat{k} \cdot \vec{B}_r / c - \psi_i(t) + \psi_j(t + \tau_m) - \omega\tau_m \tag{14}$$

for which the retarded baseline is defined by

$$\vec{B}_r = \vec{x}_j\,(t + \tau_m) - \vec{x}_i(t) \tag{15}$$

The second term in Eq. (8) has dropped out since $\omega = -\omega'$ is not covered in the region of integration.

To cast this result in more familiar form, we must define special direction parameters for $\hat{k}$. The most common approach is to define two new coordinates to replace right ascension and declination (α, δ), where the new coordinates are a linear representation of small angular displacements relative to some point near or within the source. These coordinates will be called structure coordinates and are defined by the expressions

$$\beta \equiv (\alpha - \alpha_r) \cos \delta_r \tag{16}$$

$$\gamma \equiv \delta - \delta_r \tag{17}$$

where (α_r, δ_r) is the location of the origin of structure coordinates in terms of right ascension and declination. Usually when VLBI measurements of structure are made, it is necessary to make some relatively arbitrary assignment of the origin of structure coordinates (e.g., the center of the brightest component). Further, the absolute location of this origin is not accurately known in (α, δ) coordinates. For the moment we will proceed as though this origin uncertainty did not exist and will leave discussion of its determination for Section IX. With the definition of structure coordinates in Eqs. (16) and (17), the differential solid angle becomes

$$d\Omega = \cos \delta_r \, d\alpha \, d\delta \tag{18}$$

$$= d\beta d\gamma \tag{19}$$

provided the displacements (β, γ) from the origin are small.

Up to this point, no precise definition of source location has been made, since the preceding definition only established a local coordinate frame. In Eq. (13), the positional effects of the source enter the expression as an integral over wave direction and no particular direction is the source "position." To proceed with the analysis, it is useful to pick some reference direction so that a small angle approximation can be made for the wave direction $\hat{k}$. Suppose some adopted point (β_0, γ_0) within the source has been chosen as a reference point with a reference direction given by

$$\hat{k}_0 = \hat{k}(\beta_0, \gamma_0) \tag{20}$$

If the distribution is sufficiently narrow, we can approximate the wave direction by

$$\hat{k} = \hat{k}_0 + \frac{\partial \hat{k}}{\partial \beta}\bigg|_0 (\beta - \beta_0) + \frac{\partial \hat{k}}{\partial \gamma}\bigg|_0 (\gamma - \gamma_0) \tag{21}$$

where the partials are evaluated at (β_0, γ_0). It is important to note at this point that the assignment of a particular point in the source as a reference point is quite arbitrary. Further, it should be emphasized that the point (β_0, γ_0) is not necessarily the same point as the origin of structure coordinates (β_r, γ_r) as is often implicitly assumed. Using the last relation, we obtain for the cross-correlation function:

$$\langle V_i(t) V_j(t + \tau_m) \rangle = \frac{1}{2} \int_0^\infty R(u,v,\omega) G_i G_j e^{i\phi_a} \, d\omega + \text{c.c.} \tag{22}$$

where the phase is given by

$$\phi_a = \omega \hat{k}_0 \cdot \vec{B}_r/c - \psi_i(t) + \psi_j(t + \tau_m) - \omega \tau_m \tag{23}$$

In this expression, the brightness transform is defined as

$$R(u,v,\omega) \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(\beta,\gamma,\omega)\, e^{-2\pi i[u(\beta-\beta_0)+v(\gamma-\gamma_0)]}\, d\beta d\gamma \tag{24}$$

where

$$u \equiv -\left.\frac{\partial \hat{k}}{\partial \beta}\right|_0 \cdot \vec{B}_r/\lambda \tag{25}$$

$$v \equiv -\left.\frac{\partial \hat{k}}{\partial \gamma}\right|_0 \cdot \vec{B}_r/\lambda \tag{26}$$

Note that the limits of the (β, γ) integration have been extended to infinity under the assumption that the narrowness of the brightness distribution terminates the integration.

Using the standard definition of the geometric delay, we can write

$$\tau_g \approx \frac{\hat{k}_0 \cdot \vec{B}_r}{c} \approx \frac{\hat{k}_0 \cdot \vec{B}}{c} \left(1 - \frac{\hat{k}_0 \cdot \vec{v}_j}{c}\right)^{-1} \tag{27}$$

where τ_g is the geometric delay for the direction $\hat{k}_0$ and where $\vec{v}_j$ is the velocity of station j. The phase of the cross-correlation function becomes

$$\phi_a = \omega(\tau_g - \tau_m) - \psi_i(t) + \psi_j(t + \tau_m) \tag{28}$$

In the standard VLBI measurements, the first term $\omega\tau_g$ is the quantity of interest and the others are removed in processing (by phase calibration, etc.) to the

extent allowed by the accuracy of the calibration techniques and a priori knowledge. (Had we been more thorough and considered clock synchronization, an additional unknown, τ_c, would have been added to τ_g. Such a term is present in actual VLBI experiments and its measurement is of interest to time-and-frequency specialists.) After such corrections, the measurement of delay is still corrupted by the implicit phase of the complex brightness transform R in Eq. (22). It is the purpose of this report to investigate this phase effect of structure.

III. REFORMULATION OF THE BRIGHTNESS TRANSFORM

This section reformulates the standard brightness transform in Eq. (24) to place it in a form more suitable for subsequent analysis. We will neglect the fact that the fringes are actually integrated over a narrow bandwidth (see Eq. 22) and will place the observing frequency ω at the center of the channel passband (2 MHz for the BLKO system). This will be a good approximation as long as the brightness distribution changes very slowly as a function of frequency over the channel passband. Note that a frequency change of 2 MHz would be only a change of one part in a thousand at S-band (2.3 GHz). Given the earlier definitions in Eqs. (16) and (17), one can easily show that the vectors $-\partial\hat{k}/\partial\beta$ and $-\partial\hat{k}/\partial\gamma$ are orthogonal unit vectors in the direction of increasing right ascension and declination, respectively, and form a set of basis vectors in the plane of the sky at the source. We will call these unit vectors $\hat{i}$ and $\hat{j}$, respectively, and will make them the basis vectors for the (β, γ) coordinates. Note that the basis vectors at the origin (α_r, δ_r) used to define (β, γ) coordinates in Eqs. (16) and (17) will not necessarily be exactly the same as those at (β_0, γ_0). However, we will assume that these two points are very close and the slight differences in basis vectors are negligible. The brightness transform now becomes

$$R = e^{+i\omega\vec{B}_s\cdot\vec{P}_0/c}\iint D(\vec{P},\omega,t)e^{-i\omega\vec{B}_s\cdot\vec{P}/c}\,d\Omega \tag{29}$$

where $d\Omega$ is the area differential and

$$\vec{P} \equiv \beta\hat{i} + \gamma\hat{j} \tag{30}$$

$$\vec{P}_0 \equiv \beta_0\hat{i} + \gamma_0\hat{j} \tag{31}$$

$$\vec{B}_s \equiv B_\beta\hat{i} + B_\gamma\hat{j} \tag{32}$$

A time variable has been added to the argument of the brightness distribution to allow for time variability. The vector $\vec{B}_s$ is obtained by projecting $\vec{B}$ onto the plane of the sky. (Though it is not necessary to do so, we have dropped the out-of-plane component of $\vec{B}$ to emphasize that the only in-plane component is important as a result of the dot product with $\vec{P} - \vec{P}_0$.) We have extracted the multiplicative phasor from the integral since it does not depend on the integration variables.

The phase change (structure phase) contributed to fringe phase by the brightness transform will be given by

$$\phi_B = \tan^{-1}(-Z_s/Z_c) + \omega\vec{B}_s \cdot \vec{P}_0/c \tag{33}$$

where

$$Z_c \equiv \iint D(\vec{P},\omega,t)\ \cos(\omega\vec{B}_s \cdot \vec{P}/c)d\Omega \tag{34}$$

$$Z_s \equiv \iint D(\vec{P},\omega,t)\ \sin(\omega\vec{B}_s \cdot \vec{P}/c)d\Omega \tag{35}$$

The integer-cycle ambiguity problem encountered in the computation of structure phase involves some subleties that should be mentioned. As discussed in the next section, the observables of BWS delay and delay rate are approximately given by derivatives of phase and therefore their structure corrections are unaffected by any constant offsets in structure phase caused by ambiguities. However, when fringe phase (phase delay) is the observable, ambiguities are more of a problem. If one attempts to define a unique value for structure phase for each value of (u, v), one will find it is not always possible due to the fact that the integer part of structure phase can depend on the path followed in the u-v plane to reach a given value of (u, v), provided one requires continuity in ϕ_B along each (u, v) track. This apparent problem is really of no consequence. For purposes of correcting fringe phase in the usual VLBI applications (see Section IX and X), it is not necessary to specify the integer part of structure phase since only the fractional part is needed to remove structure effects. After the subtraction of a fractional structure phase, only the corrected fringe phase has to be subjected to an ambiguity resolution process such as bandwidth synthesis. For this reason, example computations in general u-v plots will constrain structure phase to the range ±0.5 cycle. However, for display of results along a given u-v track, integer cycles can be added to maintain continuity in structure phase. (If continuity is maintained in both corrected and uncorrected fringe phase along a u-v track, as is usually the case, then the procedure for structure corrections outlined above will implicitly enforce continuity on structure phase along the track.)

IV. VLBI OBSERVABLES AND EFFECTIVE POSITION

With regard to geophysical/astrometric measurements, the primary observable in VLBI is fringe phase. As indicated in Section II, all terms but $\omega\tau_g$, a clock term, and structure phase are removed from the phase (Eq. 28) to the accuracy allowed by phase calibration, model restoration, etc. For simplicity, we will assume that all of these terms have been exactly removed so that fringe phase will be equal to a geometric delay term plus structure phase

$$\phi_T = \omega(\tau_g + \tau_c) + \phi_B \tag{36}$$

where we have now included the clock term τ_c. As indicated in Section II, τ_g is the geometric delay relative to the adopted reference point $\vec{P}_0$ in the source and ϕ_B is the structure phase given by Eq. (33).

Three observable types are derived from fringe phase, each in a different way. In each case, the brightness transform will make a contribution, but the calculation of the contribution must be made for each observable using the particular operation that is applied to fringe phase to obtain that observable. The three observables usually analyzed in interferometry work are BWS delay, phase delay and phase-delay rate. Though it can be formulated as a finite difference or as a fitted phase slope, the BWS delay is approximately equal to the partial derivative of fringe phase with respect to frequency. (For extremely wide spanned bandwidths, the derivative might be an inadequate approximation to the actual BWS delay. We will not analyze that problem in this report.) Phase-delay rate is essentially equal to the partial of fringe phase with respect to time, while phase delay is essentially the direct observation of fringe phase. Accordingly, the computation of the effect of source structure must be calculated as

$$\Delta\tau_B = \frac{\partial\phi_B}{\partial\omega} \qquad \text{for BWS delay} \tag{37}$$

$$\Delta\dot{\tau}_\phi = \frac{1}{\omega}\frac{\partial\phi_B}{\partial t} \qquad \text{for (phase) delay rate} \tag{38}$$

$$\Delta\tau_\phi = \frac{\phi_B}{\omega} \qquad \text{for phase delay} \tag{39}$$

where ϕ_B is computed from Eq. (33).

In general, for all three observables, the size of these contributions will change with baseline orientation and length so that one value for the structure effect cannot be quoted, even for a given baseline for a given day of observations. Further, the contribution to BWS delay will not necessarily be the same as that for phase delay. Note that the structure effect depends on the reference point P_0, which up to this point has not been precisely defined. A direct approach to removing this arbitrariness would be to assign a particular value for $\vec{P}_0$ (e.g., the ordinary centroid) and then directly compute with Eqs. (37)-(39) the structure effects for each observable as a function of hour angle, leaving $\vec{P}_0$ fixed. If these corrections are applied to the observables, the source would effectively become a point source at $\vec{P}_0$. A second, indirect but equivalent approach

is more useful when detailed specification of source position is desired. In this approach, the effective location of the source is explicitly computed as a function of baseline orientation and length so that deviations from a constant position are directly determined. To correct the observables, a reference position must be chosen and the difference in effective position and reference position computed. The observable correction is then computed as the product of this difference vector and the appropriate sensitivity partials relating observable to position. After all observations have been so corrected, the source effectively becomes a point source at the reference position.

The effective position of a source can be determined as follows for a given observation. The effect of structure on our usual observables is given in Eqs. (37), (38) and (39). If a value of $\vec{P}_0$ is chosen that makes the structure effect zero for a given observable, then the observable based on Eq. (36) will come purely from a delay that is equal to the geometric delay produced by a point source at $\vec{P}_0$. Thus the effective position will be the value of $\vec{P}_0$ that satisfies the following conditions:

$$\frac{\partial \phi_B}{\partial \omega} = 0 \tag{40}$$

$$\frac{\partial \phi_B}{\partial t} = 0 \tag{41}$$

$$\phi_B = 0 \tag{42}$$

In general, the same value of $\vec{P}_0$ will not satisfy all three of these equations simultaneously. More exactly, one effective position can be found for the BWS delay and delay rate but that position will not be suitable for phase. For a general distribution, the next few sections derive expressions for the structure effects in Eqs. (37), (38) and (39) and for effective position.

V. STRUCTURE EFFECT ON BWS DELAY

As indicated by Eq. (37), the effect of structure on BWS delay can be obtained by taking the frequency partial of the structure phase in Eq. (33), which gives

$$\Delta\tau_B = \frac{\left[Z_s \frac{\partial Z_c}{\partial \omega} - Z_c \frac{\partial Z_s}{\partial \omega} \right]}{Z_s^2 + Z_c^2} + \vec{B}_s \cdot \vec{P}_0/c \tag{43}$$

Based on Eqs. (34) and (35), the partials of Z become

$$\frac{\partial Z_s}{\partial \omega} = \int \left[\frac{1}{c} \vec{B}_s \cdot \vec{P}\, D \cos(\omega \vec{B}_s \cdot \vec{P}/c) + \frac{\partial D}{\partial \omega} \sin(\omega \vec{B}_s \cdot \vec{P}/c) \right] d\Omega \tag{44}$$

with a similar expression for Z_c.

If these partials are substituted into Eq. (43), we obtain

$$\Delta\tau_B = -\vec{B}_s \cdot \langle \vec{P} \rangle_R / c - B_s \, \Delta L_\omega / c + \vec{B}_s \cdot \vec{P}_0 / c \tag{45}$$

where ΔL_ω accounts for the frequency dependence of the brightness distribution and is given by

$$\Delta L_\omega \equiv \frac{Z_c I_s - Z_s I_c}{Z_s^2 + Z_c^2} \tag{46}$$

where

$$I_c \equiv \frac{c}{B_s} \int \frac{\partial D}{\partial \omega} \cos(\omega \vec{B}_s \cdot \vec{P}/c) d\Omega \tag{47}$$

with a similar expression for I_s. Further, the "centroid" of the resolved distribution is defined by the following weighted average:

$$\langle P \rangle_R \equiv \frac{Z_c^2 \langle P \rangle_c + Z_s^2 \langle P \rangle_s}{Z_c^2 + Z_s^2} \tag{48}$$

for which the "centroids" of the quadrature components of the resolved distribution are given by

$$\langle \vec{P} \rangle_c \equiv \frac{1}{Z_c} \iint \vec{P}\, D(\vec{P}, \omega, t) \cos(\omega \vec{B}_s \cdot \vec{P}/c) d\Omega \tag{49a}$$

$$\langle \vec{P} \rangle_s = \frac{1}{Z_s} \iint \vec{P}\, D(\vec{P},\omega,t) \sin(\omega \vec{B}_s \cdot \vec{P}/c)\, d\Omega \qquad (49b)$$

A special case is instructive. Suppose the brightness distribution is of the form

$$D(\vec{P},\omega,t) = f(\omega) D_0(\vec{P},t) \qquad (50)$$

In this case, one can easily show from Eq. (46) that ΔL_ω is zero so that the frequency dependence of the distribution has no effect on the BWS delay. (This result can also be obtained by considering the original brightness transform in Eq. (29). Note that a change in frequency for D would only change the multiplicative amplitude factor $f(\omega)$ and would therefore have no effect on phase.) Thus, if the brightness distribution retains the same shape but only changes in amplitude as a function of frequency, the BWS delay is the same as it would have been had the distribution had no frequency dependence. In general, since most sources change shape gradually as a function of frequency, the ΔL_ω term might be very small for many sources. More work with actual distributions is required to assess this point quantitatively. To begin to estimate this effect, Section XII develops a model for a frequency-dependent double-point source.

VI. STRUCTURE EFFECT ON PHASE-DELAY RATE

The derivation of the structure effect on delay rate observable closely follows that for the BWS delay. In analogy with the BWS delay case, the partial with respect to time in Eq. (38) leads to the result

$$\dot{\Delta\tau}_\phi = \frac{1}{\omega} \frac{\left(Z_s \frac{\partial Z_c}{\partial t} - Z_c \frac{\partial Z_s}{\partial t} \right)}{Z_s^2 + Z_c^2} + \dot{\vec{B}}_s \cdot \vec{P}_0/c \qquad (51)$$

Based on Eq. (35), the partial of Z_s will be given by

$$\frac{\partial Z_s}{\partial t} = \int \left[\frac{\omega}{c} \dot{\vec{B}}_s \cdot \vec{P}\, D \cos(\omega \vec{B}_s \cdot \vec{P}/c) + \frac{\partial D}{\partial t} \sin(\omega \vec{B}_s \cdot \vec{P}/c) \right] d\Omega \qquad (52)$$

and a similar expression for $\partial Z_c/\partial t$. These partials can be substituted into

Eq. (51) to give

$$\dot{\Delta\tau}_\phi = -\frac{\dot{\vec{B}}_s}{c} \cdot \langle \vec{P} \rangle_R - \frac{|\dot{\vec{B}}_s|}{c} \Delta L_t + \dot{\vec{B}}_s \cdot \vec{P}_0/c \tag{53}$$

where the centroid $\langle \vec{P} \rangle_R$ is given by Eq. (48), and where ΔL_t accounts for time dependence of the brightness distribution and is given by

$$\Delta L_t = \frac{Z_c H_s - Z_s H_c}{Z_c^2 + Z_s^2} \tag{54}$$

for which

$$H_s = \frac{c}{\omega |\dot{\vec{B}}_s|} \int \frac{\partial D}{\partial t} \sin(\omega \vec{B}_s \cdot \vec{P}/c) d\Omega \tag{55}$$

with a similar expression for H_c.

For most sources, the brightness distribution will change very little over an observation ($\sim$ 3 minutes) so that for these sources the term ΔL_t can be neglected. Further, in analogy with the derivation for the BWS delay, one can consider a source for which the total power changes as a function of time but the shape remains the same:

$$D(\vec{P}, \omega, t) = f(t) D_0(\vec{P}, \omega) \tag{56}$$

For such a source, one can easily show the effect of the time dependence on effective position is zero ($\Delta L_t = 0$ in Eq. 53). Thus, although such a model is generally unrealistic, it does show that, if the time dependence comes from total power change rather than shape changes, the effect on BWS delay rate will be zero.

VII. EFFECTIVE POSITION FOR BWS DELAY AND DELAY RATE

In general multiparameter fits, BWS delay and delay rate are usually analyzed together, so it makes sense to specify an effective position that is suitable for both. As explained in Section IV, conditions on effective position are obtained by setting Eqs. (45) and (53) equal to zero and solving for $\vec{P}_0$, which gives

$$\hat{B}_s \cdot \vec{P}_0 = \hat{B}_s \cdot \langle \vec{P} \rangle_R + \Delta L_\omega \tag{57}$$

for BWS delay and

$$\hat{\ell} \cdot \vec{P}_0 = \hat{\ell} \cdot \langle \vec{P} \rangle_R + \Delta L_t \tag{58}$$

for delay rate, where $\hat{\ell}$ is a unit vector in the direction of $\vec{B}_s$. The first term in each of these two equations is the "centroid" of the resolved distribution, while the second term comes from the frequency dependence and time dependence of the brightness distribution, respectively. As long as $\vec{B}_s$ is not parallel to $\vec{B}_s$, which is nearly always the case, the conditions in Eqs. (57) and (58) unambiguously provide the effective position. However, in the general case, this approach will produce coordinate distances specified along the nonorthogonal axes given by $\vec{B}_s$ and $\vec{B}_s$. To obtain coordinates along the (β, γ) axes, one will have to apply the transformation demonstrated graphically in Figure 1. As shown, the effective position is given by the intersection of two lines that orthogonally intersect the ($\vec{B}_s$ and $\vec{B}_s$) axes at the coordinate values in Eqs. (57) and (58), respectively. We will not present the trivial linear equations that express (β, γ) coordinates in terms of ($\vec{B}_s$, $\vec{B}_s$) coordinates since it adds little insight and much algebraic litter.

For many sources, it will probably be possible to make an approximation that greatly simplifies this calculation of effective position. With sources for which the time and frequency variability are negligible ($\Delta L_t = \Delta L_\omega = 0$), the effective position will be given by

$$\vec{P}_0 = \langle \vec{P} \rangle_R \tag{59}$$

This can be easily seen by considering Eqs. (57) and (58) with ΔL_t and ΔL_ω set equal to zero. Thus, in this important case, the effective position becomes the centroid of the resolved distribution.

A special case is worth giving as an example. For sources that are compact relative to fringe spacing (i.e., $B_s/\lambda \to 0$), one can easily show that the centroid of the resolved distribution becomes

$$\langle \vec{P} \rangle_R = \langle \vec{P} \rangle_0 \equiv \frac{\int\int \vec{P}\, D(\vec{P},\omega,t)\, d\Omega}{\int\int D\,(\vec{P},\omega,t)\, d\Omega} \tag{60}$$

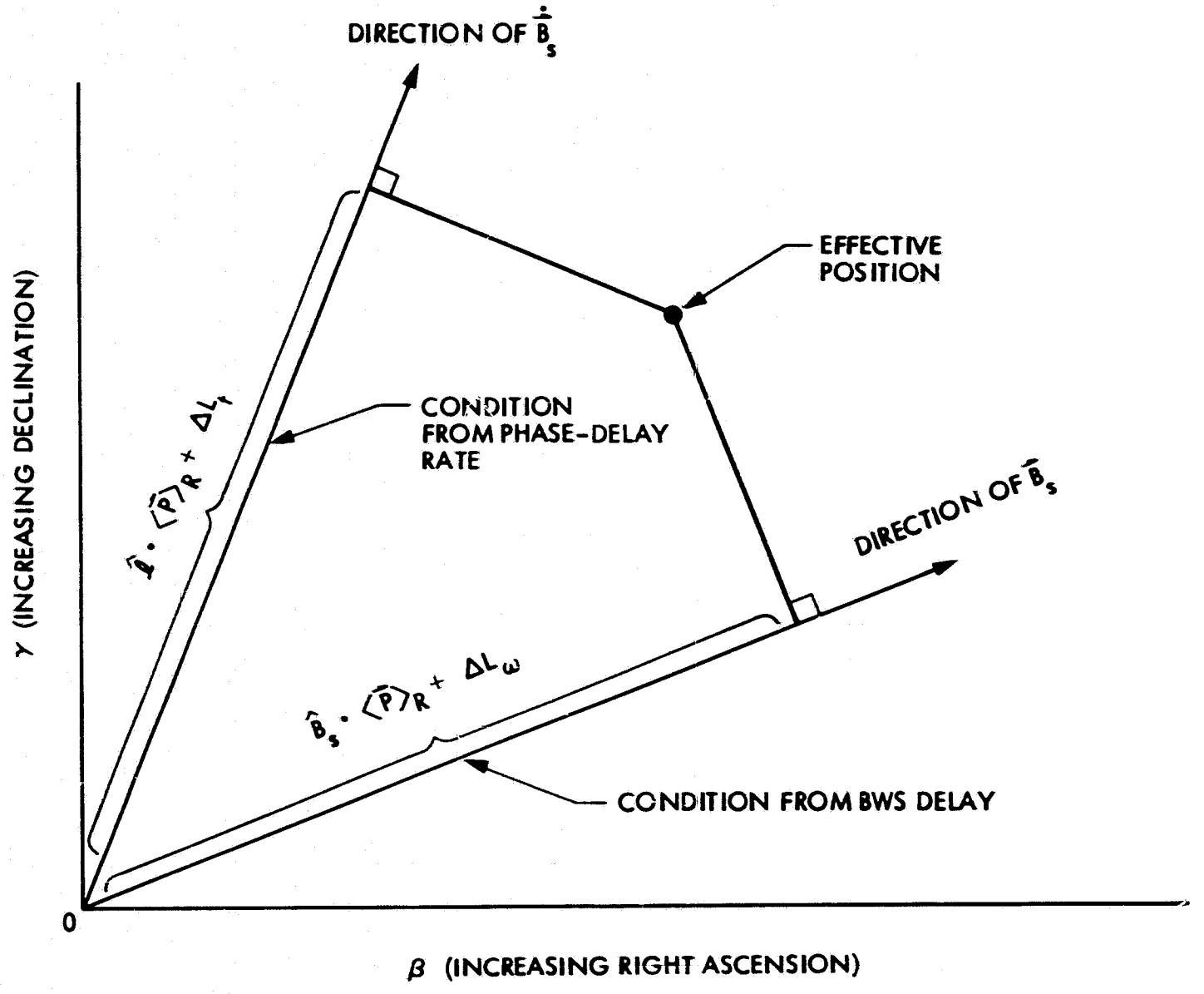


Figure 1. The Determination of Effective Position for Phase-Delay Rate and BWS Delay

Further, one can show $\Delta L_t = \Delta L_\omega = 0$ when $B_s/\lambda \to 0$. Thus, if the source is very compact, the effective position becomes the ordinary centroid of the distribution.

VIII. EFFECTIVE POSITION FOR PHASE DELAY

The effective position for the phase observable can be obtained from Eqs. (33) and (42), which gives the condition

$$\hat{B}_s \cdot \vec{P}_0 = \frac{1}{2\pi} \frac{\lambda}{B_s} \tan^{-1}\left(\frac{-Z_s}{Z_c}\right) \tag{61}$$

where Z_c and Z_s are given by Eqs. (34) and (35). Once the brightness distribution is specified, Z_s and Z_c can be readily determined, thereby yielding a value for $\hat{B}_s \cdot \vec{P}_0$. Thus, for the brightness transform to make no contribution to phase delay, the component of effective position along $\vec{B}_s$ must be made equal to the expression in Eq. (61). In general, the effective position along $\vec{B}_s$ will not be the same for the phase-delay observable as it is for the BWS delay observable. As in the case of BWS delay, however, the phase delay places no requirement on the position component in the direction orthogonal to $\vec{B}_s$. The next section will discuss the correction of phase delay for structure effects.

As was done in the last section for the effective position for BWS delay and delay rate, one can consider the special case of zero baseline. For zero baseline one can easily show that Eq. (61) becomes

$$\hat{B}_s \cdot \vec{P}_0 = \hat{B}_s \cdot \langle \vec{P} \rangle_0 \tag{62}$$

where $\langle \vec{P} \rangle_0$ is given by Eq. (60). Thus, as for the other observables, the ordinary centroid will again give the effective position when source extent is much smaller than the fringe spacing.

IX. STRUCTURE CORRECTIONS AND THE ORIGIN OF STRUCTURE COORDINATES

When source structure measurements are made, the origin of the r.a. and declination coordinates (structure coordinates) over which the brightness is distributed is not accurately known relative to absolute coordinates, even though the relative location of features in the distribution can be accurately specified. This section shows that, if the brightness distribution is "perfectly" determined in this relative sense, the absolute location of the origin of structure coordinates can be determined, in principle, to the accuracy allowed by the errors in measured delay and delay rate, as propagated through a standard multiparameter fit to determine baselines, source locations, etc. An important

limitation will be the error in aligning the S-band and X-band distributions and the resulting delay error due to S/X calibration of charged particle delays. In the process of explaining the determination of the origin of structure coordinates, this section also shows how brightness distribution measurements can be used to correct for the effect of structure on delay and delay rate. We will assume that the extent of the source is sufficiently limited and that the origin of structure coordinates is fairly close to the center of the source so that second order effects for partials w.r.t. position, etc., can be neglected. Further, it will be assumed that BWS delay and phase delay rate are the observables. The case of phase delay is discussed at the end of the section.

Since origin determination is closely associated with position determination, we will choose to use the concept of effective position to justify the procedure for absolute determination of the origin of structure coordinates. Although we will do so in this report, it will be shown that effective position does not necessarily have to be determined explicitly in practice.

We will first consider the case in which charged particle delays are negligible. As shown in Section IV, the fringes can be formulated so that the brightness transform makes no contribution to the observables of BWS delay and delay rate. This is accomplished by setting the reference position P_0 equal to the effective position discussed in Section VII. When this is done, the resulting observables of BWS delay and delay rate, which are obtained through the differentiation of Eq. (36) with respect to frequency and time, will be given theoretically by

$$\tau_{BWS} = \tau_g + \tau_c \tag{63}$$

$$\dot{\tau}_\phi = \dot{\tau}_g + \dot{\tau}_c \tag{64}$$

where τ_g is the standard geometric delay for the effective position. Explicit structure contributions ($\partial\phi_B/\partial\omega$ and $\partial\phi_B/\partial t$) have disappeared as a consequence of setting the reference position P_0 equal to the effective position. The most important aspect of this result is that both observables are the observables that would be generated by a point source at the effective position.

At this point, the absolute location of the effective position is unknown but, given the brightness distribution, the location of effective position can be calculated accurately relative to the origin of structure coordinates (or any other point in structure coordinates), as described in Section VII. If one wishes to apply a differential correction to the observables that will shift the effective position to another point in the source, he can do so by computing and applying a correction equal to the product of the vector difference in position (computed in structure coordinates) and the sensitivity partials relating observable to position. When such differential corrections are applied to the total observables in Eqs. (63) and (64), the new geometric delay and geometric delay rate will

pertain to a point source placed at the new reference position. The absolute coordinates of the new point will still be unknown while its position relative to structure coordinates will be accurately known.

The absolute coordinates of the new reference position can be determined as follows. Typically, a measurement session will involve several observations of a given source, for which the effective position changes from observation to observation. For each observation, the observables will be corrected for the displacement (the solution to Eqs. 57 and 58) of effective position from the new reference position. We will assume for each source that the new reference position is placed at the structure-coordinate origin, although another point could be chosen. When the observables are passed through the usual multiparameter fit for source location, baseline, etc., the solve-for source position for each corrected source will be the absolute location of its origin of structure coordinates. The brightness distribution for each source has now been placed absolutely on the celestial sphere to the accuracy allowed by the errors in the observables as propagated through the multiparameter fit and by the errors in the measured brightness distribution. If the data covers a long time span so that the brightness distribution of a given source changes significantly and must be remeasured, then one must face the problem of determining the origins of two or more sets of structure coordinates for one source. Two choices present themselves. One can either solve for a separate position (origin) each time a new distribution is introduced, or all origins for a source can be forced into agreement by prominent feature alignment. The choice of an approach depends on experimental goals and the nature of the source.

The preceding discussion was presented to demonstrate through position analysis how the origin of structure coordinates can be absolutely determined. As one might guess, it is not necessary to compute effective position explicitly but only structure delay (and delay rate) relative to the origin of structure coordinates (Eqs. 43 and 51 with $\vec{P}_0 = 0$). Following the procedure outlined above, one would need the sensitivity relation given by

$$\Delta\tau = -\vec{B}_s \cdot \Delta\vec{S}/c \tag{65}$$

where $\Delta\vec{S}$ is the shift in position for the source. Note that if the effective position described in Section VII is substituted in place of $\Delta\vec{S}$ as prescribed, the resulting delay will be given by Eq. (45) with $\vec{P}_0 = 0$, by design of the effective position. A similar argument can be made for delay rate. Thus, if one computes the structure delay (and delay rate) directly from the derivatives of structure phase (Eq. 33) with $\vec{P}_0 = 0$, the resulting delay (delay rate) is the correction needed to make the observable equal to that generated by a point source at the origin of structure coordinates.

When charged particle delays are considered, the procedure becomes somewhat complicated, even though the basic concepts are the same. If observations were made at one frequency and outside calibrations could supply accurate values for

charged particle delays, the procedure would be the same as that explained above. However, if we assume dual frequency (S/X) calibrations of those delays, we must complicate the procedure with two brightness distributions, one from each frequency. The main problem is that each distribution will have its own unknown origin unless steps have somehow been taken to align them or to determine their relative placement. At this time, a method has not been verified that will accurately align origins in the general case. We cannot use the preceding approach separately for S-band and X-band since the charged particle delays will corrupt the determination of absolute position. Though it may not be very satisfactory, one solution is to use the two measured brightness distributions to align prominent features via a translation of the S-band distribution. The accuracy requirements on such a translation will be very stringent but, due to the nature of S/X calibrations, it will be reduced by a factor of 13 relative to the standard error requirement. For example, a 1-cm error in delay on a 6000-km baseline would be produced by a misalignment error of 13*0.0003"/cm * 1 cm = 0".004. Determination of whether enough sources can be treated by the "prominent-feature" approach will require a study of actual S/X structures. There are apparently a number of sources for which this would be a feasible approach (e.g., Ref. 2).

Once the origin of the distribution at S-band has been made consistent with that at X-band, whatever the approach, the next step would be to correct each observable for the displacement of effective position from the origin of structure coordinates, as described above. The S-band and X-band observables would, of course, be corrected according to their respective brightness distributions. Due to differences in structure, the two bands will typically have different corrections. This procedure would, in effect, make the source a point source at both S-band and X-band, with a location equal to the (presently unknown) position of the origin of structure coordinates at X-band. After the completion of this procedure for all sources, the next step would be to perform S/X calibration for each observable type by means of the standard linear combination of S- and X-band observables. Such a calibration would now be valid since the S- and X-band positions have now been made coincident for all observations, provided the aforementioned alignment of origins has been accurately carried out. The corrected observables (the output of he S/X linear combinations) would then be simultaneously subjected to a multiparameter fit for baseline, source position, etc. The resulting solution gives for each source the absolute location of the (S/X common) origin of structure coordinates. As a result of this procedure, the S- and X-band distributions have both been accurately placed in an absolute sense on the celestial sphere for all sources.

The preceding discussion assumed that BWS delay and delay rate were the observables. If absolute phase delay is the observable, the concepts are largely the same but the equations change. In analogy with the other observables, the phase delay will be corrected by ϕ_B in Eq. (33) with $\vec{P}_0 = 0$. Unlike the structure corrections for the other observables, structure phase is afflicted with integer-cycle ambiguities. The method for handling these ambiguities consists of the following steps. First, the fractional part of ϕ_B is computed through Eq. (33) and subtracted from fringe phase, which itself still possesses an absolute ambiguity. This ambiguity in corrected fringe phase would

then be removed by the BWS technique in which an iterative process successively removes the ambiguities in (structure-corrected) BWS delays for ever-wider channel separations until KF phase is reached. For the BWS procedure to succeed, the structure-corrected BWS delays and phase delay must all be equal to the same delay, except for ambiguities and measurement uncertainties. Therefore, to equalize the geometric component, structure effects in BWS delays must be computed and removed relative to the same assigned source position as is phase. (The computation of structure effects in Section IV approximates a finite difference with a derivative. For very widely spaced channels, that approximation might fail. It is beyond the scope of this report to treat the differential analysis.) After absolute removal of integer-cycle ambiguities through the BWS technique, the resulting phase delay will be equal to the geometric delay that would be generated by a point source at the origin of structure coordinates. S/X corrections also closely follow the preceding discussion for BWS delay and delay rate.

Associated with the preceding consideration of structure effects is the issue of assigning a nominal location to a source. Once milliarcsecond accuracies are reached, two numbers will no longer be sufficient to specify the location of a complicated source. That is, the effective location will vary as a function of baseline length and orientation. As suggested above, source geometry has been completely specified once one has been supplied the brightness distribution and the absolute location of the origin of structure coordinates. Since the origin can be arbitrarily shifted to any place in the brightness distribution, one should attempt to make the two required numbers specifying the origin have as much utility as possible. One can make the point that the "compact component" should be made the origin since long-baseline solutions will apply to the compact component and will give the most accurate positions that are to be compared. However, in the most accurate measurements, the "compact component" position can be strongly dependent on the geometry of resolution and therefore only has meaning in terms of a specific geometry. One attractive possibility is the ordinary centroid, since it gives the average location in the conventional sense and is the correct effective position for the zero-baseline case. Further, it requires no specification of baseline orientation to give it meaning. It is definitely better than an arbitrary assignment, since effective positions computed in structure coordinates would then automatically be given in terms of deviations from the zero-baseline case.

X. DIFFERENTIAL POSITION MEASUREMENTS

One application of differential VLBI is the measurement of the relative positions of two natural sources. By forming the difference of the observables for closely spaced sources, one can remove many errors, thereby allowing very accurate determination of differential positions. The resulting differential observables are fit with a delay model possessing only differential right ascension and declination as solve-for parameters and perhaps a constant "clock" term (for the phase-delay case).

As indicated earlier, source structure effects can be quite large and must be corrected in some cases. The correction procedure would be like that described in the last section, but now two distributions would require corrections, one for each source. Thus, two structure corrections would be applied to each differential observable, where each correction is computed relative to some assigned reference position. In effect, this turns each source into a point source located at its respective reference position. When all observations have been corrected and then subjected to a multiparameter fit, the resulting differential coordinates will refer to the difference in the two adopted reference positions.

At this point we are no longer dependent on a particular observation and geometry for a specification of the relative location of the sources. Since we now have a complete description (i.e., the two distributions along with an accurate measure of the separation of their origins), we can now select whatever measure of separation we choose. This report will not propose a "best" definition of that measure since more consideration of the problem is required. One possibility would be to use the "compact-component centroid" for a reference geometry of resolution, where the same geometry would be applied to successive experiments. In this way, changes in resolution geometry would be eliminated as a source of error. However, changes in the compact components could still contribute substantial errors to the measured change in differential effective position between experiments. Such effects would have to be corrected through comparisons of the full distributions, perhaps by prominent compact-feature alignment between the brightness distributions of successive experiments.

A few words are necessary to explain the treatment of integer-cycle ambiguities in structure phase and fringe phase in the case of differential measurements. To obtain the final observable, it is only necessary to compute the fractional part of ϕ_R in Eq. (33) and subtract it from fringe phase. (However, one can restore integer-cycles to structure phase in a relative sense by requiring continuity along a u-v track.) The ambiguities still present in corrected fringe phase can be removed in one of three ways, depending on the type of experiment. In single-channel measurements with poor a priori values for baseline, etc., ambiguities can only be removed in a relative sense by requiring continuity between observations. The unknown overall bias created by unresolved ambiguities would be modeled and solved-for in parameter estimation. In single-channel experiments with sufficiently accurate a priori information, ambiguities can be removed absolutely by comparing each measured value with differential phase predicted on the basis of a priori information. In multiple-channel experiments with suitably placed channels and phase calibration, ambiguities can be removed through the BWS procedure. For the last two absolute schemes, there would be no need to solve for a bias term provided adequate differential cancellation of instrumental effects, etc., has occurred.

Due to small observable errors and geometric considerations, differential measurements will provide a good direct test of the theory of structure effects. Measurements of delay, for example, are directly proportional to the displacement of effective position in the direction $\vec{B}_S$ for each observations. Because of this, delay residuals and their reduction through the application of structure corrections can readily be interpreted geometrically in terms of effective position.

XI. THE SYMMETRIC BRIGHTNESS DISTRIBUTION

For a symmetric distribution, one can easily show that the structure phase is some value of $n\pi$, provided the center of symmetry is taken as the origin of structure coordinates. By definition the distribution is symmetric if

$$D(\beta,\gamma,\omega,t) = D(-\beta,-\gamma,\omega,t) \tag{66}$$

If in Eq. (24) we place β_0, γ_0 at the origin of structure coordinates ($\beta_0 = \gamma_0 = 0$) and take the complex conjugate, we obtain

$$R^*(u,v,\omega,t) = \int D(\beta,\gamma,\omega,t) e^{2\pi i(u\beta+v\gamma)} \, d\beta d\gamma \tag{67}$$

By a change of integration variables given by

$$\beta' = -\beta \tag{68}$$

$$\gamma' = -\gamma \tag{69}$$

and by the use of the symmetry relation in Eq. (66), we obtain

$$R^*(u,v,\omega,t) = \int D(\beta',\gamma',\omega,t) e^{-2\pi i(u\beta'+v\gamma')} \, d\beta' d\gamma' \tag{70}$$

which is the complex conjugate of Eq. (67). Thus, when the reference position is placed at the center of symmetry of a symmetric source, the brightness transform R is equal to its complex conjugate. This shows that, for a symmetric distribution, the transform is real and the structure phase is therefore equal to $n\pi$. In the limit of zero baseline length ($u = v = 0$), the brightness transform becomes the total power of the source, which is a positive real number. In this limit, the transform phase ϕ_B will be defined to be zero. Since the transform phase of a symmetric distribution can only be equal to the aforementioned discrete values of $n\pi$, it is clear that if phase changes, it must change in jumps. However, for a well-behaved brightness distribution, the brightness transform will be a continuous function along a (u-v) track and consequently the transform phase can only change at those u-v values for which the transform (fringe) amplitude is zero. The transform phase will be constant along tracks away from those u-v values

and, if it changes, the change must occur in the form of a discontinuous jump at a null amplitude point. We will see an example of such behavior in the following section. Since BWS delay and delay rate are derivatives of phase, the effect of a symmetric distribution on these observables will be zero except at points of zero amplitude, where the effect can be infinite (δ-function response). For the more realistic case of a nearly symmetric distribution, the structure effect on BWS delay can be quite large near (u,v) points of nearly null amplitudes and quite small away from such points. It follows from the preceding discussion that, if the brightness transform of a symmetric distribution possesses no finite (u,v) points of null amplitude, then the transform phase is zero for all (u,v).

XII. THE DOUBLE-POINT SOURCE

The next example will be a time-invariant source consisting of two close point sources whose strengths and spectral indices can be varied. The brightness distribution will be described by

$$D(\vec{P},\omega) = r_1 \left(\frac{\omega}{\omega_0}\right)^{k_1} \delta(\vec{P} - \vec{P}_1) + r_2 \left(\frac{\omega}{\omega_0}\right)^{k_2} \delta(\vec{P} + \vec{P}_1) \tag{71}$$

where we have chosen the midpoint between sources as the origin of structure coordinates, and where k_ℓ is the spectral index of source ℓ. The factors r_ℓ can be used to vary the strength of the sources, while ω_0 is a normalizing reference frequency. For this distribution the brightness transform in Eq. (29) becomes

$$R = |R| e^{i\phi_B} \tag{72}$$

$$= e^{i\omega\vec{B}_s \cdot \vec{P}_0/c} \left[g_1 e^{-i\omega\vec{B}_s \cdot \vec{P}_1/c} + g_2 e^{i\omega\vec{B}_s \cdot \vec{P}_1/c} \right] \tag{73}$$

$$\text{where } g_\ell = r_\ell \left(\frac{\omega}{\omega_0}\right)^{k_\ell} \tag{74}$$

Fringe visibility can be easily computed from this expression by taking the RSS of the real and imaginary parts and dividing by the total flux, $g_1 + g_2$, which gives

$$v_s = \left[v_m^2 + (1 - v_m^2) \cos^2 (\omega\vec{B}_s \cdot \vec{P}_1/c)\right]^{1/2} \tag{75}$$

where the minimum fringe visibility is given by

$$v_m = \frac{g_1 - g_2}{g_1 + g_2} \tag{76}$$

(Note that v_m can be either positive or negative so that, strictly speaking, the absolute value of v_m is the minimum fringe visibility.) Further, one can easily solve for structure phase to obtain

$$\phi_B = -\tan^{-1}\left[v_m \tan\left(\omega\vec{B}_s \cdot \vec{P}_1/c\right)\right] + \omega\vec{B}_s \cdot \vec{P}_0/c \tag{77}$$

It is convenient to rewrite this equation in the form

$$\phi_B = -\tan^{-1}\left[v_m \tan(\theta/2)\right] + \theta/2 \tag{78}$$

where

$$\theta = 2\pi S_B/\Delta \tag{79}$$

The quantity S_B is the component of the vector $\delta\vec{S} = 2\vec{P}_1$ containing the two sources projected onto the baseline vector and is given by

$$S_B = \delta\vec{S} \cdot \hat{B}_s \tag{80}$$

The quantity Δ is the interferometer resolution (spacing of fringes on the sky) and is given by

$$\Delta = \frac{2\pi c}{\omega B_s} \tag{81}$$

$$= \frac{\lambda}{B_s} \tag{82}$$

We have chosen this representation because it conveniently makes θ (in cycles) equal to the projected separation of the source divided by the interferometer resolution. Further, in Eq. (78) we have set $\vec{P}_0 = \vec{P}_1$ so that the phase that will be plotted will be referenced to the location of point 1. With this choice of

reference point, the phase oscillates about zero, remaining between $\pm\pi/2$, provided source 1 is the stronger source. Figure 2 shows the behavior of ϕ_B for various values of the strength ratio g_2/g_1. This phase behavior can be easily understood by plotting the phasors in Eq. (73) in the complex plane with $\vec{P}_0 = \vec{P}_1$, as schematically shown in Figure 3. For g_2/g_1 near 1.0 (i.e., equal strengths), the phase forms a sawtooth pattern as a function of S_B/Δ with jumps of π whenever the projected source separation is an integer multiple of one-half of the fringe spacing. In terms of phasors in Fig. 3a, such jumps occur whenever the second phasor has a phase of $n\pi$ so that the sum amplitude is nearly zero and the sum vector moves rapidly from the first to the fourth quadrant as a function of the phase of the source - 2 phasor. When the strength of source 2 is very weak compared to source 1($g_2/g_1 < < 1$), the phase oscillates by a small amount about zero, as shown for $g_2/g_1 = 0.1$ in Fig. 2 and illustrated with phasors in Fig. 3b.

Effective position for phase-delay, relative to the origin of structure coordinates, can easily be obtained by setting Eq. (77) equal to zero and solving for P_0:

$$\frac{\hat{B}_s \cdot \vec{P}_0}{S} = \frac{1}{2\pi} \frac{\Delta}{S} \tan^{-1}\left[v_m \tan(\pi S_B/\Delta)\right] \tag{83}$$

where S is the separation of the two sources ($S = 2|\vec{P}_1|$). Thus the component of effective position along $\vec{B}_s$ is a fairly complicated function of resolution geometry and source characteristics. Numerical examples of effective position are given in Fig. 4 under the assumptions that the baseline vector $\vec{B}_s$ points along the vector connecting the sources (S_B=S) and that structure phase is a continuous function of S/Δ that equals zero when the baseline length equals zero. As indicated in Eq. (83), the results are given in Fig. 4 in units of source separation so that deviations from the position of the stronger source (#1) are to be compared with the source separation shown in the figure. Note that the effective position varies from source 1 by less than one-half of the source separation and decreases as the strength of source 2 decreases. The decrease in the oscillation of effective position with increasing S/Δ is a result of increasing sensitivity of phase to position change (e.g., $\partial\phi/\partial\delta$ becomes larger as baseline length increases). Though the effective positions shown here for phase delay are not as constant as one would like, they are much more stable than the effective positions that will be derived for BWS delay later in this section.

The effect of structure on BWS delay can be obtained by differentiating Eq. (77), with respect to ω, which gives

$$\Delta\tau_B = \frac{\partial\phi_B}{\partial\omega} = \tau_D + \tau_p + \vec{B}_s \cdot \vec{P}_0/c \tag{84}$$

where

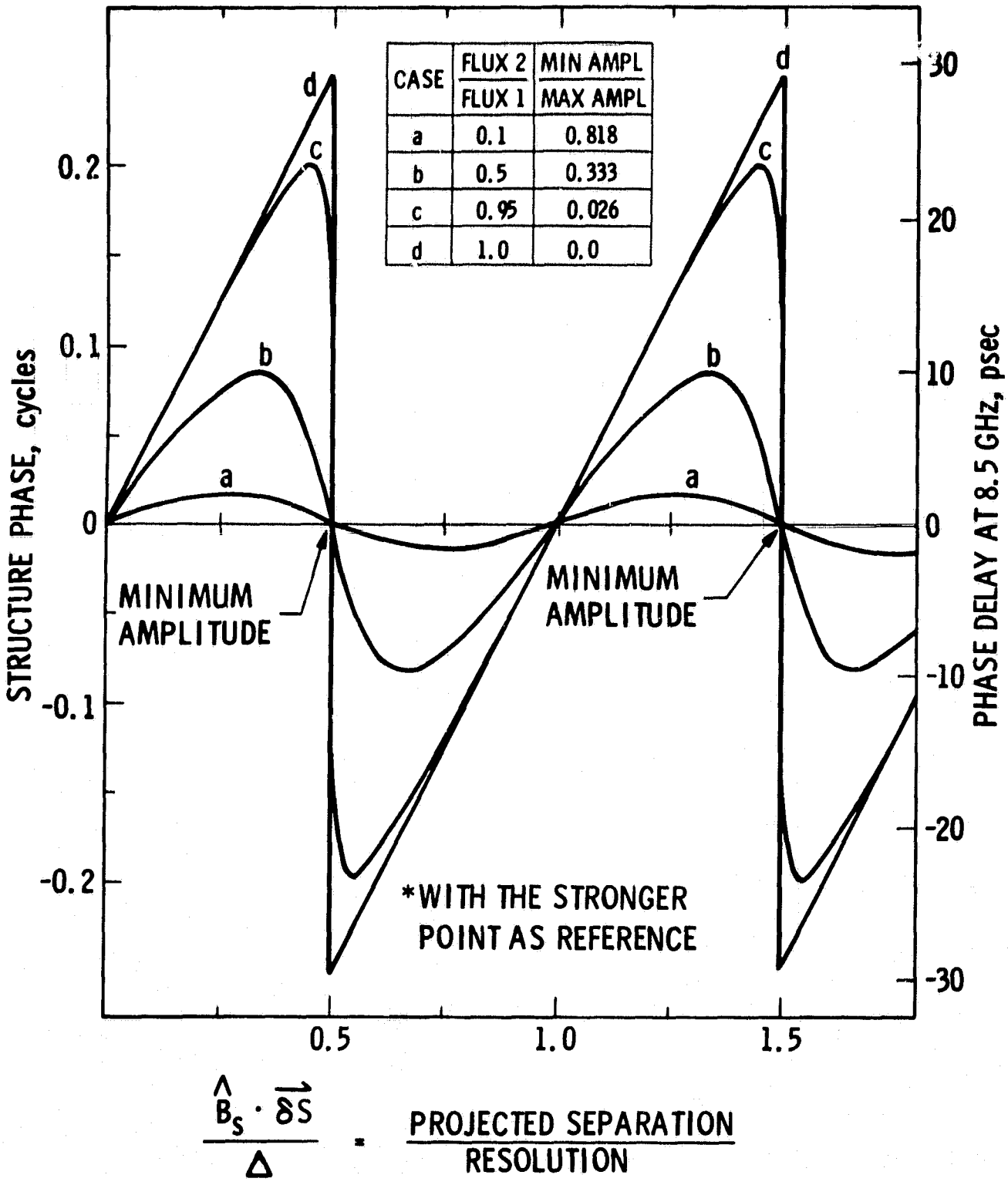


Figure 2. Structure Phase* for a Double-Point Source

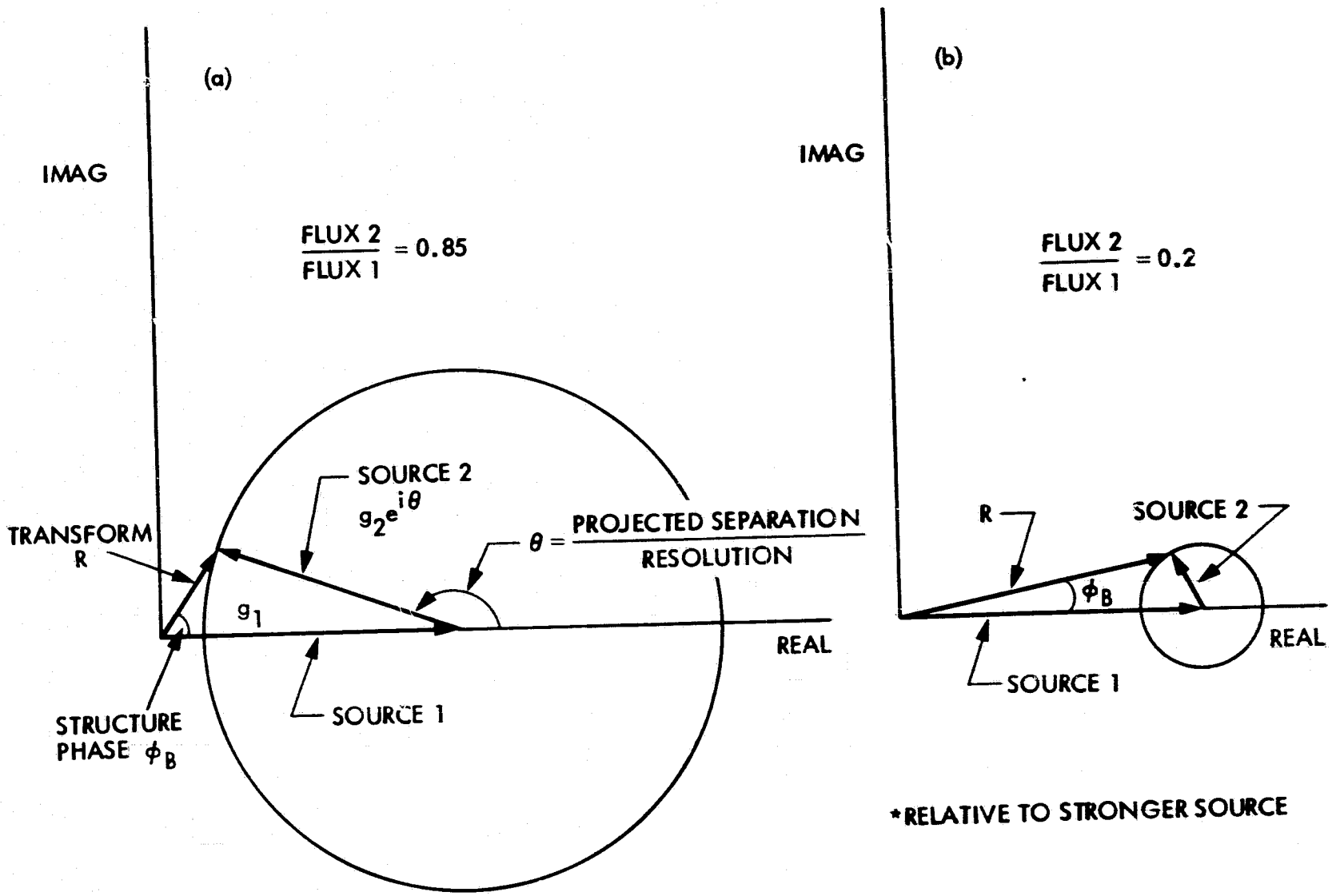
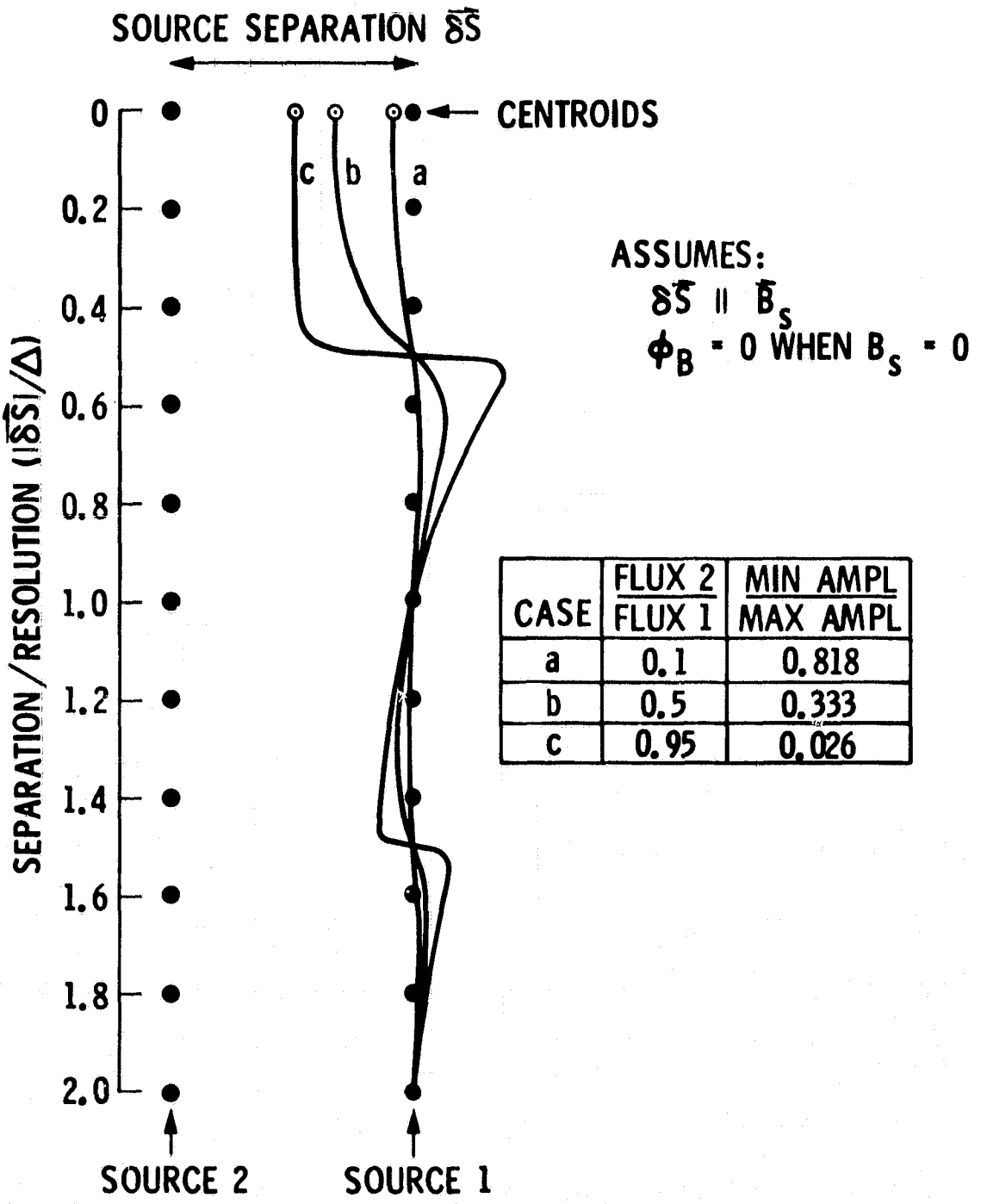


Figure 3. Schematic Example of the Brightness Transform* for a Double-Point Source



CASE	FLUX 2 / FLUX 1	MIN AMPL / MAX AMPL
a	0.1	0.818
b	0.5	0.333
c	0.95	0.026

Figure 4. Effective Position for a Double-Point Source: Phase Observable

$$\tau_D \equiv \frac{1}{2\omega} \frac{(k_2 - k_1)(1 - v_m^2)}{1 + h^2} \tan(2\pi \vec{B}_s \cdot \vec{P}_1/\lambda) \tag{85}$$

$$\tau_p \equiv -\frac{v_m}{1 + h^2} \frac{\vec{B}_s \cdot \vec{P}_1}{c} \sec^2(2\pi \vec{B}_s \cdot \vec{P}_1/\lambda) \tag{86}$$

$$h \equiv v_m \tan(2\pi \vec{B}_s \cdot \vec{P}_1/\lambda) \tag{87}$$

The first term τ_D comes from the frequency dependence of the brightness distribution while the second term τ_p comes from the frequency dependence of u, v in the transform or, in other words, from the frequency dependence of the sky fringes. Note that, as one would expect, the first term is zero when $k_1 = k_2$. (When $k_1 = k_2$, the shape of the distribution does not change with frequency so that the BWS delay effect due to frequency dependence is zero, as explained in Section V.) Further, the second term τ_p is zero if $g_1 = g_2$, except at singular points. (When $g_1 = g_2$, the distribution is symmetric about the origin of structure coordinates. As explained in Section XI, such symmetry leads to zero effect on BWS delay, except for possible singular points.)

The effective position associated with the BWS delay can be obtained by finding the value of $\vec{P}_0$ that makes the delay in Eq. (84) zero. The value of the component of $\vec{P}_0$ along B_s that will do this given by

$$\hat{B}_s \cdot \vec{P}_0 = \frac{1}{1+h^2}\left[\frac{\lambda}{B_s}\frac{(k_1-k_2)(1-v_m^2)}{4\pi}\tan(2\pi\vec{B}_s \cdot \vec{P}_1/\lambda) + v_m \hat{B}_s \cdot \vec{P}_1 \sec^2(2\pi\vec{B}_s \cdot \vec{P}_1/\lambda)\right] \tag{88}$$

More will be said about this result for effective position after the delay rate has been analyzed below.

Before proceeding to the delay rate observable, numerical examples will be presented for the two delay terms in Eq. (84) that give the effect of double-point structure on BWS delay. The first term in Eq. (84), the one due to the frequency

dependence of the point strengths, can be rewritten as

$$\tau_D = \frac{(k_2 - k_1)\ (1 - v_m^2)}{2\omega} \quad \frac{\tan(\theta/2)}{1 + v_m^2 \tan^2 (\theta/2)} \tag{89}$$

where θ was defined above as the ratio of projected source separation and resolution. Since this function is periodic and antisymmetric about $\theta = 0$, one plot of τ_D for $0 < \theta/2 < \pi/2$ suffices for all θ. Figure 5 plots τ_D as a function of θ for various values of relative source strength g_2/g_1, with the spectral index difference equal $(k_2 - k_1)$ to 1.0. Since the effect is proportional to $k_2 - k_1$, the delay for other values of $k_2 - k_1$ can be obtained by multiplying the results from Fig. 5 by $k_2 - k_1$. The delay effect is zero for $\theta = 0$ (or $2n\pi$) and passes through a maximum toward π (or $(2n + 1)\pi$). Note that the largest values (e.g., > 0.2 nsec) are obtained when the source strengths are nearly equal ($0.95 < g_2/g_1 < 1.0$) and the spectral index difference is large ($k_2 - k_1 > 1.0$). An important quantity is the duration of the delay "pulse" as a function of time (through θ) on an intercontinental baseline. For a 6000-km baseline observing a source with separation of 0".001 at X-band, θ can vary by no more than 0.005 cycles in 1 minute. Thus, in this case, only a fraction of the "pulse" width will be traversed in a 3-minute observation, when $g_2/g_1 < 0.95$.

A general expression can be easily computed for the maximum value of τ_D as a function of θ. Such a maximum occurs whenever

$$\tan(\theta/2) = \pm \frac{1}{v_m} \tag{90}$$

at which point the maximum of τ_D is

$$\tau_D\Big|_{max} = \pm \frac{k_1 - k_2}{4\omega} \quad \frac{(1 - v_m^2)}{v_m} \tag{91}$$

This function is plotted in Fig. 6 at 8.5 GHz as a function of the spectral index difference $(k_1 - k_2)$ and of the strength ratio (g_2/g_1). Note that the delay effect can become infinite as the strengths converge $(g_2 \to g_1)$. The reason for this singularity can be seen in Fig. 3a. For $g_2 \approx g_1$ and $\theta \approx (2n + 1)\pi$, the brightness transform is the sum of two phasors of nearly equal length pointing in nearly opposite directions. Consequently, the amplitude of the sum phasor is very small compared to the amplitude of the component phasors, and a slight change in the amplitude of the component phasors (i.e., of the point strengths g_1 and g_2) as a function of frequency (as in the partial $\partial\phi_B/\partial\omega$) can cause large changes in the phase of the sum phasor. As we shall see, similar singular behavior is exhibited by the second term τ_p in Eq. (84) as a consequence of the

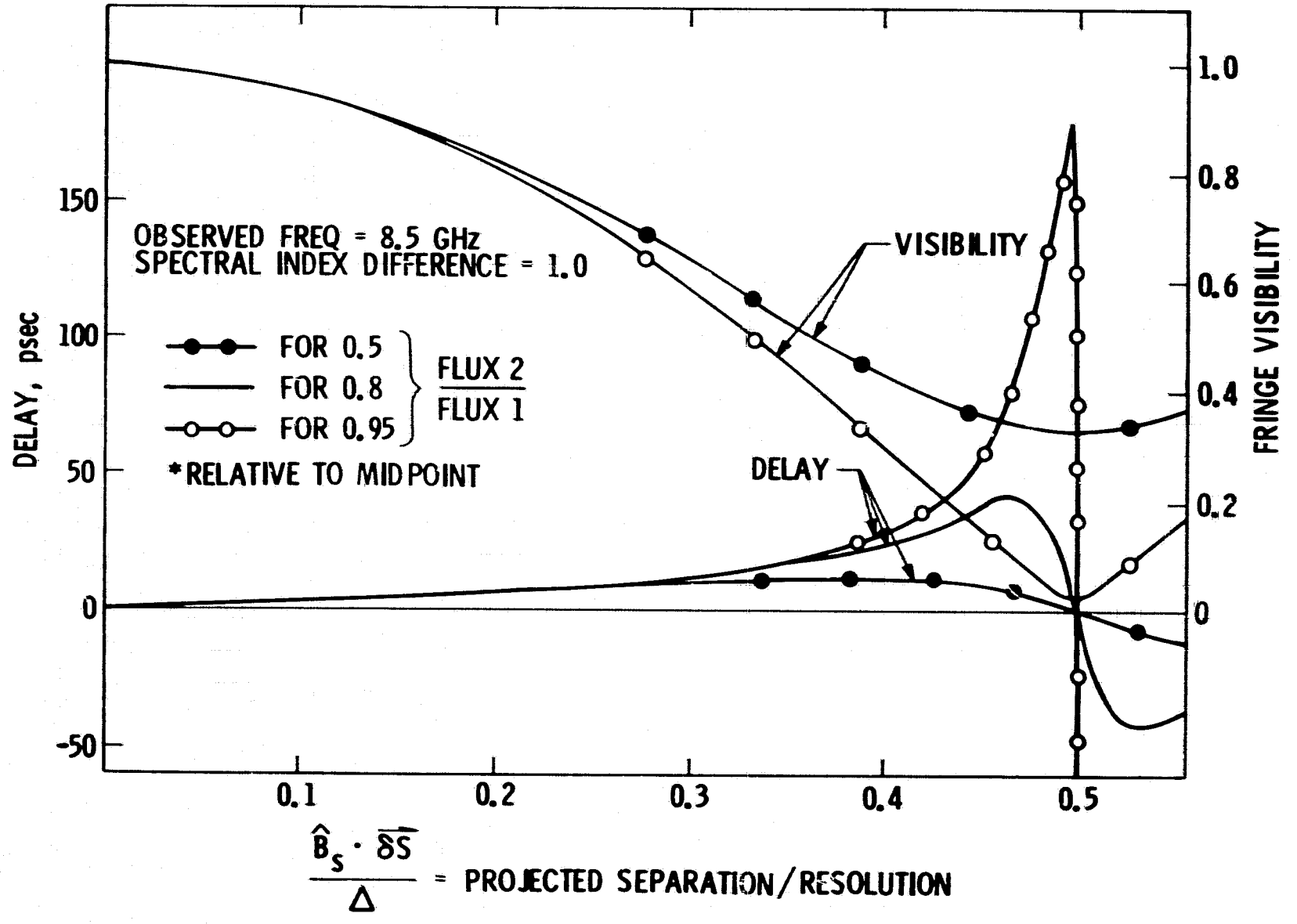


Figure 5. BWS Delay* Due to Frequency Dependence of a Double-Point Source

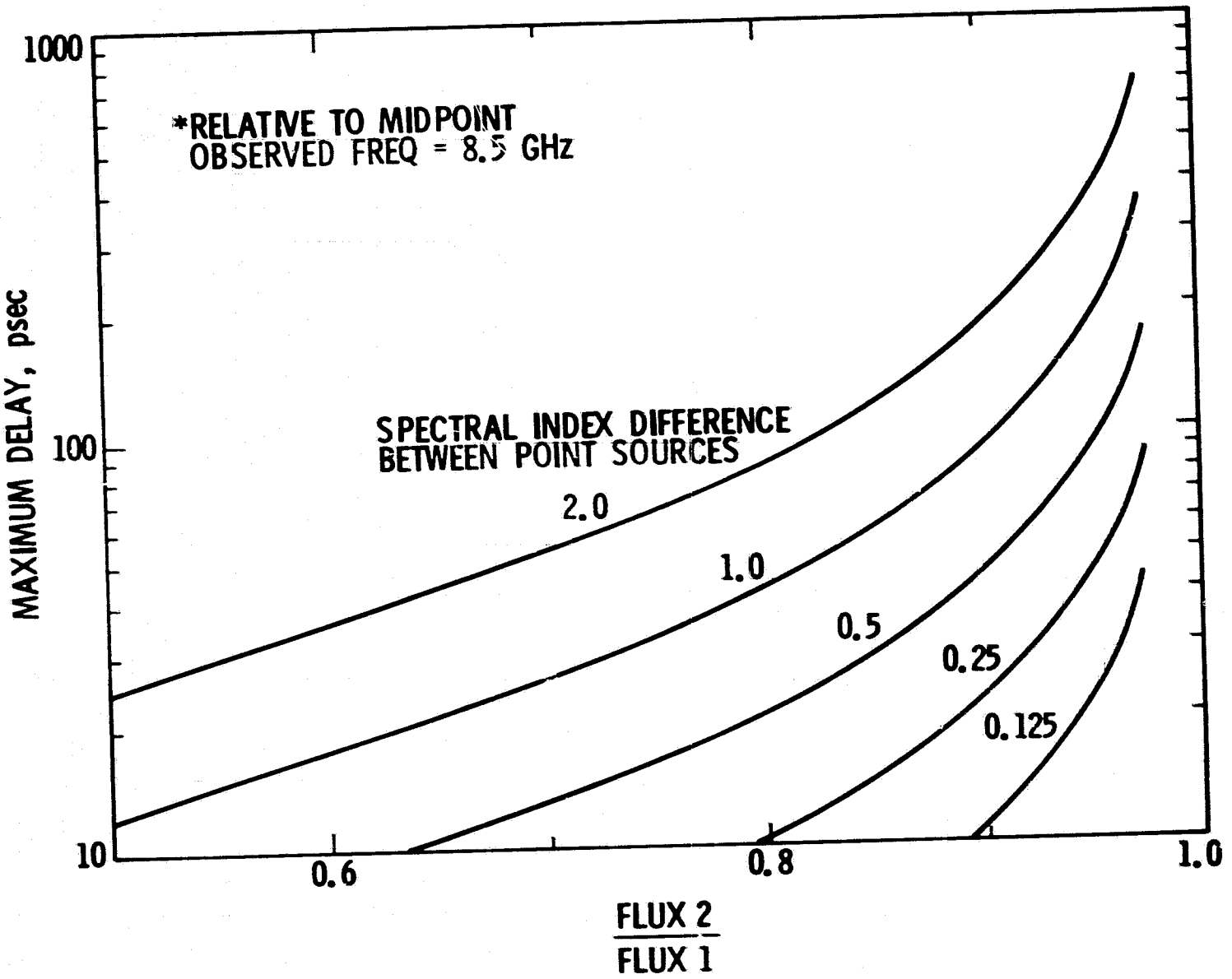


Figure 6. Maximum BWS Delay* Due to Frequency Dependence of a Double-Point Source

phase changes in the component phasors. The special case suggests that if a given source exhibits very small fringe amplitude relative to the amplitude for other (u,v) points, then that source should be investigated for undesirable behavior in structure phase at that amplitude minimum. For the particular case of a double point source with $k_2 - k_1 = 1.0$, the maximum value of the structure delay term τ_D reaches 0.18 nsec (5.5 cm) at x-band when $g_2/g_1 = 0.95$ (or visibility = 0.026).

To summarize the results for this delay term, the frequency dependence of structure can theoretically cause significant delay errors in the case of special double-point sources, if they are not corrected. However, these worst-case sources would probably be unusual since they would have to possess both nearly equal point strengths and a large spectral index difference. Given adequate structure measurements as a function of frequency, corrections can be applied to the delay observables to remove this effect adequately.

For the same example of a double point source, the second structure effect (τ_p) on BWS delay is caused by the frequency dependence of the sky fringes, given by Eq. (86). This term can be rewritten as

$$\tau_p' \approx -\frac{v_m}{\omega}\,\frac{\theta/2}{\cos^2(\theta/2) + v_m^2 \sin^2(\theta/2)} + \frac{v_m}{\omega}\,\frac{\theta}{2} \tag{92}$$

where v_m and θ are defined above and where the second term has been included in order to reference the delay to the brightness centroid given by

$$\vec{P}_c = v_m \vec{P}_1 \tag{93}$$

Though such a subtraction is not necessary, it removes a distracting linear drift (with respect to θ) from the delay and comes close to referencing the structure delay to an "average" or "typical" source position. This is equivalent to assuming that the last term in Eq. (84) has been combined with τ_p where we have made the definition $P_0 = P_c$. Figure 7 plots τ_p' over a range of θ at 8.5 GHz for few values of the strength ratio g_2/g_1. Relative maxima occur near $\theta = (2n+1)\pi$, with the maximum values increasing with increasing θ. A fairly good approximation for these maxima can be obtained by setting $\theta = (2n + 1)\pi$, which gives

$$\tau_p'\Big|_{max} \approx -\frac{\pi(1 - v_m^2)}{\omega v_m}\,\frac{\theta_n}{2\pi} \tag{94}$$

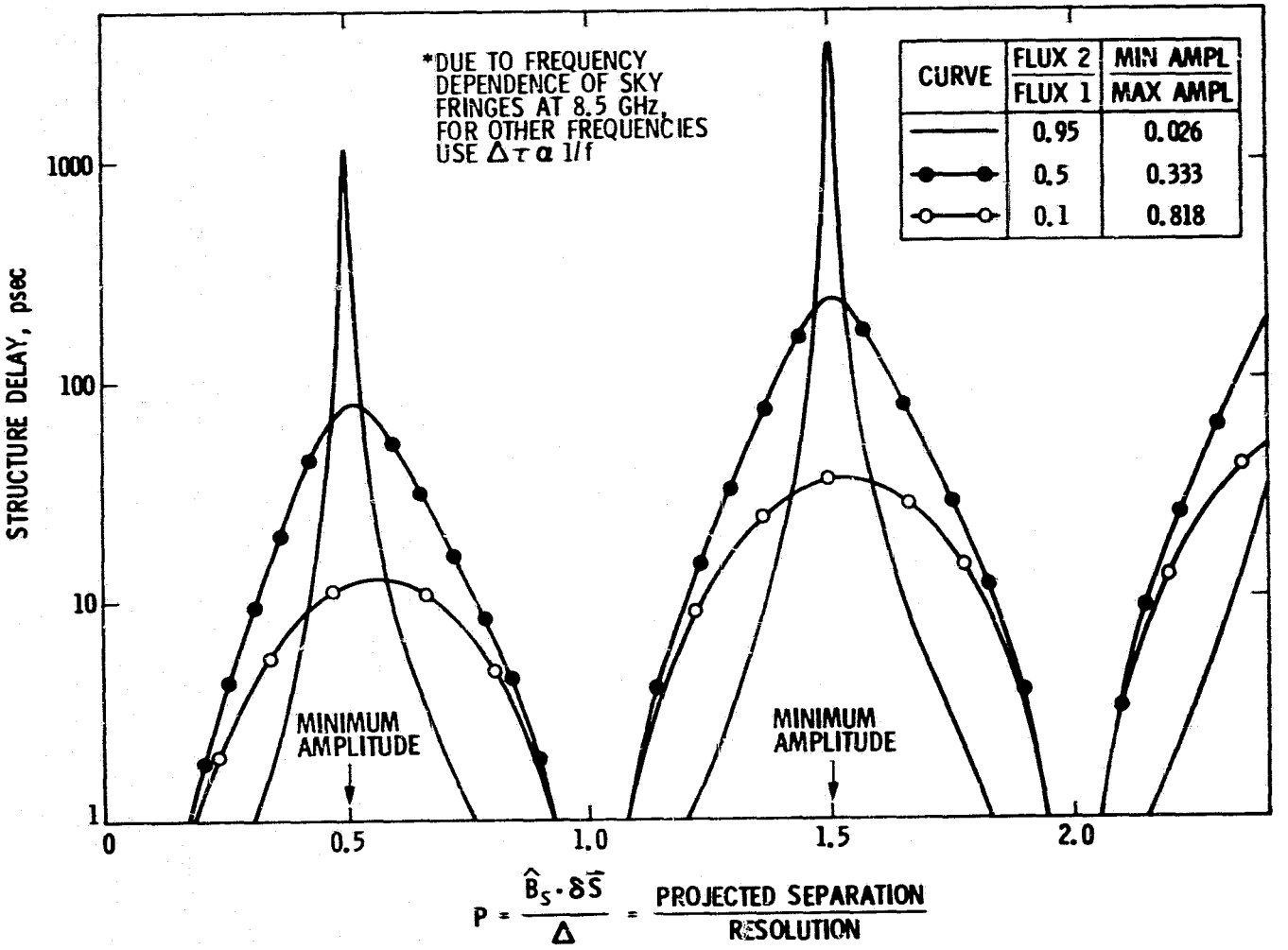


Figure 7. BWS Delay* Relative to Brightness Centroid for Double-Point Source

where

$$\frac{\theta_n}{2\pi} = \left(\frac{S_B}{\Delta}\right)_n = \frac{2n+1}{2} \qquad n = 0, \pm 1, \pm 2, \ldots \tag{95}$$

Thus the maxima increase in magnitude as the source strengths converge (for the reasons discussed above) and as the separation/resolution ratio increases.

Figure 8 plots $\tau_p'\Big|_{max}$ at 8.5 GHz as a function of S_B/Δ and g_2/g_1. For $g_2/g_1 = 0.95$ and separation/resolution = 0.5, the maximum delay effect is about 1 nsec. For a wide range of strength ratio and S_B/Δ, the effect is greater than 30 psec (1 cm). Thus, it appears that, for many sources that are nearly double point, this structure effect can be very significant and must be corrected if accuracies at the centimeter level are required. The problem at its worst levels would be alleviated, of course, by avoiding observations in the regions of relatively small fringe amplitude.

Given the preceding analysis for a double-point source, one can compare the magnitude of the delay effect due to the frequency dependence of the distribution (τ_D) with that due to the frequency dependence of the sky fringes (τ_p). A comparison of Eqs. (91) and (94) indicates that τ_p is about an order of magnitude larger than τ_D. This result suggests that one measurement of the brightness distribution at the observing frequency will be adequate to remove most of the structure delay and that a second brightness measurement to determine frequency dependence will be necessary only for the most accurate applications, if at all.

In the same example of a double-point source, the effect of structure on delay rate can be obtained by taking the partial of ϕ_B in Eq. (77) with respect to time:

$$\dot{\Delta\tau}_\phi = -\frac{v_m}{1+h^2}\,\frac{\dot{\vec{B}}_s \cdot \vec{P}_1}{c}\,\sec^2(2\pi\vec{B}_s \cdot \vec{P}_1/\lambda) + \dot{\vec{B}}_s \cdot \vec{P}_0/c \tag{96}$$

where we have assumed the brightness distribution is independent of time. Since the delay rate observable is much less important than the BWS delay as a consequence of larger measurement errors, we will not present numerical examples of delay rate. However, comparison of Eqs. (86) and (96) will reveal that, with respect to occurrence of maxima, the behavior of $\dot{\tau}_\phi$ will be similar to the BWS delay. When converted to equivalent baseline error, the sizes of these maxima are of the same order of magnitude as the maxima in τ_p in Fig. 7.

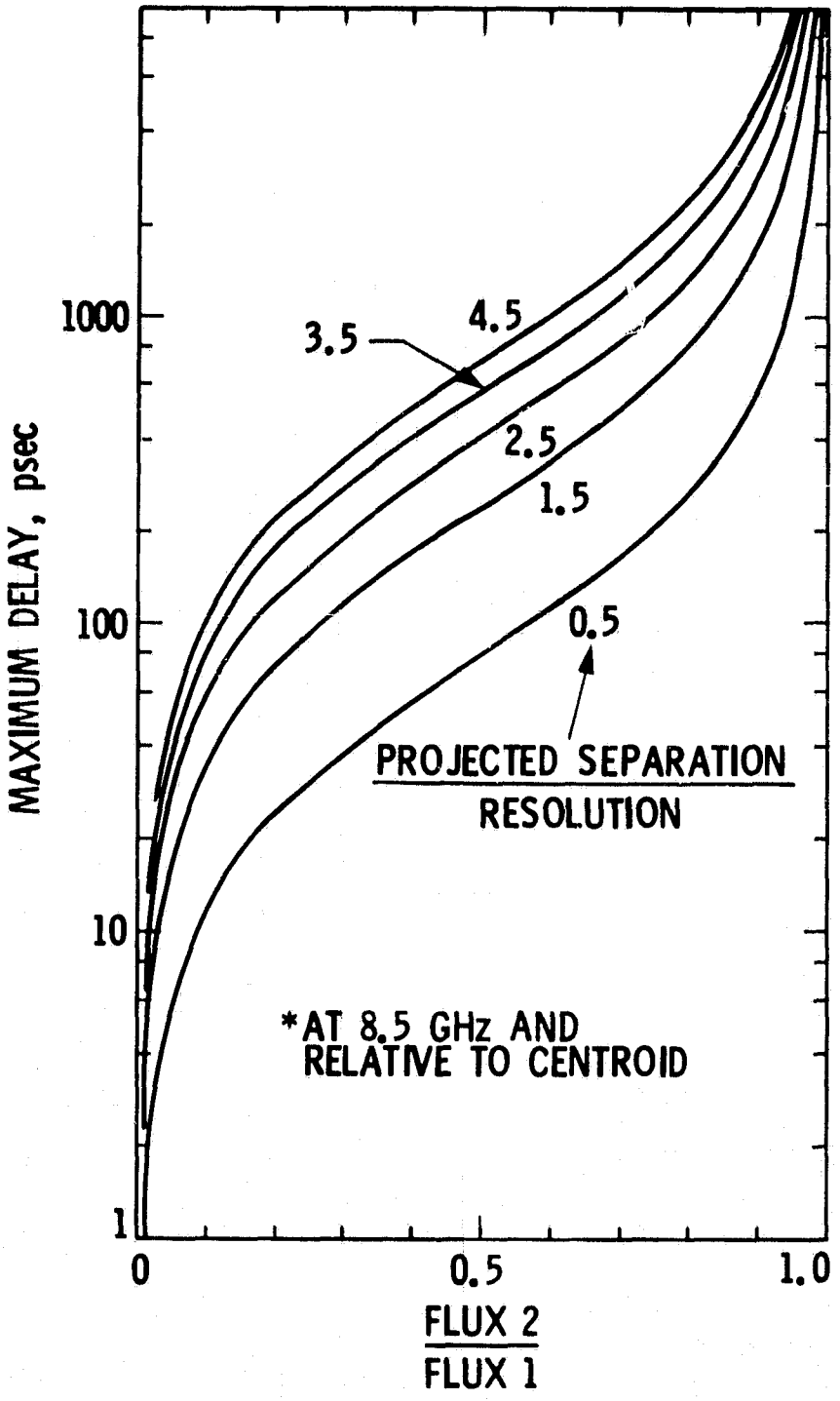


Figure 8. Maximum BWS Delay* for Frequency-Independent Double-Point Source

By setting the derivative in Eq. (96) equal to zero and solving for $\vec{P}_0$, we obtain another condition on a component of $\vec{P}_0$:

$$\dot{\vec{B}}_s \cdot \vec{P}_0 = v_m \frac{\sec^2 (2\pi \vec{B}_s \cdot \vec{P}_1/\lambda)}{1+h^2} \dot{\vec{B}}_s \cdot \vec{P}_1 \tag{97}$$

Both BWS delay and delay rate are normally used together, in which case Eq. (88) and Eq. (97) can be used to determine uniquely the effective position P_0. As explained in preceding sections for a general distribution, that procedure is usually somewhat involved if the frequency dependence of the brightness distribution is not negligible ($\tau_D \neq 0$). When such dependence is negligible, the effective position has been shown to be equal to the centroid of the resolved distribution. In the present example, this centroid is easily obtained by inspecting Eq. (97) and the second term in Eq. (88) which, along with Eq. (87), gives

$$\langle \vec{P} \rangle_R = \frac{v_m \vec{P}_1}{\cos^2(2\pi \vec{B}_s \cdot \vec{P}_1/\lambda) + v_m^2 \sin^2 (2\pi \vec{B}_s \cdot \vec{P}_1/\lambda)} \tag{98}$$

As emphasized after Eq. (71), this position calculation is referenced to the midpoint between point sources. As one would expect, the centroid lies on the line connecting the two sources, but is a rather complicated function of the source strengths and resolution geometry. Three special cases are of interest. For various values of the minimum fringe visibility, it is easily shown that

$$\langle \vec{P} \rangle = 0 \quad \text{where } g_1 = g_2 \ [v_m = 0 \text{ and } \theta \neq (2m+1)\pi] \tag{99}$$

$$\langle \vec{P} \rangle = \vec{P}_1 \text{ when } g_1 >> g_2 \ (v_m = 1) \tag{100}$$

$$\langle \vec{P} \rangle = -\vec{P}_1 \text{ when } g_1 << g_2 \ (v_m = -1) \tag{101}$$

$$\langle \vec{P} \rangle = \vec{P}_c \text{ when } \vec{B}_s = 0 \tag{102}$$

as one would guess.

To compute numerical examples of the effective position, it is convenient to rewrite the components of Eq. (98) in the form

$$\frac{\langle\beta\rangle}{S} = \frac{v_m}{2} \frac{1}{\cos^2(\theta/2) + v_m^2 \sin^2(\theta/2)} \tag{103a}$$

$$\langle\gamma\rangle = 0 \tag{103b}$$

where S and θ have been defined above. Without loss of generality, we have placed $\vec{P}_1$ along the β-axis and have expressed results in units of S, the distance between sources. In this form, the result depends only on the strength ratio (g_2/g_1) and on the ratio of projected separation to resolution (θ). Figure 9 plots effective position as a function of θ for a few values of the strength ratio. As can be seen, the effective position is a periodic function of separation/resolution that goes through a maximum whenever separation/resolution = $(2n + 1)/2$ cycles. The function becomes singular at the same points if the point strengths are equal, for the reasons discussed earlier. Even when g_2/g_1 = 0.5, the effective position can move outside the region between sources to the extent that it is separated from the stronger (closer) source by a distance equal to the source separation. Thus, the movement of the effective position is generally much more damaging in the case of BWS delay and delay rate than it is in the case of phase-delay (Fig. 4). This result suggests that, with respect to structure effect, the phase delay will tend to be superior to the BWS delay.

Due to the importance of null points and minimum amplitudes, it is useful to demonstrate their location in the u-v plane for a double-point source. As can be readily discerned from Eq. (75), minimum amplitude will occur on lines for which $\vec{B}_s/\lambda$ satisfies

$$\frac{\vec{B}_s}{\lambda} \cdot \hat{P}_1 = \frac{2n + 1}{2S} \tag{104}$$

where S is the spacing in radians between point sources ($S = 2|\vec{P}_1|$). Figure 10 schematically plots these lines of minimum amplitude in the u-v plane. As indicated by Eq. (104), the lines are perpendicular to $\hat{P}_1$ and have a spacing of S^{-1}, with the first lines separated from the origin by $(2S)^{-1}$. It is clear from this result that, for a double point source, a long u-v track will usually cross a line of minimum amplitude if $\vec{B}_s/\lambda$ is greater than $(2S)^{-1}$. Thus for nearly equal point sources, it will not be unusual in such cases to apply fairly large structure corrections to BWS delay for some observations.

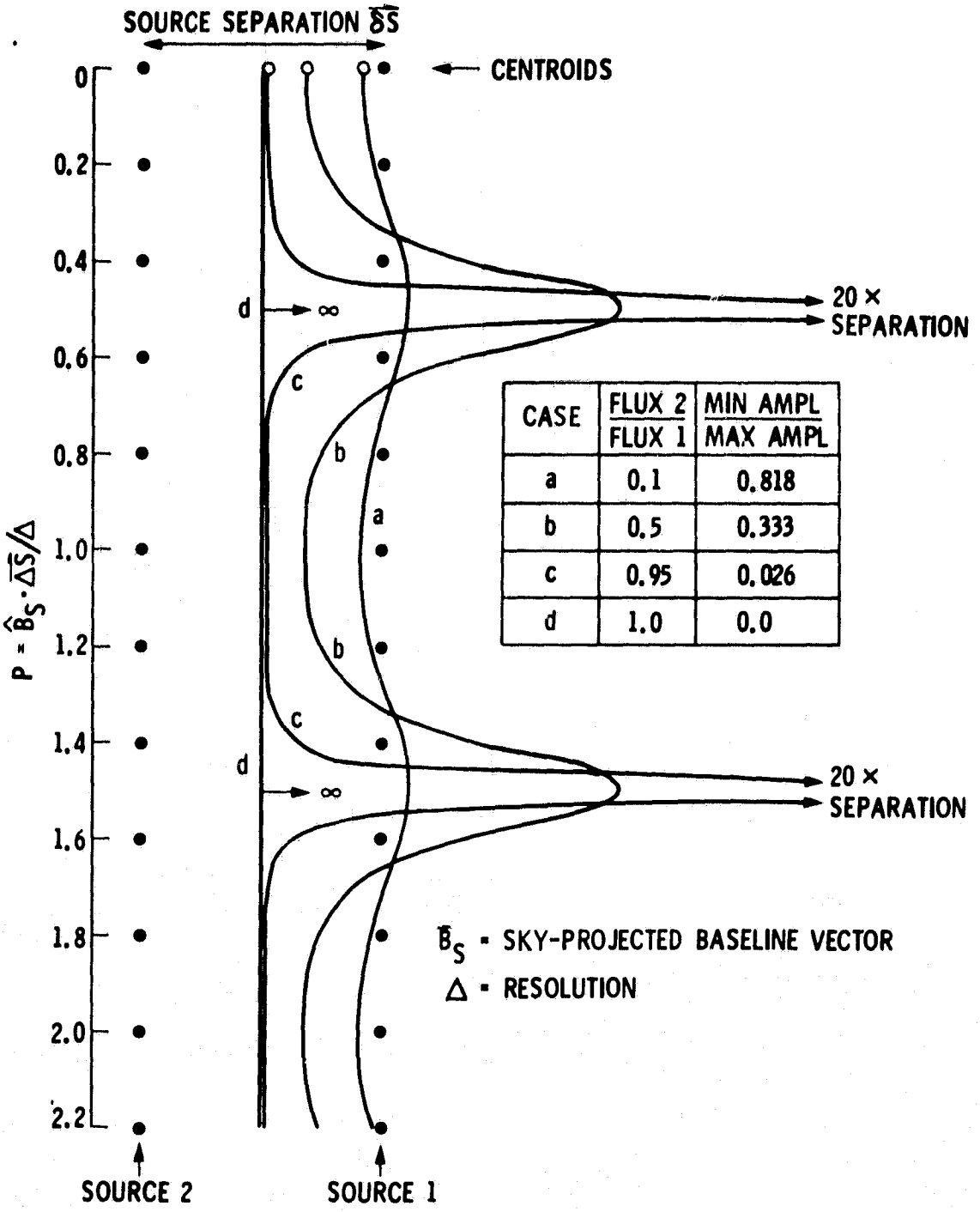


Figure 9. Effective Position for a Frequency-Independent Double-Point Source: BWS Delay and Phase-Delay Rate

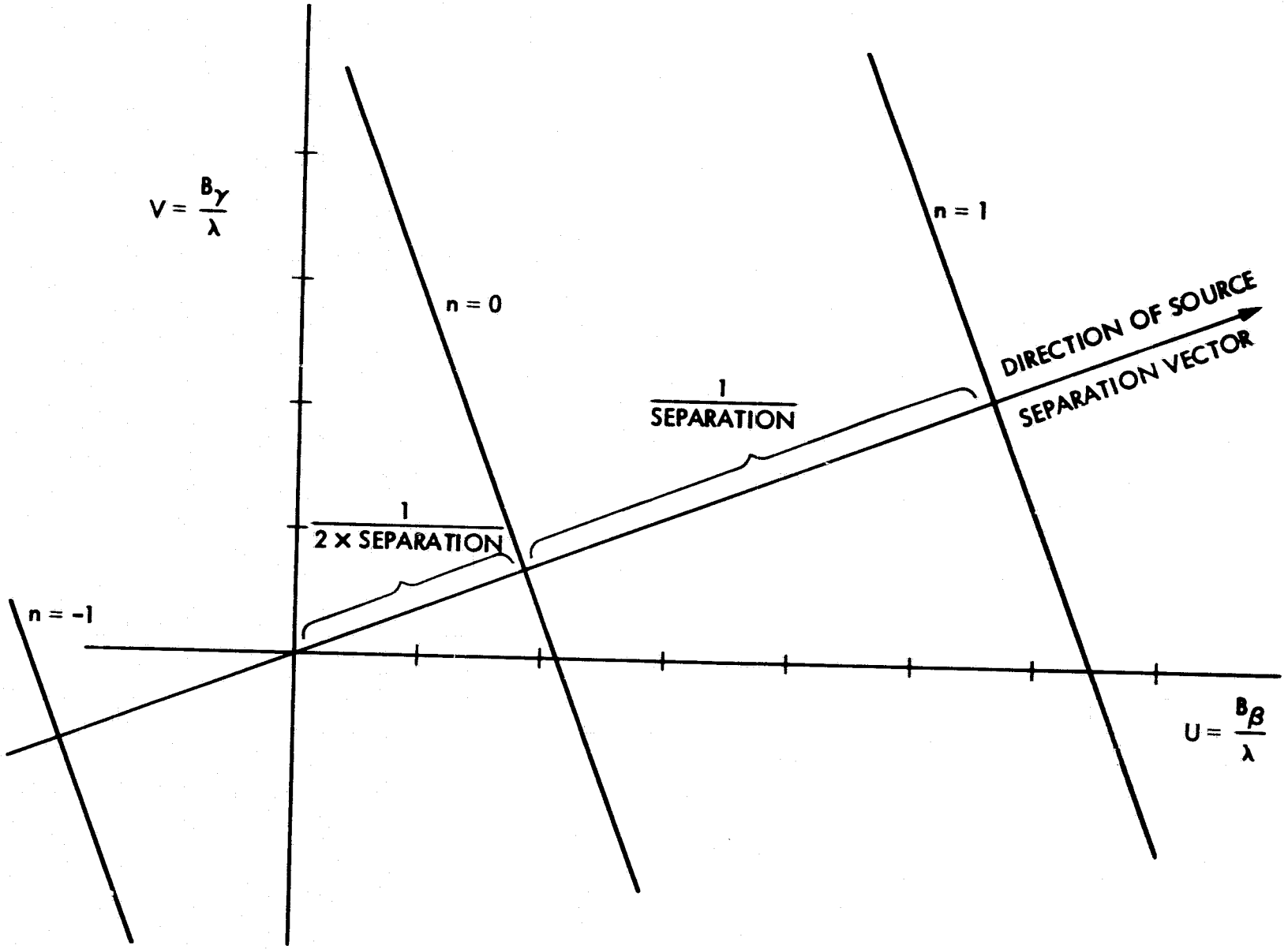


Figure 10. Lines of Minimum Fringe Amplitude for a Double-Point Source

XIII. THE N-POINT SOURCE

This section analyzes the structural effects of a frequency-independent source consisting of N-point sources of arbitrary strength and position. The special case of a triple point source is considered in some detail.

The brightness distribution for an N-point source can be written as

$$D(\vec{P}) = \sum_{k=1}^{N} g_k \, \delta(\vec{P} - \vec{P}_k) \tag{105}$$

where g_k and $\vec{P}_k$ are the strength and position of point source k. For this distribution, the brightness transform in Eq. (29) is given by

$$R = e^{i\omega\vec{B}_s \cdot \vec{P}_o/c} \sum_{k=1}^{N} g_k \, e^{-i\omega\vec{B}_s \cdot \vec{P}_k/c} \tag{106}$$

for which the structure phase becomes

$$\phi_B = \tan^{-1}(-Z_s/Z_c) + \omega\vec{B}_s \cdot \vec{P}_0/c \tag{107}$$

where

$$Z_c = \sum_k g_k \cos \theta_k \tag{108}$$

$$Z_s = \sum_k g_k \sin \theta_k \tag{109}$$

$$\theta_k = \omega\vec{B}_s \cdot \vec{P}_k/c \tag{110}$$

The fringe visibility becomes

$$v_s = \left[\left(\sum_k g_k \cos \theta_k\right)^2 + \left(\sum_k g_k \sin \theta_k\right)^2\right]^{1/2} \Big/ \left(\sum_k g_k\right) \tag{111}$$

The BWS delay due to structure is obtained through the use of Eq. (45) with $\Delta L_\omega = 0$, which gives

$$\Delta\tau_B = -\left[\langle\vec{P}\rangle_N - \vec{P}_0\right] \cdot \vec{B}_s/c \tag{112}$$

where the effective position is given by

$$\langle\vec{P}\rangle_N = \frac{1}{Z_c^2 + Z_s^2} \sum_k (Z_c \cos \theta_k + Z_s \sin \theta_k)\, g_k \vec{P}_k \tag{113}$$

Thus the effective position is equal to a linear combination of the point-source position vectors $\vec{P}_k$ where the coefficients are complicated functions of source and resolution geometry. For the special case of N colinear point sources, this result shows that the effective position lies on the line joining the sources, as one would expect. The special case of two point sources can be used as a check by showing Eq. (113) reduces to Eq. (98) for N = 2.

Unlike sources for which N is greater than 3, a triple point source can be characterized by a sufficiently small number of variables to allow a reasonably general display of results. As indicated in Eq. (106), the brightness transform for a triple point source will consist of the following sum of three phasors:

$$R = 1 + g_2\, e^{-i\theta_2} + g_3\, e^{-i\theta_3} \tag{114}$$

where, for simplicity, we have used the strength and position of point source 1 as reference by setting $g_1 = 1$ and $\vec{P}_0 = \vec{P}_1$. Without loss of generality, we will assume point source 1 is the strongest source. Thus, four variables can be used to parametrize a triple point source: g_2, g_3, θ_2, θ_3. Figure 11a schematically illustrates the formation of the brightness transform from the three phasors. As the baseline changes for a given source, the directions of the phasors (i.e., θ_2 and θ_3) will change while, of course, their lengths (i.e., point strengths) remain constant. The resulting length (amplitude) and orientation of the sum phasor will change as a function of θ_2 and θ_3. This parameterization in terms

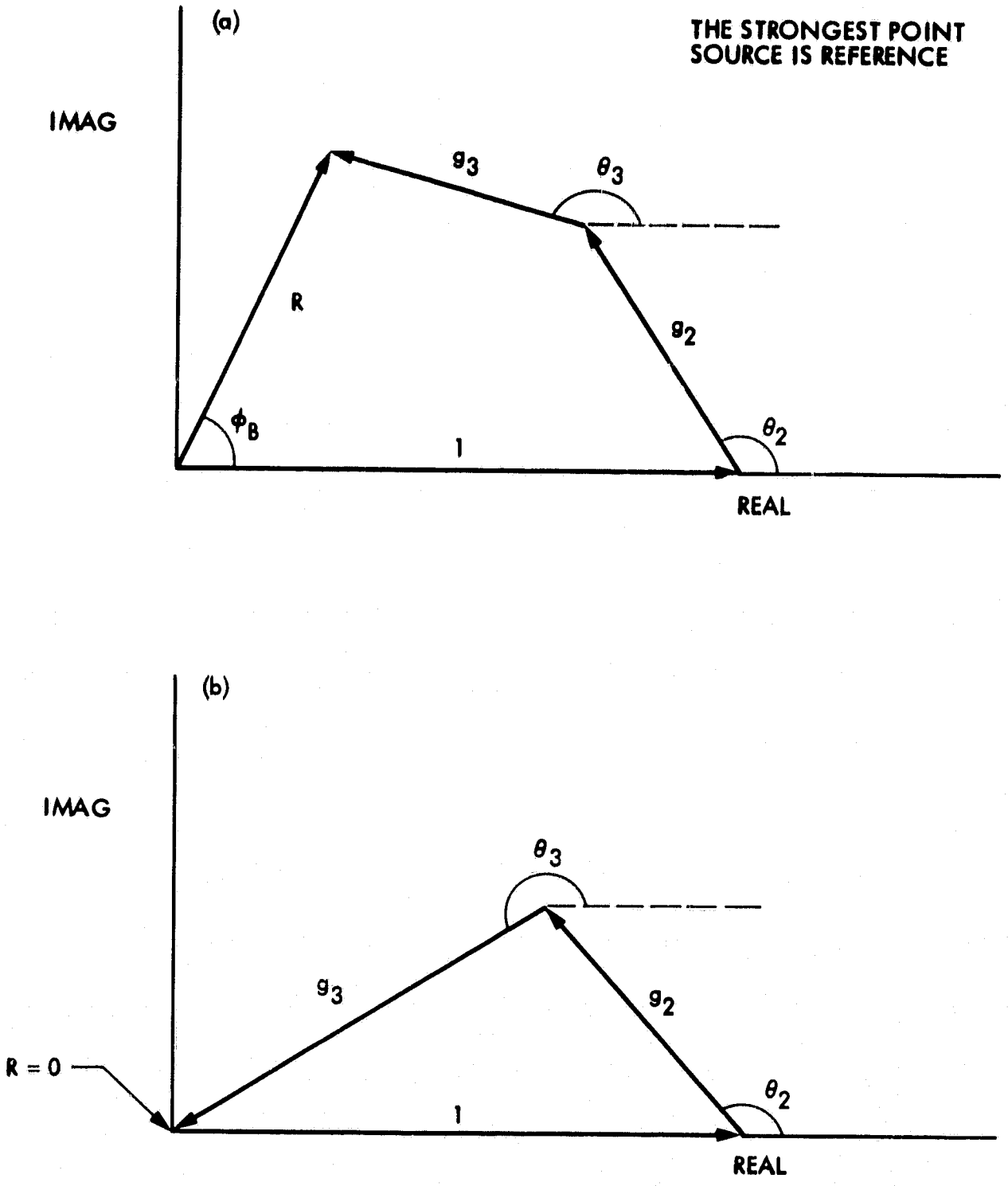


Figure 11. Schematic Example of (a) the Formation of the Brightness Transform for a Triple-Point Source and (b) Zero Fringe Amplitude for a Triple-Point Source

of (θ_2, θ_3) will be general in that the resulting phase and delay equations will be valid for any values of source positions and baseline. To use the result, one merely makes the conversion

$$\theta_2 = 2\pi \frac{\vec{B}_s}{\lambda} \cdot \vec{P}_2 \tag{115}$$

$$\theta_3 = 2\pi \frac{\vec{B}_s}{\lambda} \cdot \vec{P}_3 \tag{116}$$

where $\vec{B}_s/\lambda$ and $(\vec{P}_2, \vec{P}_3)$ are the given values for the baseline vector (i.e., u, v) and the source positions, respectively.

Since large delay effects occur where the amplitude of the brightness transform approaches zero, we will first determine where points of zero amplitude occur. Zero amplitude occurs whenever the phasors sum to zero as shown in Fig. 11b. For given values of point strength (g_2, g_3), there are unique values of (θ_2, θ_3) for which such closure will occur. One can easily show that those values are given by

$$\overset{\sim}{\theta}_2 = \pm\cos^{-1}\left(\frac{g_3^2 - g_2^2 - 1}{2g_2}\right) + 2m\pi \qquad m = 0, \pm 1, \pm 2, \ldots \tag{117}$$

$$\overset{\sim}{\theta}_3 = \mp\cos^{-1}\left(\frac{g_2^2 - g_3^2 - 1}{2g_3}\right) + 2n\pi \qquad n = 0, \pm 1, \pm 2, \ldots \tag{118}$$

where the leading plus (minus) sign in the second equation is used if the leading minus (plus) sign is used in Eq. (117). Such null points can occur only if the two weakest point sources have a combined strength at least as strong as the strongest source (i.e., $g_2 + g_3 \geq 1$). Figure 12 schematically demonstrates how the locations of these null points can be calculated in the u-v plane. In the plane of the sky at the source, draw two new axes in the directions of point sources (2, 3) starting at point source 1 (i.e., along $\vec{P}_2$, $\vec{P}_3$). Then for point source 2, draw perpendiculars to the $\vec{P}_2$ axis that intersect that axis at distances from the u, v origin equal to

$$\frac{\vec{B}_s}{\lambda} \cdot \hat{P}_2 = \frac{1}{|\vec{P}_2|} \frac{\overset{\sim}{\theta}_2(m)}{2\pi}, \quad m = 0, \pm 1, \pm 2, \ldots \tag{119}$$

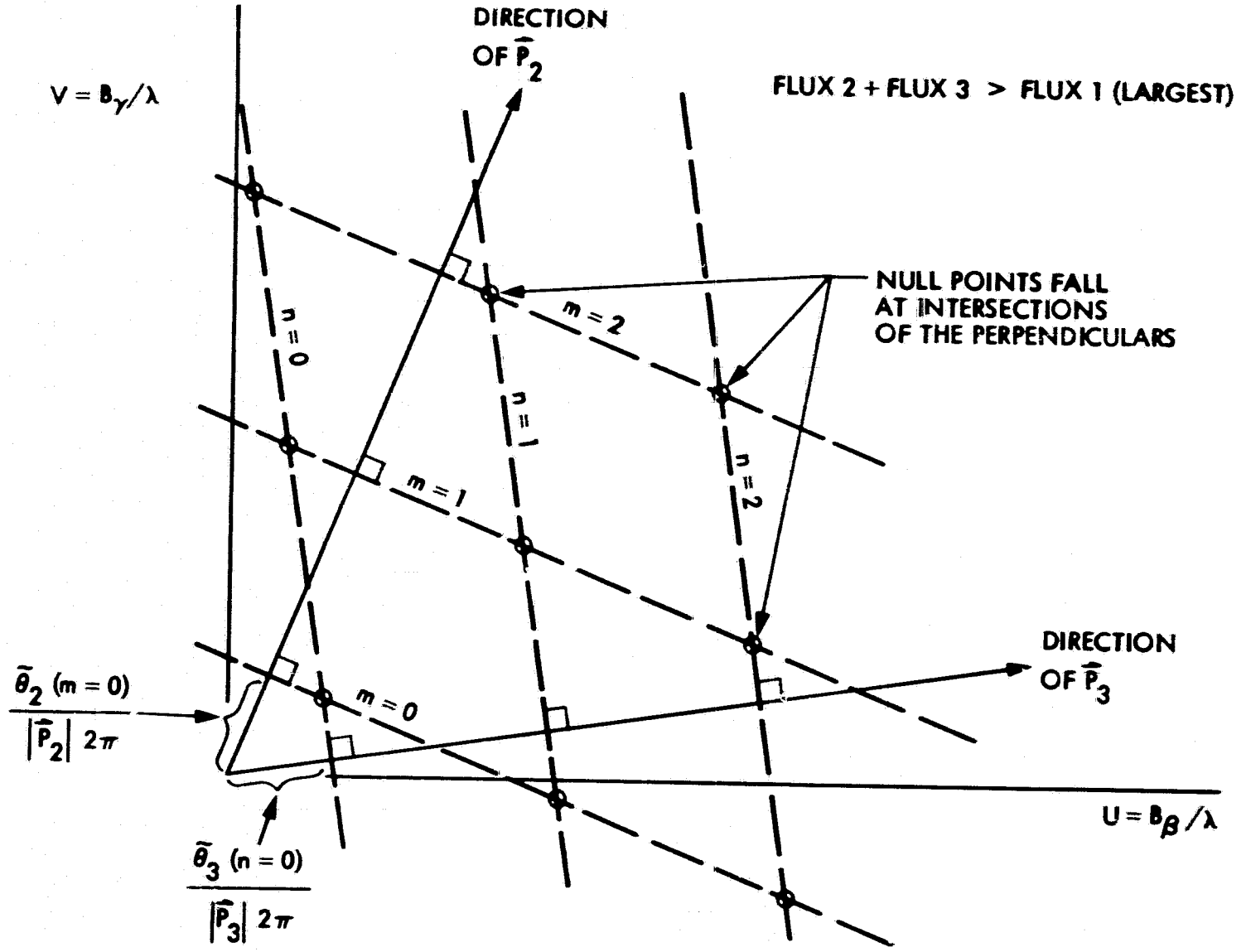


Figure 12. Schematic Example of the Locations of u-v Points of Zero Fringe Amplitude for a Triple-Point Source

Next, draw similar lines perpendicular to the $\vec{P}_3$ axis, based on θ_3, etc. The intersections of these two sets of perpendiculars are the location of the null points in the u-v plane. As can be seen, the null points occur at discrete points on a regular lattice. For simplicity, the figure shows only the null points associated with the leading plus sign in Eq. (117). The other set of null points that comes from the leading minus sign is also latticelike but is generally not coincident with the first set.

The null-point behavior of a triple-point source is to be contrasted with the result for a double-point source for which the zero amplitudes fall on lines in the u-v plane (see Fig. 10). For a double-point source, the u-v points of minimum amplitude depend only on the source separation vector and not on point strengths. On the other hand, for a triple-point source (and for N > 3), the locations of zero amplitude can depend on both the locations and the strengths of the point sources. For a triple-point source, the first null points to occur as baseline length B_s increases will satisfy the inequality:

$$\frac{B_s}{\lambda} \geq \frac{0.3}{|\vec{P}_m|} \tag{120}$$

where $|\vec{P}_m|$ is either $|\vec{P}_2|$ or $|\vec{P}_3|$.

To obtain fairly general results for structure effects for a triple point source, we will parameterize amplitude and phase as suggested by Eq. (114):

$$|R|^2 = (1+g_2 \cos\theta_2 + g_3 \cos\theta_3)^2 + (g_2 \sin\theta_2 + g_3 \sin\theta_3)^2 \tag{121}$$

$$\tan\phi_B = \frac{-g_2 \sin\theta_2 - g_3 \sin\theta_3}{1+g_2 \cos\theta_2 + g_3 \cos\theta_3} \tag{122}$$

where $g_k < 1$. The BWS delay can be obtained as a function of the same four parameters by either numerically or analytically differentiating Eq. (122) with respect to ω. For plots, the following numerical computation has been performed:

$$\Delta\tau_1 \equiv \frac{\partial\phi_B}{\partial\omega} \approx \frac{\phi_B(\theta'_2,\theta'_3) - \phi_B(\theta_2,\theta_3)}{\varepsilon\omega} \tag{123}$$

where

$$\theta'_k = (1+\varepsilon)\,\theta_k \tag{124}$$

and ϵ is a very small number. An analytic expression for $\Delta\tau_1$ can also be readily derived in terms of $(g_2, g_3, \theta_2, \theta_3)$, but that result will not be presented here due to its uninformative complexity. Due to the definition given to $\vec{P}_0$ and $\vec{P}_1$ in forming ϕ_B, this BWS delay will be relative to the strongest point source. To change the reference point for BWS delay to the brightness centroid, one must set $\vec{P}_0$ in Eq. (106) equal to the centroid given by

$$\vec{P}_c = \frac{g_2 \vec{P}_2 + g_3 \vec{P}_3}{1 + g_2 + g_3} \tag{125}$$

The phase shift due to this change in $\vec{P}_0$ becomes

$$\phi_c = \omega \vec{B}_s \cdot \vec{P}_c / c \tag{126}$$

$$= \frac{g_2\theta_2 + g_3\theta_3}{1 + g_2 + g_3} \tag{127}$$

Since (θ_2, θ_3) are proportional to ω, the BWS delay due to ϕ_c will be given by

$$\Delta\tau_{1c} = \frac{\partial\phi_c}{\partial\omega} = \frac{1}{\omega} \frac{g_2\theta_2 + g_3\theta_3}{1+g_2 + g_3} \tag{128}$$

Thus the total structure delay becomes

$$\Delta\tau_B = \Delta\tau_1 + \Delta\tau_{1c} \tag{129}$$

when referenced to the brightness centroid, where $\Delta\tau_1$ and $\Delta\tau_{1c}$ are obtained from Eqs. (123) and (128).

Before proceeding to example plots, it is useful to derive an expression for the extrema in BWS delay as a function of θ_2 and θ_3. As discussed above, when $g_1 + g_2 > 1$, the amplitude will go to zero for some values of (θ_2,θ_3), and the BWS delay will be infinite at those points. For $g_1 + g_2 < 1$, the BWS delay does not become infinite but peaks for certain values of (θ_2,θ_3). The extrema in BWS delay at those points can be computed as follows. A consideration of phasors shows that, when $g_2 + g_3 < 1$, the most rapid change in ϕ_B as a function of θ_2, θ_3 occurs when

$$\theta_2 \approx (2m + 1)\pi \quad m = 0, \pm 1, \pm 2, \ldots \tag{130}$$

$$\theta_3 \approx (2n + 1)\pi \quad n = 0, \pm 1, \pm 2, \ldots \tag{131}$$

The extrema of BWS delay will occur near these points. To obtain approximate values of those extrema, Eq. (129) can be analytically evaluated as a function of (θ_2, θ_3) and computed for the specific values of (θ_2, θ_3) in Eqs. (130) and (131). Such a calculation gives

$$\Delta\tau_B(m,n) \approx \frac{2\pi}{\omega} \frac{g_2(2m + 1) + g_3(2n + 1)}{1 - (g_2 + g_3)^2} \tag{132}$$

where it is understood that $g_2 + g_3 < 1$. Figure 13 plots $\Delta\tau_B$ as a function of (g_2, g_3) at X-band for the first extremum, $m = n = 0$. From this example plot, one sees that $\Delta\tau_B = 0.26$ nsec for the particular case $g_1 = g_2 = 0.4$. As we saw for the case of a double-point source, structure effects can also be considerable for triple-point sources.

For examples, Figs. 14 through 22 display amplitude, phase and BWS delay as a function of (θ_2, θ_3) for two cases of (g_2, g_3). The axes ($p = \theta_2/2\pi$ and $q = \theta_3/2\pi$) are in units of cycles rather than radians and are equal to the projections of the two separation vectors in the direction of the sky-projected baseline vector, with both in units of resolution (i.e., $\hat{B}_s \cdot \vec{P}_k/\Delta$). Interpretation of these plots in terms of phasor sums such as those in Fig. 11 will be left to the reader. To help clarify results, contour and surface plots of BWS delay are both shown. Even though the conversion of (θ_2, θ_3) to (u, v) might complicate geometric interpretation of these results, the values for the extrema in BWS delay shown in the plots will not change with variable transformation or with changes in $(\vec{P}_2, \vec{P}_3)$. Thus, given only point strengths (g_2, g_3), one can readily determine for a triple-point source the extremal BWS delays (relative to the brightness centroid) either by inspecting plots such as Fig. 14 or by evaluating Eq. (132).

As shown in Figs. 14 and 15, the structure delay for the case $g_2 = 0.2$ and $g_3 = 0.4$ passes through an extremum in delay whenever p and q are both multiples of 1/2, as explained earlier. The first extremum at $p = q = 0.5$ has a magnitude of about 0.1 nsec. A check on results for $g_2 + g_3 < 1$ is supplied by a comparison of the extremal values predicted by Eqs. (130) - (132) with the results numerically computed to obtain these figures. The extrema in delay correspond to minima in fringe visibility shown in Fig. 16. These minima are all equal to about 0.25. Structure phase for this example is given in Fig. 17 and ranges between $\pm$ 0.1 cycles.

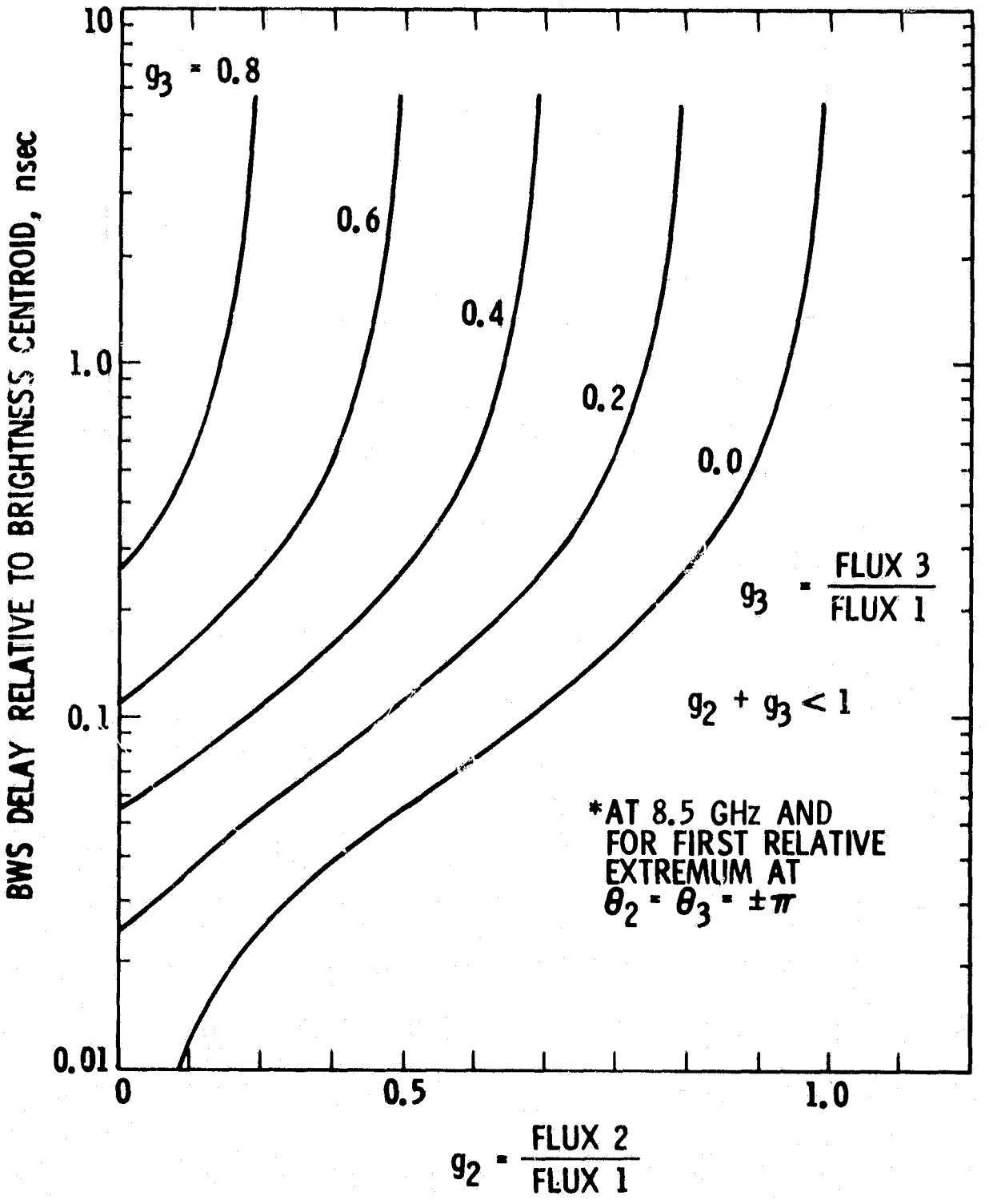


Figure 13. Magnitude of First Relative Extremum in BWS Delay* for a Triple-Point Source

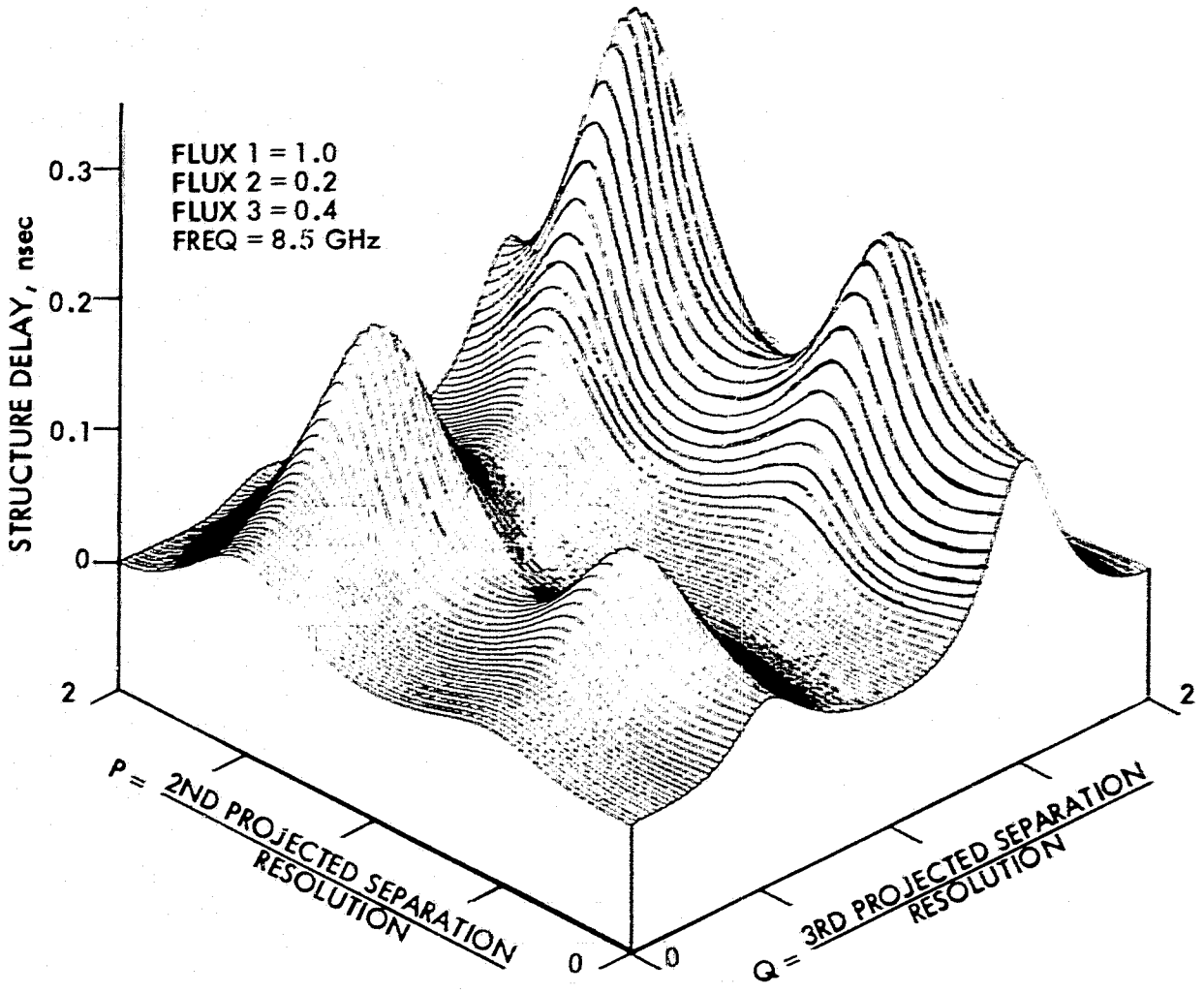


Figure 14. BWS Delay Relative to Brightness Centroid for a Triple-Point Source, Case 1

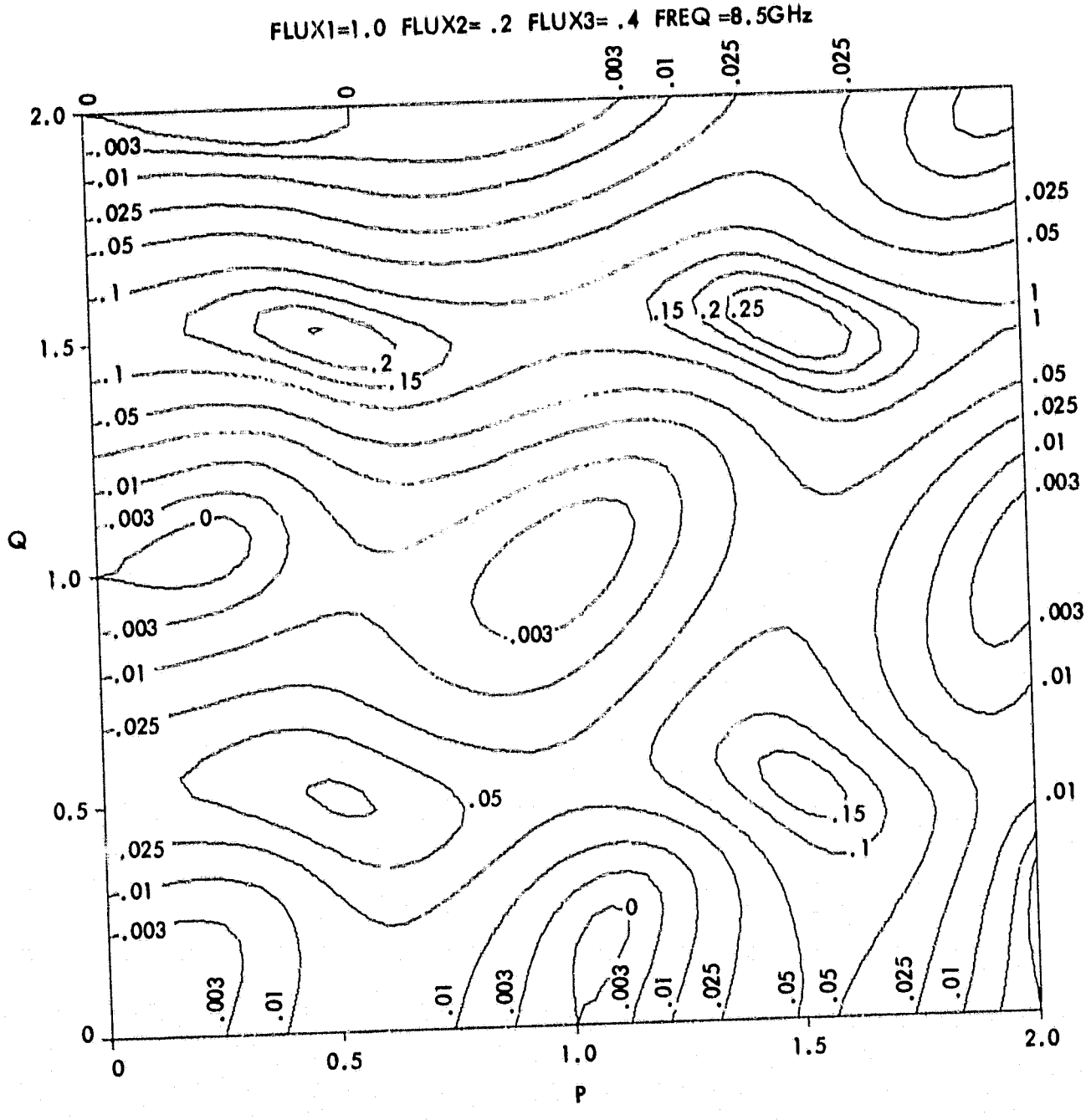


Figure 15. Contour Plot of BWS Delay Relative to Brightness Centroid for a Triple-Point Source, Case 1

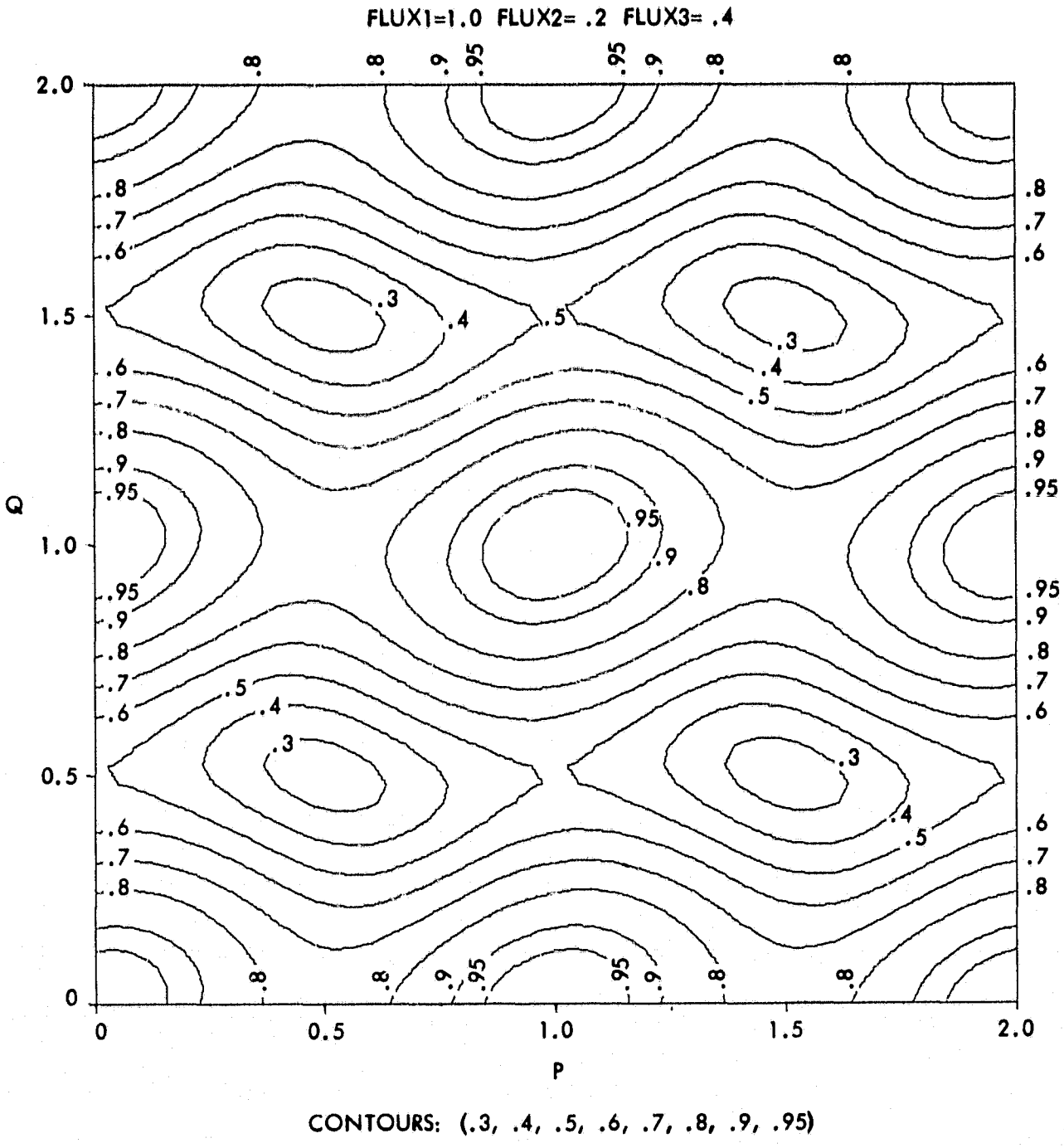


Figure 16. Fringe Visibility for a Triple-Point Source, Case 1

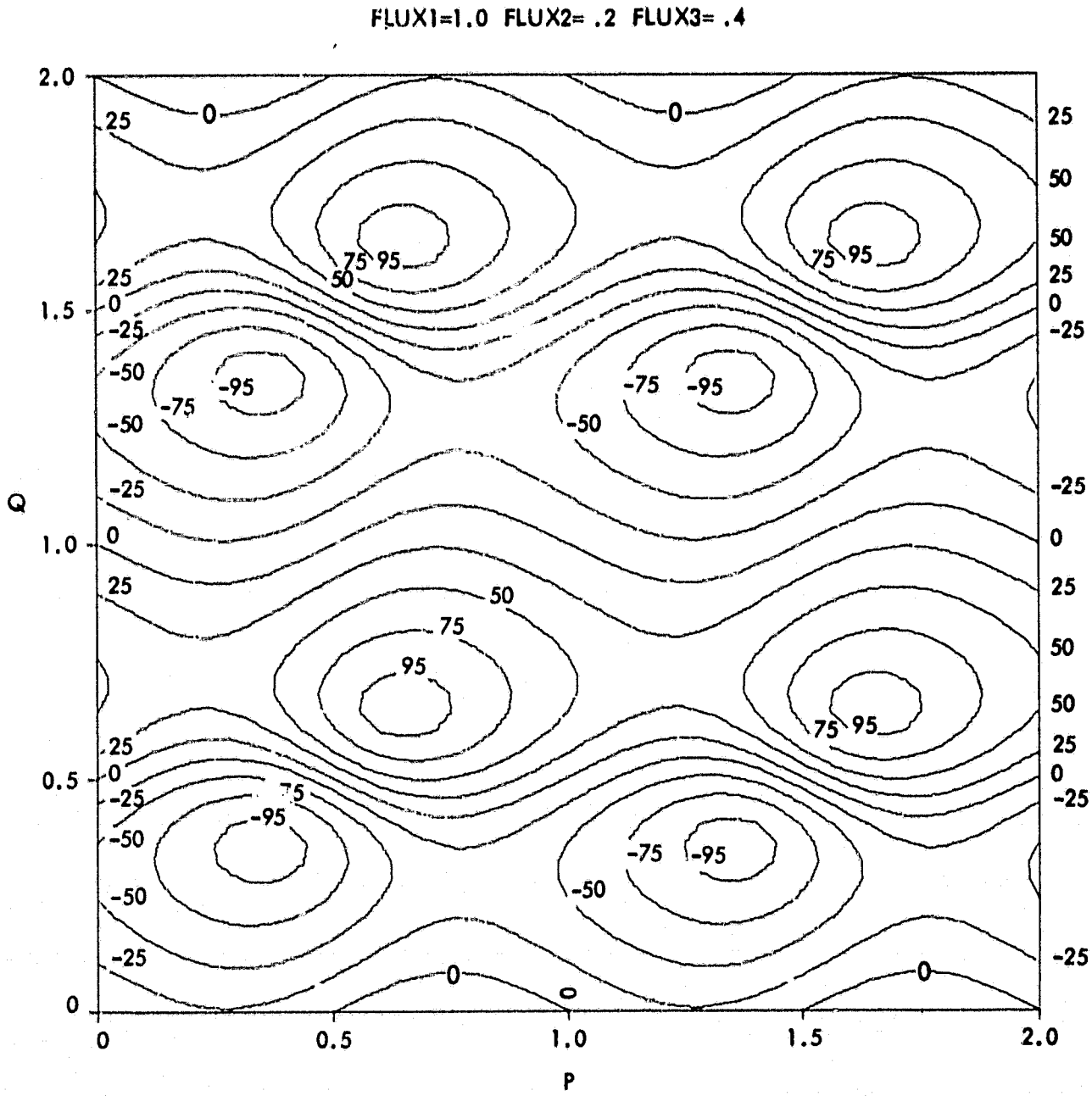


Figure 17. Structure Phase Relative to Strongest Point for a Triple-Point Source, Case 1

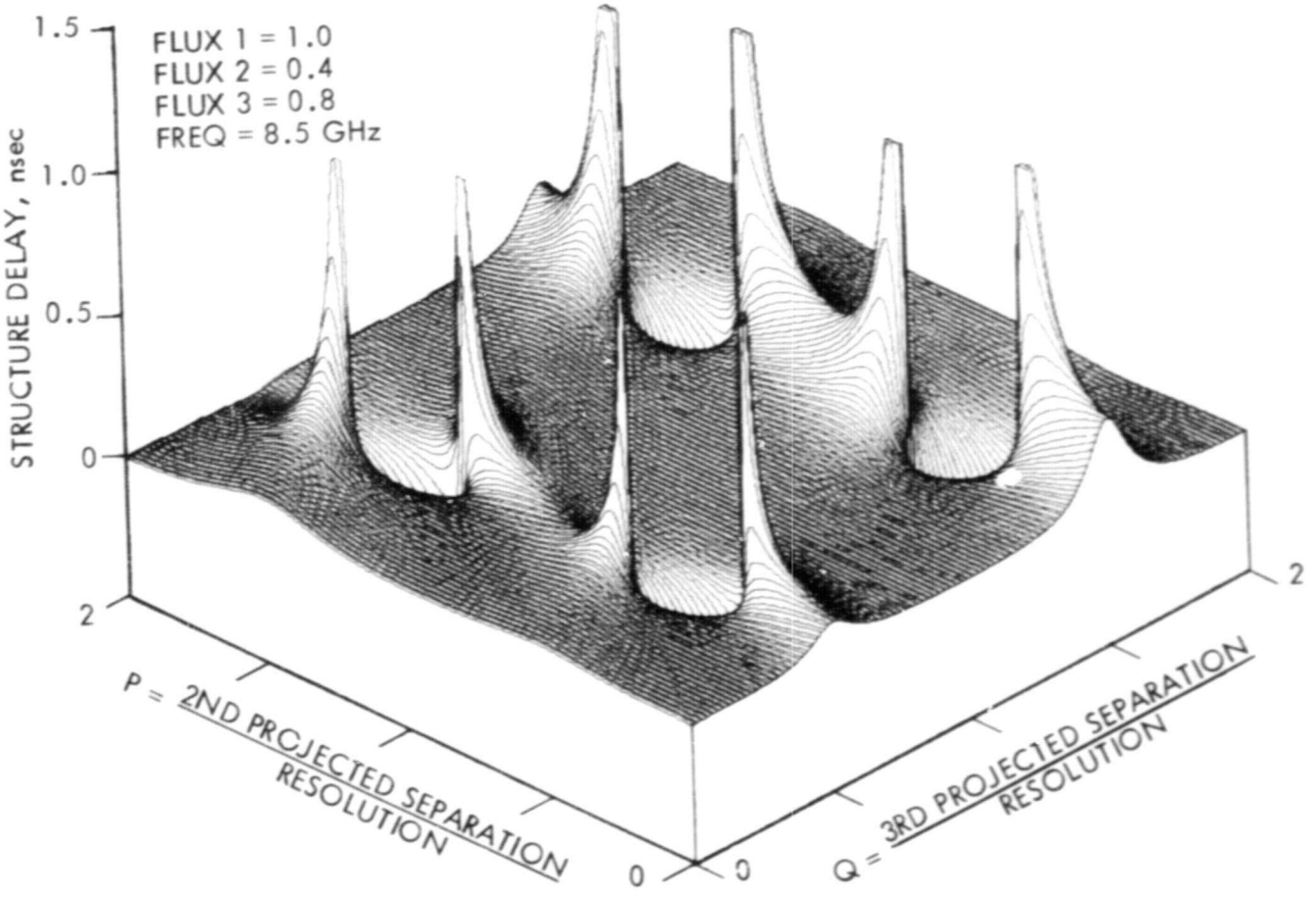


Figure 18. BWS Delay Relative to Brightness Centroid for a Triple-Point Source

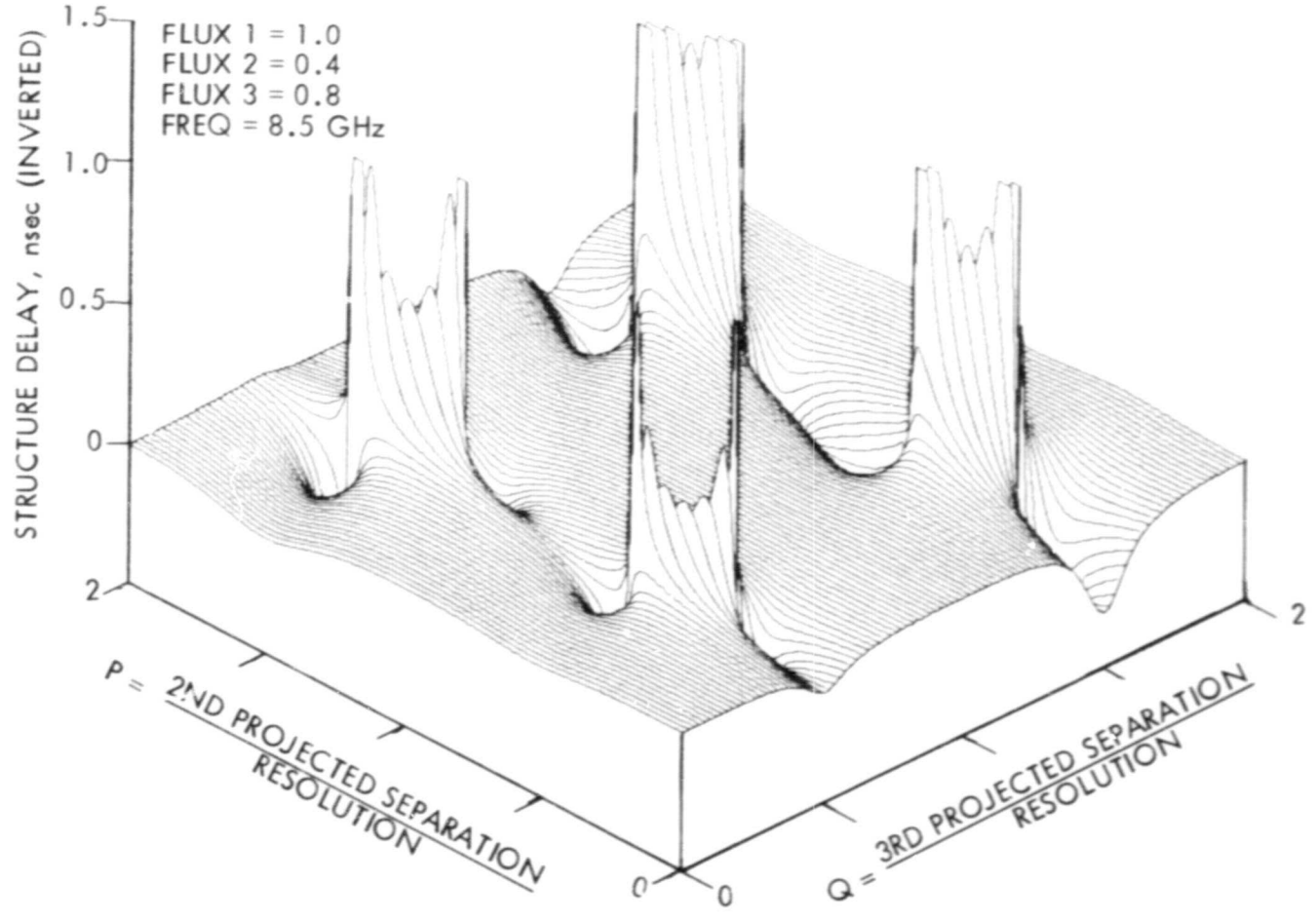


Figure 19. Inverted BWS Delay Relative to Brightness Centroid for a Triple-Point Source, Case 2

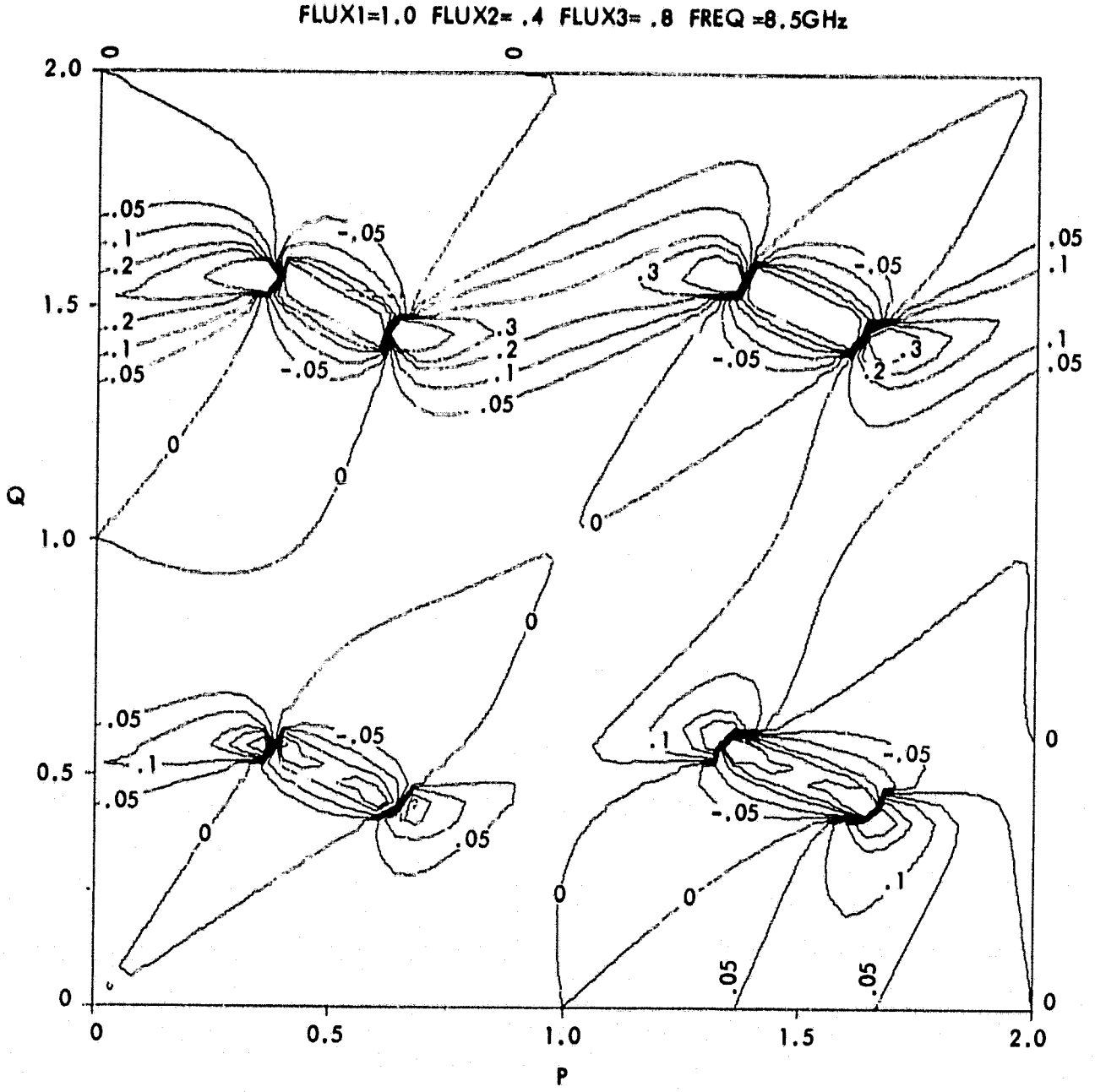


Figure 20. Contour Plot of BWS Delay Relative to Brightness Centroid for a Triple-Point Source, Case 2

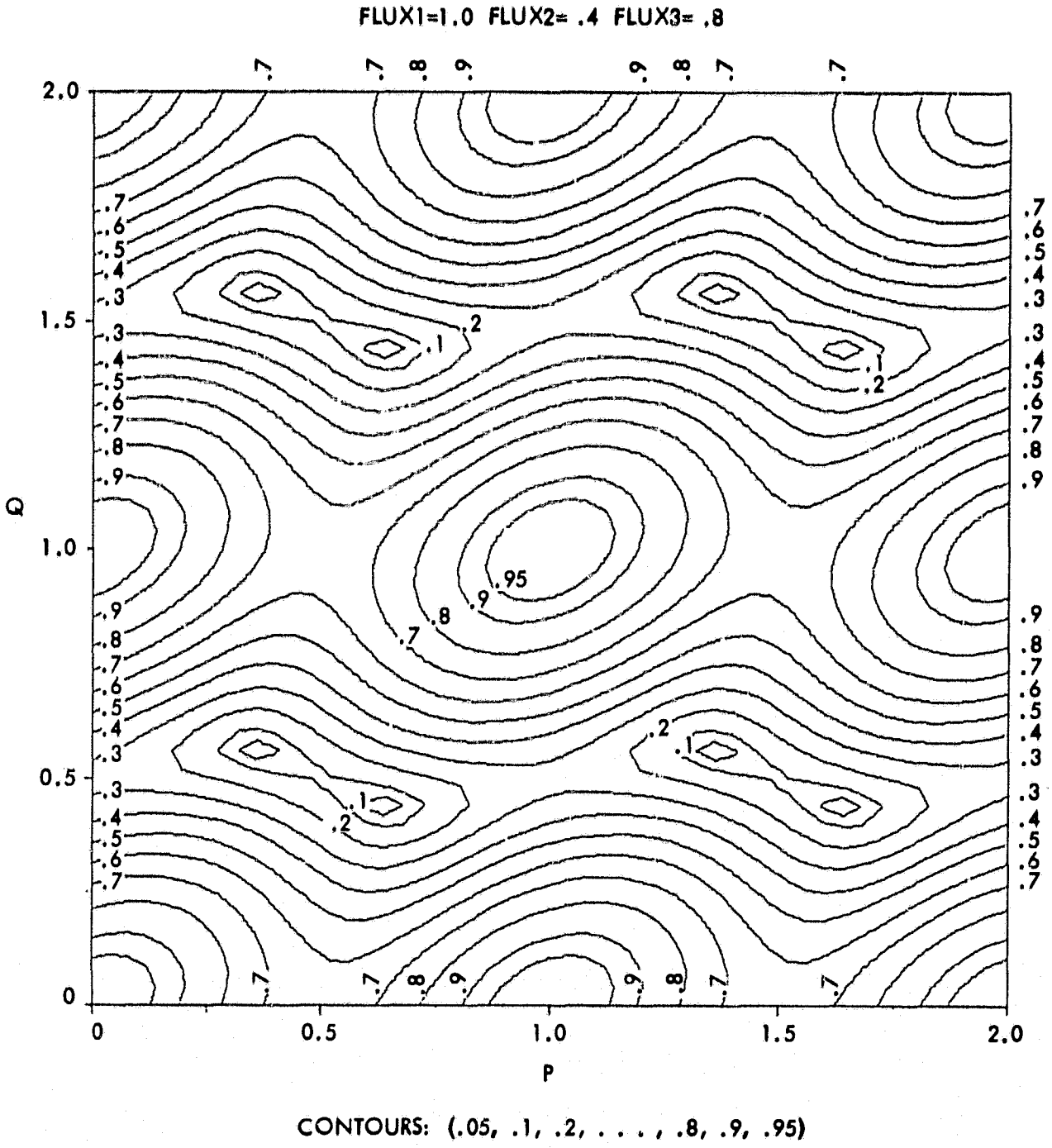


Figure 21. Fringe Visibility for a Triple-Point Source, Case 2

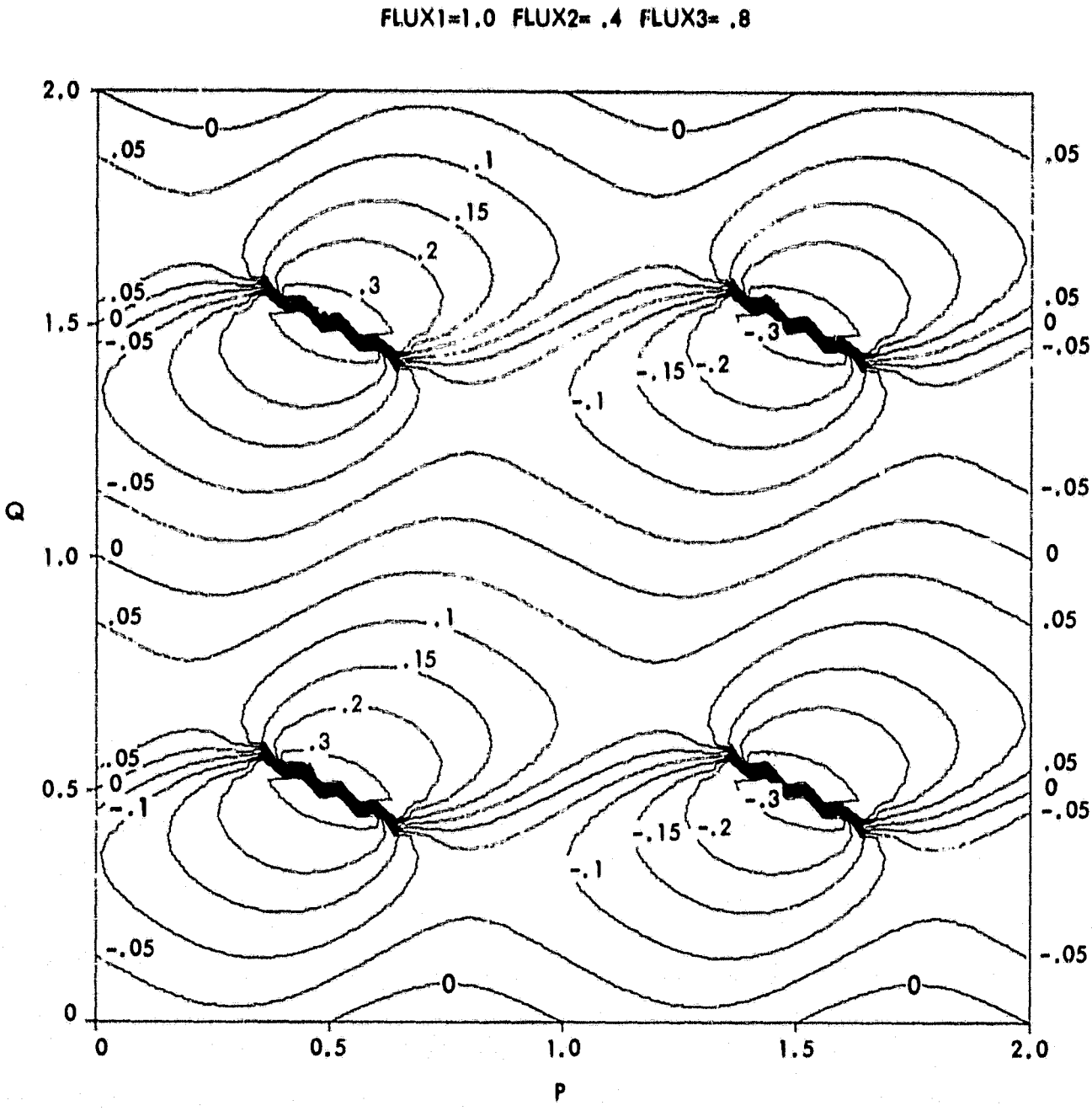


Figure 22. Structure Phase Relative to Strongest Point for a Triple-Point Source, Case 2

The second example is the case $g_2 = 0.4$ and $g_3 = 0.8$. The delay values for this case are plotted in Figs. 18, 19 and 20. These figures are all plots of the same function and are given together to help clarify the rather complicated behavior of delay. The surface plots in Figs. 18 and 19 are truncated above 1 nsec to remove infinite spikes. When $g_2 + g_3 > 1$, the amplitude will always be zero for some baseline vectors, as explained earlier. Null amplitudes in this example occur at adjacent pairs of points that are regularly placed in the p-q plane, as shown in Figure 21. Each pair corresponds to the $\pm$ signs in Eqs. (117) and (118). For each null point, the delay approaches plus infinity on one side of the point and minus infinity on the other, as shown in Figs. 19 and 20. Structure phase for this example is shown in Fig. 22. The heavy jagged lines in that figure indicate an integer-cycle jump in phase from -0.5 cycle to 0.5 cycle. This discontinuous behavior has no ultimate significance and exists in this form due to the decision to keep phase between ± 0.5 cycle, as discussed earlier. Given that BWS delay is the partial of phase with respect to frequency, it is an interesting exercise to predict the plotted delay behavior on the basis of the structure phase plotted in Fig. 22.

This example suggests that, if a source can be exactly represented by a triple-point source, it will be necessary to avoid certain critical values of baseline vector (provided the point sources are sufficiently far apart). Since the fringe visibility becomes very small at these points, it would be possible to alleviate this problem by placing a lower limit on allowed values of fringe visibility.

With regard to observed sources, it is unrealistic to assume that a real source can be exactly modeled with three point sources. Actual distributions possess extended components rather than "point" components and/or have additional weaker components or background features. These characteristics make the truly singular behavior of discrete-point sources an unusual occurrence. Nevertheless, real sources can possess regions in which the fringe visibility becomes very small and the structure delay passes through extrema of considerable magnitude. Examples of real sources are given in the next section.

XIV. EXAMPLES OF MEASURED BRIGHTNESS DISTRIBUTIONS

The preceding sections have treated the analytical cases of a symmetric source, a double-point source and a triple-point source. Although analytical examples are instructive, only a thorough study of many actual source distributions can give a complete picture of source structure effects. To begin such an investigation, this section reports on an analysis of ten brightness distributions measured by the CIT VLBI group. For each distribution, fringe visibility, fringe phase, BWS delay and effective position have been computed as a function of the sky-projected baseline vector (u, v), with the last three effects specified relative to the ordinary centroid. Although instructive, this set of sources is not necessarily a typical set since sources specifically selected for structure investigations are often chosen on the basis of unusual characteristics. Another potential difficulty is that the u-v coverage in these measurements is not as complete is some u-v regions as one would like. In such cases, computed structure

effects must be used with caution, particularly BWS delays.

The brightness distributions resulting from interferometry measurements can be represented in a number of forms. As indicated in Appendix A, one representation is a set of delta functions placed on a regular grid. For this representation, the derivation for an N-point source in Section XIII (see Eqs 105-113) can be directly applied but the equations for phase, delay, amplitude and effective position should first be expressed as a function of the usual (u, v) coordinates. This is easily accomplished by rewriting Eq. (110) in the form

$$\theta_k = 2\pi(u\beta_k + v\gamma_k) \tag{133}$$

where (β_k, γ_k) are the plane-of-the-sky coordinates of the k^{th} point, expressed in radians. A similar expression can be written for the $\vec{\omega}\vec{B}_s \cdot \vec{P}_0/c$ term in Eq. (107). Further, the delay expression (Eq. 112) can be rewritten as

$$\Delta\tau_B = -\frac{1}{f}\left[u\left(\langle\beta\rangle_N - \beta_0\right) + v\left(\langle\gamma\rangle_N - \gamma_0\right)\right] \tag{134}$$

where

$$\vec{P}_0 \equiv (\beta_0, \gamma_0) \tag{135}$$

$$\langle\vec{P}\rangle_N \equiv \left(\langle\beta\rangle_N, \langle\gamma\rangle_N\right) \tag{136}$$

The vector $\vec{P}_0$ is the reference position while $\langle P\rangle_N$ is the effective position of the source, where both positions are expressed relative to structure coordinates. In this work, the reference position $\vec{P}_0$ will be set equal to the ordinary centroid, as computed on the basis of the measured distribution. The quantity f is the observing frequency in hertz. Once a given source has been specified by a flux array $[(g_k, \beta_k, \gamma_k),\ k = 1,N]$, the aforementioned equations can be used to compute phase, amplitude, delay and effective position. In the present computations of BWS delay and effective position, the frequency dependence of the distributions will be neglected.

Table 1 summarizes the 10 brightness distributions measured by the Caltech group (Refs. 2 and 3) and gives the source name, date of the measurement, observing frequency, maximum values of u and v, total flux, interferometer stations, and a quantity k_s to be explained below. Because the total number of plots is excessive,

Table 1. Summary of Brightness Distribution Measurements

Distribution number	Source	Measurement date (Reference)	Observing wavelength, cm	UMAX/VMAX $10^{6}\lambda$	Total flux Jy	Interferometer stations	k_s
1	3C119	May 1976 (3)	18	19/12	7.5	1, 2, 3, 4	0.28
2	3C286	May 1976 (3)	18	19/10	13.5	1, 2, 3, 4	0.18
3	3C345	May 1976 (3)	18	19/12	6.5	1, 2, 3, 4	0.21
4	3C454.3	May 1976 (3)	18	19/8	12.0	1, 2, 3, 4	0.13
5	CTA102	May 1976 (3)	18	19/7	6.4	1, 2, 3, 4	0.11
6	3C273	July 1977 (2)	2.8	135/40	41	1, 2, 3, 5	0.31
7	3C345	July 1977 (2)	2.8	140/80	7.5	1, 2, 3, 5	0.23
8	3C273	December 1977 (2)	6	60/15	46.2	1, 2, 3, 4	0.29
9	3C345	December 1977 (2)	6	60/40	7.8	1, 2, 3, 4	0.25
10	3C120	December 1977 (2)	6	60/20	6.6	1, 2, 3, 4	0.18

Station code: 1 = NRAO; 2 = FDVS; 3 = OVRO; 4 = HCRK; 5 = HSTK

complete results for only one of the measured distributions will be presented here. Brightness contours for the example distribution (9 in Table 1) are given in Fig. 23. The structure effects for this distribution are given in Figs. 24a-e in the form of contour plots as a function of u,v. In the plots, u is defined to be positive to the east and v positive to the north. Since the u-v coverage associated with a given distribution should not extend beyond the maximum allowed by baseline length considerations, each contour plot is marked with an approximate boundary outside of which the results are deleted. One important point concerning this analysis is that the BWS delay has been computed on the basis of an "artificial" frequency of 8.3 GHz (see f in Eq. 134). In effect, this assignment of f pretends that, for BWS delay computation, the distribution was measured for the specified u-v values at 8.3 GHz (λ = 3.6 cm), even though it was not. The reason for this assignment is that most of our work will be carried out at X-band and it is therefore important to obtain BWS delay results at that frequency. To obtain the BWS delay at the actual frequency, one can easily scale the results by f^{-1} (or by λ) as in Eq. (134). All of the other plotted quantities require no explicit frequency assignment in computation, although they actually pertain, of course, to the observing wavelength given in Table 1 for distribution 9.

The results for the 10 distributions were fairly complex but some general descriptive statements can be made. The magnitudes of the structure delays (the BWS delays computed for X-band) relative to the centroid were as large as a nanosecond but typically were less than 150 psec. The largest delays (~ 1 nsec) occurred in very localized regions in the u-v plane where very small fringe visibilities (~ 0.03) occurred. On average, structure delay increased as the fringe visibility decreased.

With regard to effective position for BWS delay and phase-delay rate, the position of some sources changed as a function of u,v by as much as 10-30 marcsec while other sources remain fixed to within a few milliarcseconds. Thus, in applications that require position determination at 10 marcsec level, one must be careful in selecting sources and/or in correcting for source structure. For the specific case of ΔVLBI applications, a final assessment cannot be made until intercontinental structure measurements (preferably at S-band and X-band) are available. As indicated in Table 1, the distributions analyzed here were produced by transcontinental baselines and therefore have inadequate resolution for intercontinental applications.

The process of calibrating source positions for short-baseline measurements through the use of positions obtained on long baselines can also be investigated on the basis of the effective position plots. A more direct approach, however, is to study the structure delay plots along the anticipated u-v tracks for the proposed baselines. In general, the bounds on structure delay discussed above also apply to this calibration process.

XV. AN UPPER LIMIT FOR THE SOURCE STRUCTURE EFFECT IN BWS DELAY

The preceding section has shown that source structure can sometimes contribute errors in BWS delays in the 10-30 cm range. Thus, if centimeter-level

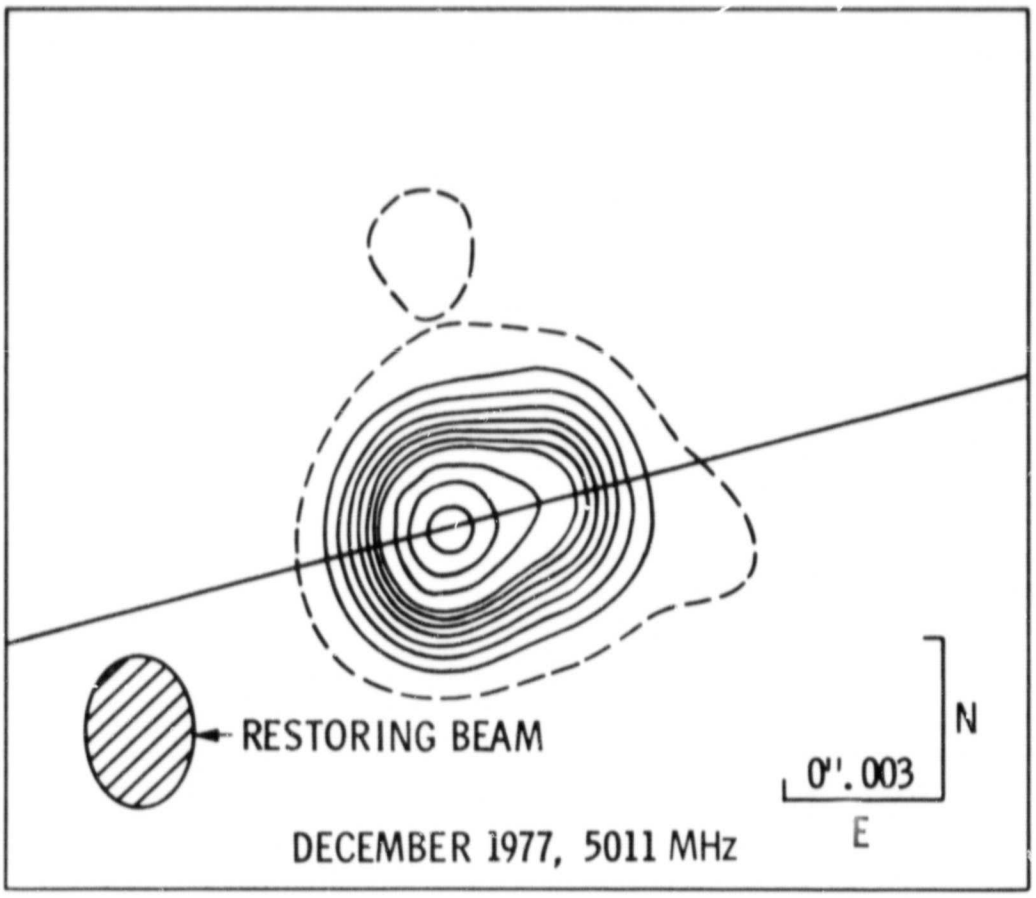


Figure 23. Measured Brightness Distribution for 3C345, Distribution 9

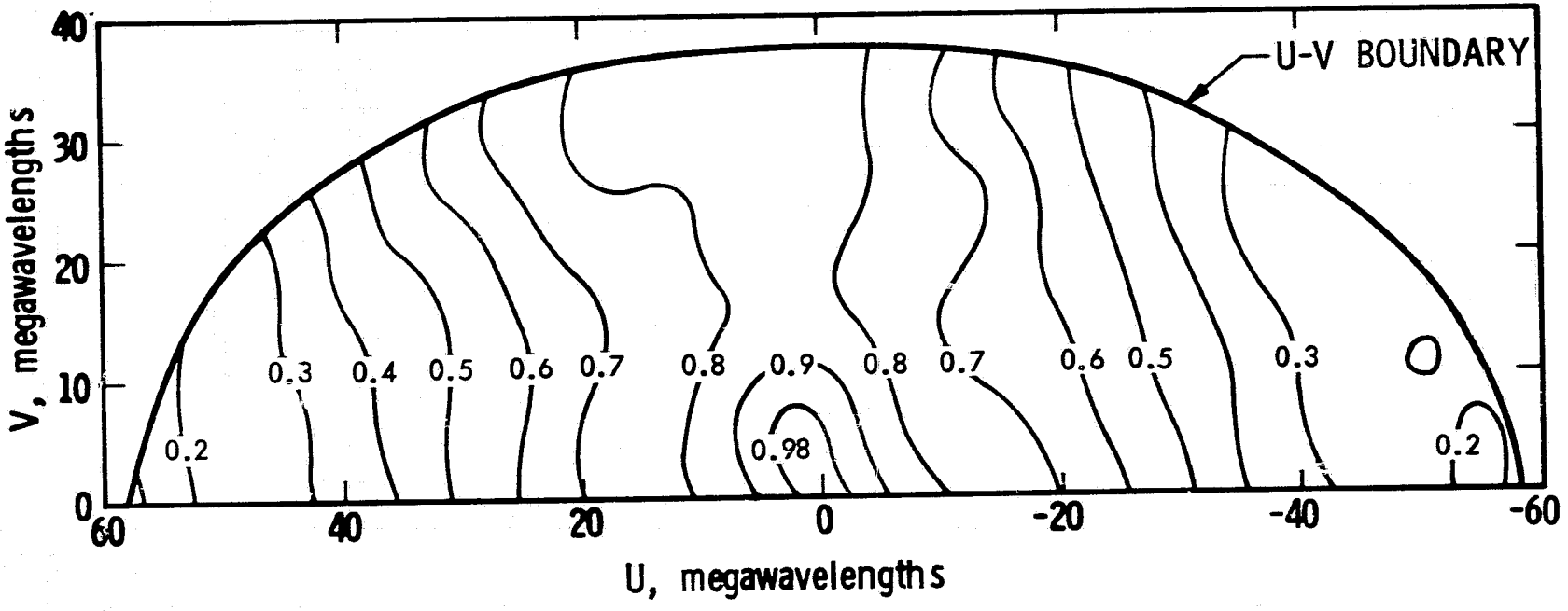


Figure 24a. Fringe Visibility for 3C345, Distribution 9

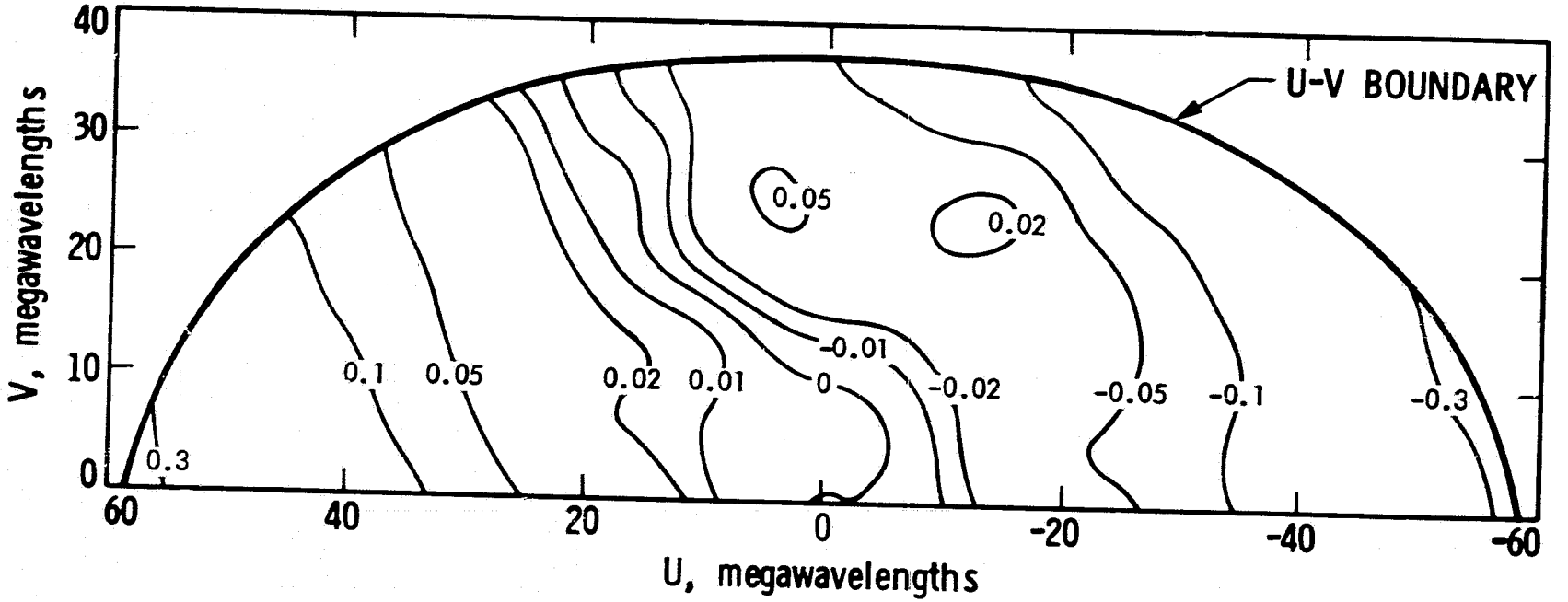


Figure 24b. Structure Phase (cycles) Relative to Brightness Centroid for 3C345, Distribution 9

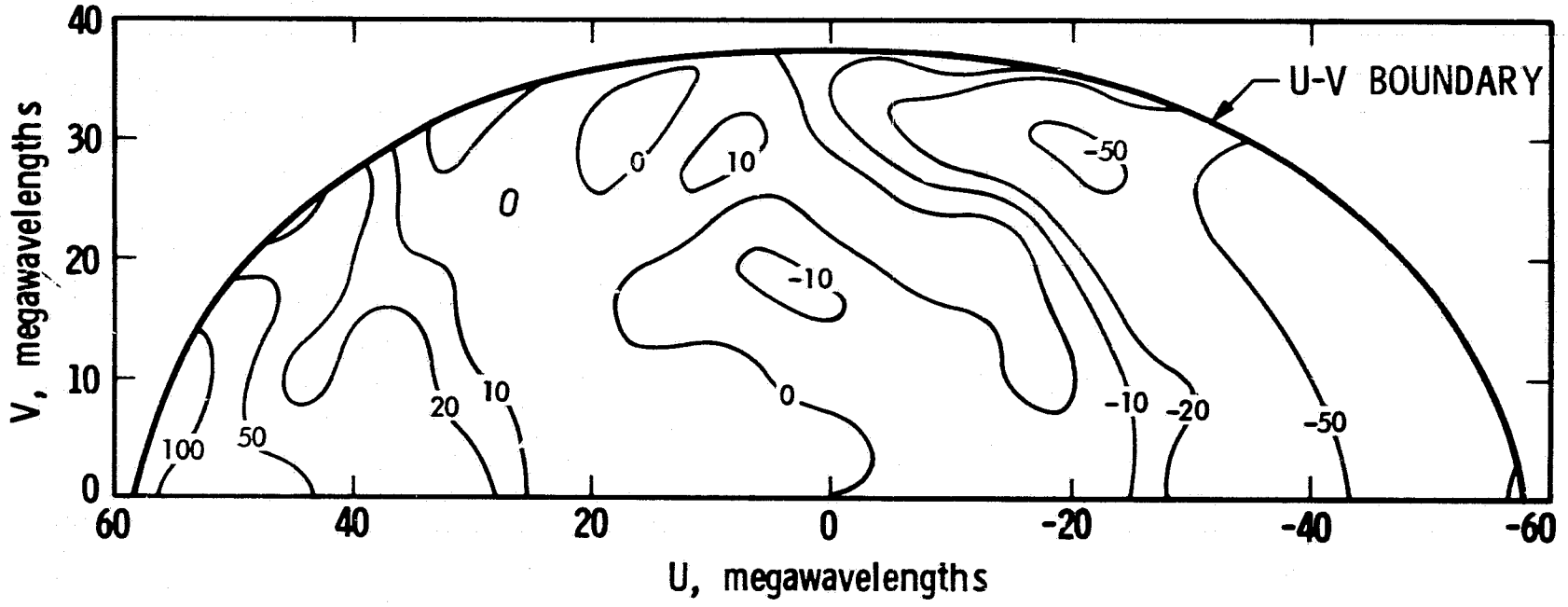


Figure 24c. Structure Delay (psec) Relative to Brightness Centroid for 3C345, Distribution 9

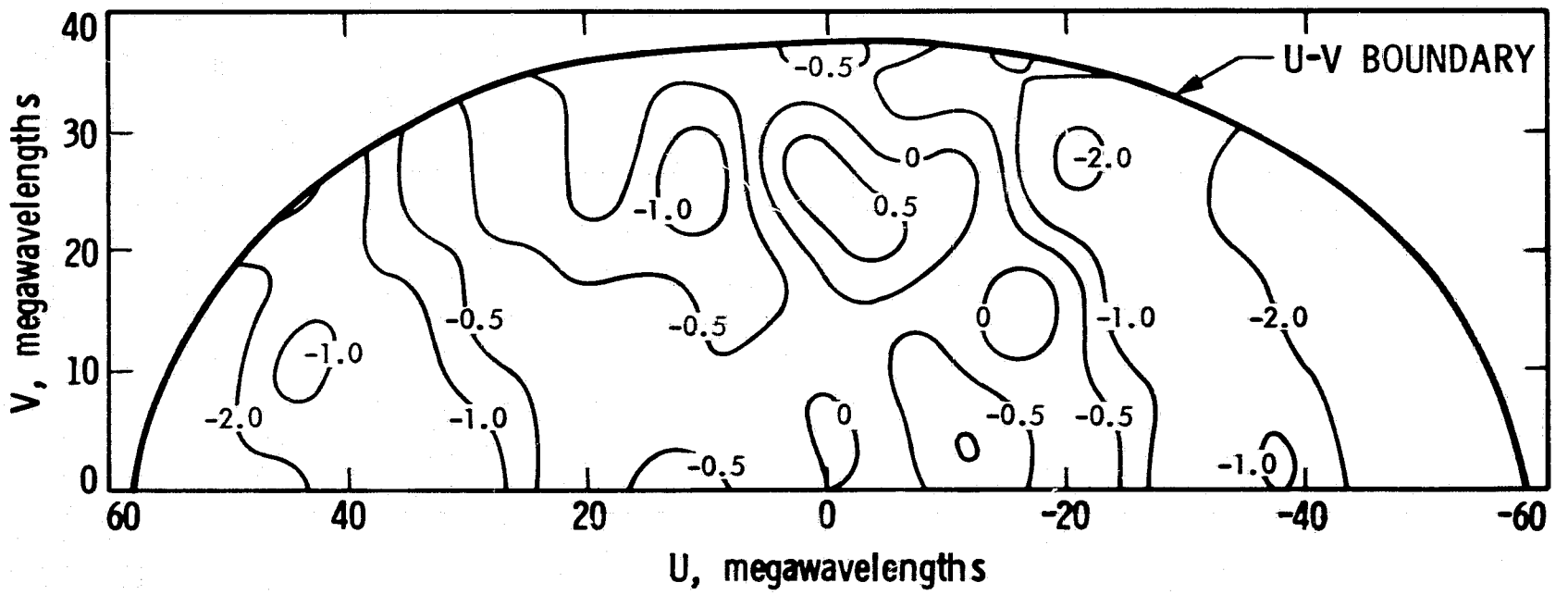


Figure 24d. Effective Position (milliarcsec), K Component, Relative to Brightness Centroid for 3C345, Distribution 9

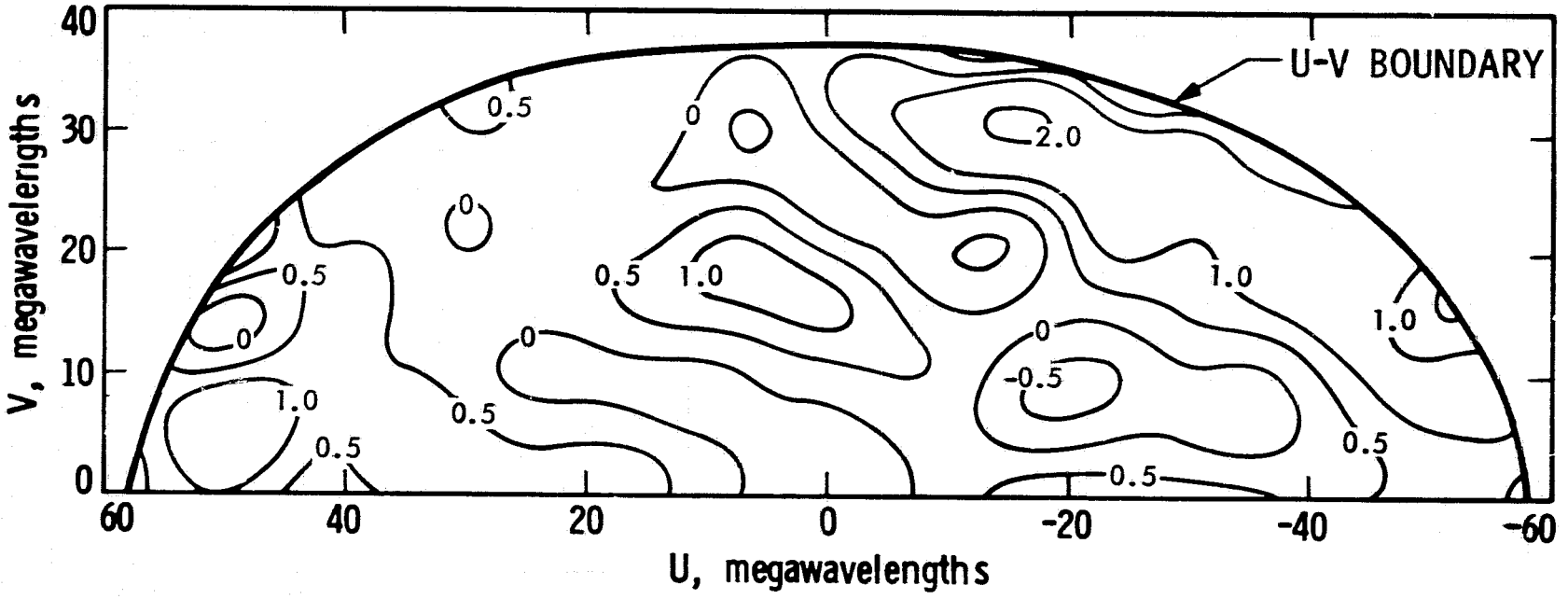


Figure 24e. Effective Position (milliarcsec), Y Component, Relative to Brightness Centroid for 3C345, Distribution 9

baseline measurements are a goal, some method must be devised for reducing or calibrating structure effects. A possible method is calibration through the use of measured brightness distributions, as discussed in the preceding sections. The primary difficulty with this approach is that an individual VLBI structure measurement is currently an expensive and time-consuming process. The prospect of working with very large catalogues, possible containing many time-varying members, makes the calibration approach even less inviting. It is therefore worthwhile to investigate alternate methods for overcoming structure problems.

Another approach is suggested by the general tendency of BWS structure delay to increase with decreasing fringe visibility. Suppose a general formula could be established that, purely on the basis of the value of fringe visibility, sets an upper limit on structure effects in BWS delay. Then, if the limit turned out to be sufficiently small for some upper range of visibility values, the larger, unacceptable structure delays could be eliminated by merely deleting observations with the smaller visibility values. Such an approach is attractive since fringe visibility can often be obtained along with each VLBI observation. Thus, if the visibility determinations were accurate enough, the experiment would not depend on outside measurements for structure corrections. One important point that can be made is that the proposed upper limit would not have to be an absolute limit, valid for every source. For example, it would be useful to establish, if possible, an approximate 3σ statistical limit so that structure effects could be treated like other errors. Another example would be a limit that was valid for all sources except for infrequent pathological cases. In fits with redundant observations, the pathological cases could be discovered and deleted through residual analysis.

To begin development of a limit approach, this section derives and tests a limit formula for BWS delay due to structure. First, a fairly general analysis is used to derive an expression for an upper limit on structure delay. Next, special analytical sources are used to generate another expression for the upper limit. These theoretical results are compared with results from 10 measured distributions.

The general derivation of a delay limit proceeds as follows. Since the results for a double-point source suggest that possible frequency dependence in a given brightness distribution will produce a fairly small delay effect relative to that produced by the frequency dependence of the sky fringes, we will ignore the former dependence in the following derivation. As shown in Fig. 25, plot a given value of the brightness transform R in the complex plane along with a vector indicating the change in R produced by a change in frequency $\Delta\omega$. The quantities Z_c and Z_s are the real and imaginary components of R defined in Section III. The associated BWS structure delay will be given by

$$\Delta\tau_B = \frac{\Delta\phi_B}{\Delta\omega} \tag{137}$$

where $\Delta\phi_B$ is the change in phase of the brightness transform. Among the possible directions for ΔR, the maximum value for $\Delta\tau_B$ occurs when ΔR is perpendicular to R.

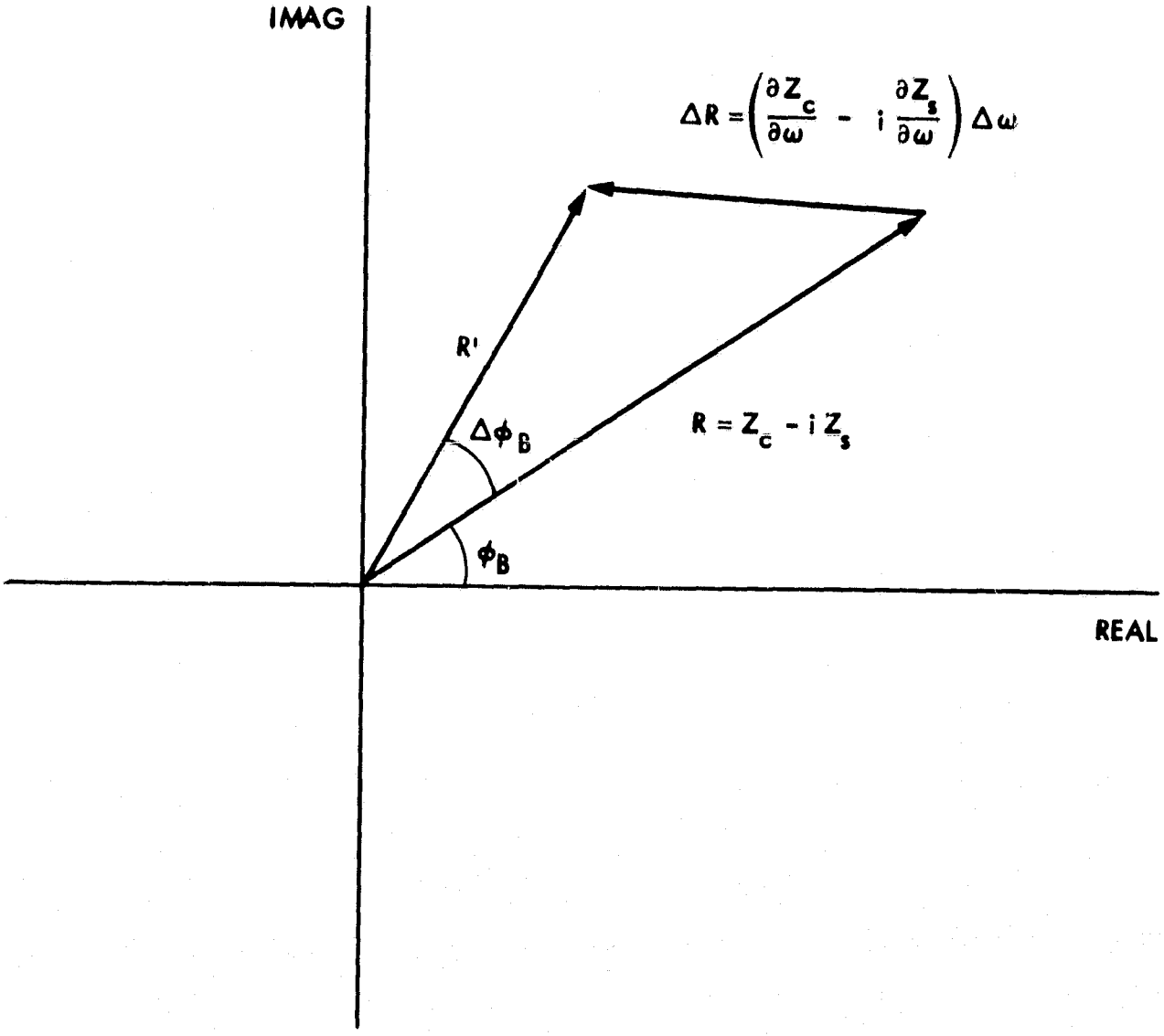


Figure 25. Schematic Illustration of the Change in Structure Phase with Frequency

Thus, an upper limit for $\Delta\tau_B$ is given by

$$\Delta\tau_B \leq \left[\frac{(\partial Z_c/\partial\omega)^2 + (\partial Z_s/\partial\omega)^2}{Z_c^2 + Z_s^2} \right]^{1/2} \tag{138}$$

where the inequality holds if ΔR is perpendicular to R. This expression can be further reduced for a frequency-independent N-point source. For such a source, the frequency derivatives can be obtained from Eqs. (108) and (109), which give

$$\frac{\partial Z_c}{\partial\omega} = -\frac{1}{\omega} \sum g_k\theta_k \sin\theta_k \tag{139}$$

$$\frac{\partial Z_s}{\partial\omega} = \frac{1}{\omega} \sum g_k\theta_k \cos\theta_k \tag{140}$$

where we have used the fact that θ_k is proportional to ω. If we momentarily neglect ω, the numerator in Eq. (138) becomes

$$\text{Numerator} = \left[\left(\sum_k g_k\theta_k \cos\theta_k \right)^2 + \left(\sum g_k\theta_k \sin\theta_k \right)^2 \right]^{1/2} \tag{141}$$

This expression can be viewed as the length of a vector formed by summing N vectors of length $g_k|\theta_k|$. The length of the sum vector can be no greater than the sum of the lengths of the component vectors. Thus Eq. (138) becomes

$$\Delta\tau_B \leq \frac{\sum g_k|\theta_k|}{\omega \left(Z_c^2 + Z_s^2 \right)^{1/2}} \tag{142}$$

For most real sources at larger u-v values, this limit will be much larger than the actual structure delay since the point-source vectors will add destructively to a large extent.

Up to this point, the derivation has involved no approximations and placed no restrictions on the structure of the sources. Unfortunately, it now becomes necessary to make assumptions since it is possible to concoct a source that will generate any arbitrary set of θ_k values. Assume a valid limit can be obtained by pulling outside of the sum a factor $2\pi k_s$ that represents an upper limit for the overall effective value of the θ_k terms. This operation gives

$$\Delta\tau_B \leq \frac{k_s}{f} \frac{\sum g_k}{(z_c^2 + z_s^2)^{1/2}} \tag{143}$$

which can be rewritten as

$$c\Delta\tau_B \leq k_s \frac{\lambda}{v_s} \tag{144}$$

where we have expressed the results in dimensions of length and have used the definition of fringe visibility:

$$v_s = \frac{(z_c^2 + z_s^2)^{1/2}}{\sum g_k} \tag{145}$$

This result indicated that the upper limit for structure delay varies in inverse proportion to fringe visibility and in direct proportion to observing wavelength. However, the limit is not of quantitative value unless an estimate for k_s is obtained.

The results for measured brightness distributions in Section XIV can be used to make an estimate of k_s. The simplest procedure is to determine for each distribution the value of k_s that makes the limit equation an absolute limit. This can be carried out by computing for each u-v point the value of k_s required to make the limit equation give the BWS delay obtained at that point from the brightness distribution. The resulting k_s values can be searched over the allowed region of the u-v plane to find a maximum value for each distribution. The results are given in Table 1, which shows that the maximum values for k_s fell in the range of 0.11 to 0.31.

Four interesting features became evident in the analysis of this particular limit equation. First, the u-v points at which the limit value was reached were usually not at fringe visibility minima although they were associated with the

lower range of visibility values. Second, these points occurred near the u-v boundary for eight out of the 10 brightness distributions while even larger (but necessarily disallowed) values occurred outside that boundary. This result suggests that some of the largest allowed values (see Table 1) may be merely an artifact of "boundary effects." More information is required to test this possibility. Third, for the important visibility values in the range 0.1 - 0.2, as well as for larger visibility values, the actual BWS delays were almost all less than about half the values predicted by the limit equation with $k_s = 0.25$. Thus, for these sources, the RMS delay error at a given visibility value was much less than the limit value, perhaps 1/3 as large. Fourth, for u-v points near zero, the delay effects were a very small fraction of the value predicted by Eq. 144. This result is not unexpected since the BWS delay is actually zero at $u = v = 0$, while the limit formula gives $k_s\lambda$. This last feature suggests that a better limit equation should be found.

Another limit formula can be obtained from an apparently unrelated result for a double-point source. At the first fringe visibility minimum, the BWS delay effect (see Eq. 94) can be written as

$$c\Delta\tau_B \approx \frac{\lambda}{4} \frac{1 - v_s^2}{v_s} \tag{146}$$

where the strict definition of v_s in this case is the value of the fringe visibility at the minimum. Although Eq. (146) was not derived as a limit, we will test it as a limit by letting v_s be any given value of visibility. In comparing the two limits in Eqs. (144) and (146), we find they are approximately the same at small visibilities ($v_s \lesssim 0.2$) if we set $k_s = 0.25$. Since this is the approximate value for k_s obtained above through comparisons with measured distributions, we see that Eq. (146) is approximately valid as a limit at small visibilities for those distributions. As visibility approaches 1.0 near $u = v = 0$, Eq. (146) correctly approaches zero, unlike Eq. (144). As discussed in Appendix B, one can be even more thorough and show that Eq. (146) sets a valid (though loose) upper limit for a range of small u-v values. The encouraging results for these two regimes support the use of Eq. (146) as a possible limit formula.

Although it is not conclusive, more support for Eq. (146) as a limit formula can be gained from theoretical considerations. First, the limit in Eq. (146) actually applies to a class of sources. For example, the result in Eq. (132) for a particular class of triple-point source can be reformulated at the first visibility minimum ($m = n = 0$) so that it matches Eq. (146). Further, although is will not be shown here, a particular class of N-point sources can also conform to this equation. Second, despite the fact it is based on special cases, Eq. (146) provides a more general limit than is immediately obvious. For both of the limit steps (i.e., Eqs. 138 and 142) established in the preceding derivation, these special cases require the equal sign. That is, ΔR is perpendicular to R for Eq. (138) and, for Eq. (142), all phasors resulting from the frequency derivative of the point-source phasors (i.e., $\omega^{-1}\theta_k g_k e^{i\theta_k}$) add coherently. Thus, the full vector strength of the derivative phasors is realized, as represented in the numerator of Eq. (142). Although there are these vector advantages, one can point out that the

derivative amplitudes are limited since the θ_k values never exceed π for the special cases. Thus for larger u-v values where the $|\theta_k|$ values are quite large, the limit might fail. Whether failure occurs very often with real sources must be tested with many measured brightness distributions. One factor working against failure for many sources is the likelihood that increases in the θ_k values will be counteracted by rather severe deviations from the coherent addition of phasor derivatives.

At least two categories of distributions can be specified that would violate the limit formula. The first category consists of sources with a large, strong diffuse component displaced by a considerable distance from a weaker compact component. In this case, the effective position would move from the centroid of both components to the center of the compact component as baseline length increased. If the separation of components were great enough, large changes in structure delay could occur. The second category includes any source with two components that are separated by a distance much larger than the extent of either component. As baseline length increased, such a source could be modeled approximately as a double-point source until one of the components began to resolve out. Since many "sky fringes" could be placed between the components, the relative maxima for BWS delay in Eq. (94) might be present for several values of n. Since the limit formula is based on the first and smallest relative maximum (n = 0), the larger maxima would exceed the limit formula. As mentioned earlier, the existence of such nonconforming sources does not necessarily invalidate the limit approach. As long as the limit formula only fails for a small percentage of sources, it still would be of great value in reducing and estimating structure effects in BWS delay for the other sources. As an example of the limit approach, suppose one set a lower limit of 0.2 on allowed fringe visibilities. Then the maximum structure delay predicted by the limit equation would be about 4.5 cm at X-band, as shown in Figure 26. Many observations would have visibilities larger than 0.2 and would therefore have smaller structure delay limits. Further, as discussed above, actual structure delays for the analyzed sources were almost always much less than the limit, with the RMS value being about 1/3 the limit at a given value of visibility. If we hypothesize on the basis of this preliminary statistical information that the maximum delay of 4.5 cm is an approximate estimate of the 3-σ delay for all observations, then the 1-σ structure delay for all observations would be about 1.5 cm. This calculation suggests one might be able to reduce the 1-σ structure delay to about 1.5 cm simply by deleting observations with fringe visibilities less than 0.2. For shorter regional baselines, one would be able to place a higher cutoff on allowed visibility values and therefore could reduce structure delays even more. For example, with visibilities greater than 0.5, the trial limit formula and the same statistical assumptions indicate a 1-σ structure delay of about 0.5 cm at X-band.

As mentioned in Section XIV, in some applications the source positions obtained from long baseline measurements are used as a priori for short baseline measurements. One important feature of such position calibrations is that any structure errors in delay encountered on long baselines will be reduced in effect by about the ratio of baseline lengths. This feature allows relaxation of the structure delay requirement that one must place on the long-baseline behavior of sources to be calibrated for use with regional baselines. For example, if 1-σ

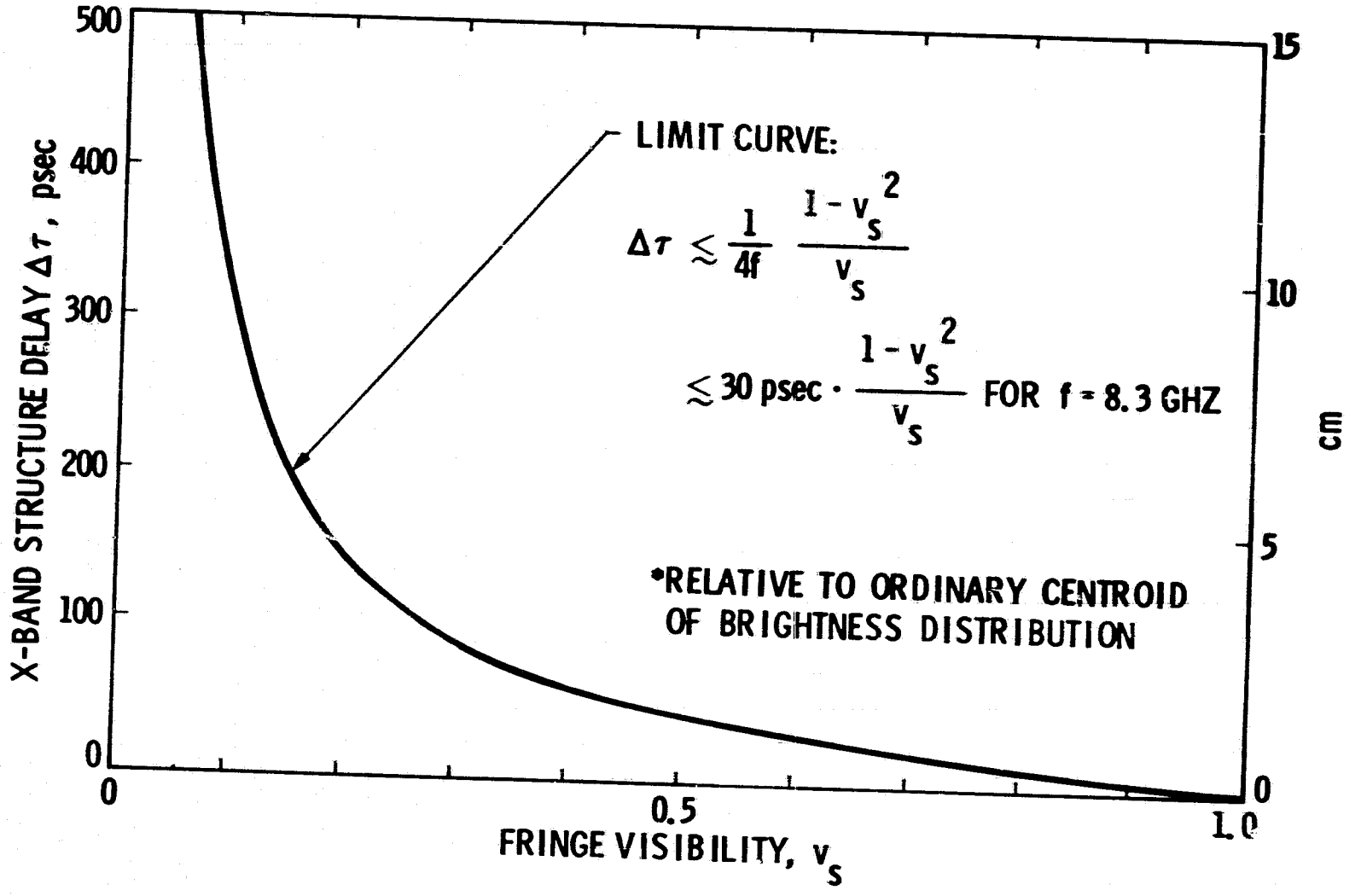


Figure 26. A Limit Formula for Structure Effects* in BWS Delay

structure delays equal to about 5 cm corrupt the source positions determined on a 6000-km baseline, those position errors would show up as about a 0.5-cm error (1-σ) in the calibrated delays on a 600-km baseline. Since the preceding calculation indicates that a 1-σ structure delay less than 5 cm at X-band on a long baseline should not be difficult to realize, position calibration errors due to long-baseline structure can probably be restricted to 0.5 cm or less for regional baselines. The other structure error in such calibrations, that due to variations in effective position on the regional baselines themselves, might be reduced to about 0.5 cm by a visibility cutoff, as discussed above.

The preceding examples of the limit approach are based on the analysis of only a few sources and must therefore be treated with caution. To conclusively test the validity of Eq. (146) or any other limit formula, many more measured distributions must be analyzed using more sophisticated techniques. Improvements in the analysis would include computation of the following quantities for each distribution: (a) the RMS structure delay for u-v points with visibilities greater than a variable lower cutoff, (b) the number of u-v points with visibilities falling within separate visibility intervals (e.g., with 0.1 spacing) and (c) the maximum and RMS structure delay for each of these bins. The results could be used to improve the limit equation by empirically determining a new multiplicative function of visibility to better account for the general dependence of the limit on visibility. Further, such an analysis would allow a delay value from the limit equation to be more reliably connected with a typical or RMS value for delay at a given visibility.

XVI. SUMMARY AND CONCLUSIONS

This report investigates structure effects in VLBI measurements and includes (a) a theoretical background, (b) an approach to structure corrections based on measured brightness distributions, (c) two analytical examples of sources, (d) the results of an analysis of measured brightness distributions, and (e) a limit formula for reducing and estimating BWS structure delay. In the theoretical derivation, the concept of effective position is formally introduced. It is shown that, for a given observation of any source, the source can be analytically and conceptually replaced by a point source located at the formal effective position. When BWS delay and delay rate are the observables, one unique effective position, valid for both observables, can be specified for each observation. For a given source, structure corrections are carried out by applying to each observable the correction that results from moving the effective position for that observation to an assigned reference position (e.g., the origin of structure coordinates). One insight offered by this approach is that the unknown origin of structure coordinates for an independently measured brightness distribution can be absolutely placed on the celestial sphere by means of standard VLBI measurements and multi-parameter fits.

An analysis of the examples of a double-point source and a triple-point source shows that the largest and most troublesome corrections to BWS delay (>1 nsec at X-band) and delay rate will occur in regions of the u-v plane near points of zero or nearly zero fringe visibility (<0.03). Near these points, structure

phase undergoes rapid variations as a function of (u,v) and the BWS delay correction can even (theoretically) approach infinity. In the example of a double-point source, the first line of zero amplitude in the u-v plane can be crossed only when the sky-projected baseline length satisfies $B_s/\lambda > 0.5/S$, where S is the point-source separation. For a triple-point source, the first null point can be reached only when $B_s/\lambda \gtrsim 0.3/S_k$, where S_k is the separation of point source 2 or 3 from the strongest point source. The analyses for double-point and triple-point sources also show that their effective positions (for BWS delay) can lie far outside the extent of the source near u-v points with very small fringe visibilities. In general, similar behavior in effective position will probably be observed in the u-v regions near local deep minima in amplitude for most sources possessing such minima. With regard to possible frequency dependence of brightness distributions, it is shown that a source consisting of two points with dissimilar spectral indices can require non-negligible additional corrections to BWS delay at X-band (~150 psec = 4.5 cm) near points of very small fringe visibility (~0.03) if the difference in the two spectral indices is large ($\Delta k = 1.0$). However, this effect is about an order of magnitude smaller than the delay effect due to the frequency dependence of the sky fringes mentioned above. For larger, more common values of fringe visibility (>0.2), the delay effect will be less than 1 cm. This result suggests that it may be possible in general to neglect calibration of this particular structure effect whenever the goal for total delay error is greater than 1-2 cm. Since this hypothesis is based on results for a double-point source, the frequency dependence of many actual distributions should be analyzed to see if the hypothesis is generally valid. Such a study has not been carried out to date.

To begin a study of structure effects for real sources, 10 brightness distributions measured by the Caltech group have been analyzed. Again, the largest BWS delay effect (~1 nsec = 30 cm at X-band) occurred at points in the u-v plane where fringe visibility fell to very small values (~0.03). More typically, BWS structure delay fell in the range of 1-10 cm. These results indicate that some method must be developed for handling source structure when delay accuracies of a few centimeters are desired. Effective position (for BWS delay) in some cases varied by as much as 10-30 marcsec, which indicates that, in applications concerned with 0.01" position measurements, one must be careful in selecting sources and/or correcting for structure effects.

Since structure effects can be quite large, this report has begun an investigation of a limit approach to source structure effects in BWS delay. A trial limit formula (Eq. 146) has been constructed that might correctly set an upper limit on BWS structure delay. Theoretical considerations show that the formula sets a valid (though loose) upper limit for all sources as u and v approach zero. At present, the formula has been compared with too few real sources to establish its general validity. However, the seven real sources analyzed at this time suggest that the limit is approximately valid and that it is approximately equal to three times the RMS structure delay at a given visibility. If these results turn out to be valid for nearly all real sources over all u-v values, then the limit approach might be of great value in both reducing and estimating structure effects in BWS delay. For example, in the last section, rough calculations based on these results indicated that the 1-σ structure delay at X-band might be reduced

to about 1.5 cm by avoiding observations with visibilities less than 0.2 (e.g., on intercontinental baselines) and to about 0.5 cm by avoiding observations with visibilities less than 0.5 (e.g., on regional baselines). The structure delays calculated in these two examples are less than the total error budgets for BWS delay in some VLBI projects. Thus some projects might avoid, at least temporarily, the large potential burden imposed by the present difficulty in obtaining VLBI brightness distributions.

A final word of caution would be that the limit approach in its present form has not yet been verified. It may turn out that, at the large u-v values associated with intercontinental baselines, too few natural sources possess sufficiently large fringe visibilities in the required u-v regions. Further, the measured distributions considered here are too few in number and too limited in u-v coverage and observing frequency to provide a general verification. The u-v coverage is too limited because (a) the absence of short baselines (< 300 km) can allow important large-scale structure to be resolved out and missed and (b) the absence of intercontinental baselines can allow important small-scale structure to remain unresolved. Finally, in the present limit approach, there are two statistical assumptions that must still be generally verified: (a) that structure delays as an ensemble can be statistically modeled and (b) that the limit equation is equal to about three times the RMS structure delay at a given visibility. In the last section, two categories of sources were specified that would violate the trial limit formula. More survey data is required to determine whether the percentage of sources that fall in those categories, or other disallowed categories, is too large for the limit approach in its present form to be valid. Even if the present approach is invalid it should be possible to construct another visibility-dependent limit that is valid for nearly all sources. How useful the final form will be in overcoming source structure problems remains to be determined. As a possible refinement to the limit approach, future studies should determine whether the nature of visibility signatures along a few u-v tracks can help eliminate undesirable sources. For example, one might show that sources with fairly smooth, constant signatures generally have smaller structure delays than those with widely varying signatures. To answer these questions, we plan to investigate additional sources in the near future and apply more detailed and sophisticated analysis, with the hope that the present limit approach, or some refined version, can be generally verified.

ACKNOWLEDGEMENTS

I would like to thank S. G. Finley for her enthusiastic and skillful assistance in generating the software for the contour and projection plots. I am also grateful to G. H. Purcell, Jr. and J. L. Fanselow for reading an advance version of this report and making numerous constructive comments and to the Caltech radio astronomy group for supplying their measured brightness distributions.

REFERENCES

1. Thomas, J. B., "An Analysis of Long Baseline Radio Interferometry", DSN Progress Report 32-1526, Vol. III, pp. 37-50, February 1972.

2. Readhead, A. C. S., Pearson, T. J., Cohen, M. H., Ewing, M. S., and Moffet, T. A., "Hybrid Maps of the Milliarcsecond Structure of 3C120, 3C273 and 3C345", Astrophys. J., 231, 299, July 1979.

3. Pearson, T. J., Readhead, A. C. S., and Wilkinson, P. N., "Maps of the Quasars 3C119, 3C286, 3C345, 3C454.3 and CTA102 with a Resolution of 5 milliarcseconds at 1.67 GHz", Astrophys. J., 236, 714, March 1980.

4. Readhead, A. C. S., and Wilkinson, P. N., "The Mapping of Compact Radio Sources from VLBI Data", Astrophys. J., 223, 25, 1978.

APPENDIX A. THE MEASUREMENT OF BRIGHTNESS DISTRIBUTIONS

The determination of a brightness distribution for a given source by VLBI involves the measurement of fringe amplitude and fringe phase for that source for many baselines, with the various baselines preferably spanning a wide range of lengths. Measured amplitudes and phases (actually closure phase) are passed through an inversion procedure to determine the brightness distribution. For this inversion, the CIT group used the "hybrid-map" approach (Ref. 4).

In VLBI measurements, the spatial resolution can be no better than about λ/B, where λ is the observing wavelength and B is the maximum length among the baseline vectors. Consequently, without loss of information, a measured distribution can be described by specifying a flux strength at each point on a grid in the plane of the sky where the dimensions of each square are set equal to a fraction of the best resolution. The output of the hybrid-map procedure specifies only the nonzero values and consists of a sequence of three numbers - a flux strength followed by two coordinates specifying the relevant grid point. Since the grid representation is unrealistically discontinuous and possesses more resolution than the interferometer merits, the final hybrid map for display is computed by convolving the grid distribution with a Gaussian restoring beam whose dimensions are approximately equal to the measurement resolution. This operation smooths out the distribution so that the fine detail is more consistent with the resolution limits of the interferometer. Since it gives a more accurate visual impression, the smoothed distribution will be shown for each source in this report.

In the analysis to obtain structure phase, etc., one has the choice of working with the grid distribution or the smoothed distribution. An important fact concerning this choice is the theorem that the Fourier transform of a convolution is equal to the product of the Fourier transforms of the two convolved functions. Since the Fourier transform of the Gaussian restoring beam has zero phase, the structure phase values resulting from the two representations will be identical, as will the BWS delay and effective position. With regard to amplitude, the transform of the grid distribution will possess amplitudes that closely approximate the measured values along the u-v tracks of the interferometer. However, the transform of the smoothed distribution will possess amplitudes that are smaller than the measured values along the u-v tracks for the longest baselines of the interferometer. Since one would want the transform amplitudes to equal the measured values, the grid distribution will be analyzed in this report.

Since sampling in the (u,v) plane is always incomplete, a measured brightness distribution is only approximately equal to the true distribution. Sampling in the (u,v) plane is sometimes incomplete in a local region. One example occurs near $u = v = 0$, which can cause large-scale structure to be missed. As mentioned above, sampling is always incomplete above the maximum (u,v) values, which means that the most compact components will not be evident. This report will not consider such deficiencies but will merely use the measured distributions to mechanically compute structure effects for (u,v) points below the maximum values.

The problem of incomplete sampling in the u-v plane might be alleviated in some cases by the inclusion of four observables not currently used in structure measurements: BWS structure delay ($\partial\phi_B/\partial\omega$), phase rate ($\partial\phi_B/\partial t$) and the frequency and time derivative of fringe visibility ($\partial v_s/\partial\omega$, $\partial v_s/\partial t$). The phase derivatives would be separately summed around each baseline triangle to reveal structure information, while the visibility derivatives could be obtained for each baseline. Measurement of the two derivatives with respect to frequency would require multichannel observations, such as those described in the text for bandwidth synthesis.

To demonstrate how these new observables might complement the standard observables of visibility and phase, first consider the phase derivatives and rewrite BWS structure delay in the form

$$\Delta\tau_B \approx \frac{\partial\phi_B}{\partial\omega} = \frac{\partial\phi_B}{\partial u}\,\frac{\partial u}{\partial\omega} + \frac{\partial\phi_B}{\partial v}\,\frac{\partial v}{\partial\omega} \tag{A1}$$

$$= \frac{1}{\omega}\,\nabla\phi_B \cdot \frac{\vec{B}_s}{\lambda} \tag{A2}$$

and phase-delay rate in the form

$$\frac{\partial\phi_B}{\partial t} = \frac{\partial\phi_B}{\partial u}\,\frac{\partial u}{\partial t} + \frac{\partial\phi_B}{\partial v}\,\frac{\partial v}{\partial t} \tag{A3}$$

$$= \nabla\phi_B \cdot \frac{\dot{\vec{B}}_s}{\lambda} \tag{A4}$$

where ∇ denotes the gradient with respect to u-v coordinates. In the equations, we have only considered the time and frequency dependence due to the sky fringes and have neglected the time and frequency dependence of the brightness distribution. As discussed in the text, results for a double-point source suggest that the variation of the sky fringes with frequency might produce an effect in the derivative that is about an order of magnitude greater than that due to frequency dependence of the brightness distribution (compare Eq. (91) with Eq. (94) and see p. 78). A study of a variety of real sources will be required to resolve this issue. On the other hand, the effect of the time dependence of the brightness distribution on phase rate is expected to be negligible. With regard to the utility of these observables, the phase rate observable might help to extrapolate into data gaps along a u-v track or might improve sampling in regions

of rapidly varying phase. The BWS delay observable might help to extrapolate into unsampled regions between tracks in the u-v plane since it is sensitive to the change in phase in the direction of such regions. The primary question is whether the accuracy of these closure observables, particularly BWS structure delay, will be sufficient to provide significant new information. Study of that issue is beyond the scope of this report.

With regard to the visibility derivatives, one can substitute v_s for ϕ_B in the above equations and show that the gradient of visibility in u-v coordinates could be determined by measuring $\partial\phi_B/\partial\omega$ and $\partial\phi_B/\partial t$. We again assume that the time and frequency dependence due to the sky fringes greatly exceeds the time and frequency dependence of the brightness distribution. This is expected to be a valid assumption for the derivative with respect to time. Results for a frequency-dependent double-point source suggest that the frequency dependence of a source will contribute a much smaller effect to the visibility derivative than the sky-fringe effect, provided differences in spectral indices of the components of the source are less than about 0.3. Again, more work is required to resolve this question for real sources. The extrapolation benefits described above for the phase derivatives also pertain to the visibility derivatives. To obtain accurate measurements of $\partial v_s/\partial\omega$, one would have to make accurate relative calibrations of the visibilities measured across the RF band. Again, it is beyond the scope of this report to estimate possible accuracies for the observables.

One strong indication of the positive contribution that could be made by accurate determination of the derivatives of visibility and phase is provided by ordinary sampling theory. If one attempts to obtain a complete measure of a function on the basis of a uniform sample rate, then the measurement of both the value and the derivative of the function at each sample point will allow a reduction of the minimum sample rate to 1/2 the sample rate required for value-only sampling. That is, the sample rate can be reduced from the Nyquist rate (R_N = 2 x highest frequency) to $R_N/2$. Thus, accurate measurements of the derivatives of visibility and phase would substantially increase the effectiveness of a given set of inadequate u-v tracks if the contribution due to the frequency dependence of the distribution were sufficiently small for the source being measured.

APPENDIX B. FRINGE VISIBILITY, BWS DELAY AND PHASE FOR SMALL U-V VALUES

For small u-v values, the θ_k values in Eq. (110) become small and the fringe visibility in Eq. (111) can be shown to approach

$$v_s^2 \approx 1 - < \theta_k^2 > \tag{B1}$$

where

$$< \theta_k^2 > \equiv \frac{\Sigma g_k \theta_k^2}{\Sigma g_k} \tag{B2}$$

We have assumed that the θ_k values (which are expressed in radians) are evaluated relative to the ordinary centroid so that

$$\Sigma g_k \theta_k = 0 \tag{B3}$$

Thus, for small u-v values, the fringe visibility decreases by an amount equal to the RMS spread of angles of the point-source phasors. Further, one can show that the BWS delay derived from Eqs. (112) and (113) can be reduced to the form

$$c\Delta\tau \approx \frac{\lambda}{4\pi} < \theta_k^3 > \tag{B4}$$

where $< \theta_k^3 >$ is given by an expression similar to Eq. (B2) and where the reference position is again the ordinary centroid.

The BWS delay result in Eq. (B4) can be used to test the limit formula in Eq. (146) for small u-v values. Although it was not derived for small θ_k, Eq. (146) is used in limit applications by merely substituting a given value of visibility. Thus, that limit can be connected with the θ_k values through the use of Eq. (B1). This operation gives

$$c\Delta\tau \approx \frac{\lambda}{4} < \theta_k^2 > \tag{B5}$$

Due to the factor of π and due to the fact that the θ_k values are all assumed to be a small fraction of a radian, one can readily see that the actual delay given by Eq. (B4) will be much smaller than the limit predicted by Eq. 146 (i.e., Eq. B5). Thus, at small u-v values, Eq. (146) is a valid upper limit for BWS delays.

Since this limit is not as tight as one might desire, the form of Eq. (B4) suggests a new limit formula that might be tested for small u-v values:

$$c\Delta\tau \lesssim \frac{\lambda}{4\pi}\left(1 - v_s^2\right)^b \quad , \quad v_s \gtrsim 0.9 \tag{B6}$$

where the constant b is to be empirically determined but should be greater than 1.0. This form will be tested along with other limit formulas in future work.

Finally, a similar derivation can be carried out for phase in Eq. (107) to give

$$\phi_B \approx \frac{1}{6} < \theta_k^3 > \text{ (in radians)} \tag{B7}$$

This result suggests a limit formula for phase delay that should be tested for small u-v values:

$$c\Delta\tau_\phi \lesssim \frac{\lambda}{12\pi}\left(1 - v_s^2\right)^b \quad , \quad v_s \gtrsim 0.9 \tag{B8}$$

where b was defined above. Since the delay in Eq. (B8) is 1/3 as large as the delay in Eq. (B6), this result for phase delay is another indication that phase delay can be less susceptible to structure than BWS delay.

NASA–JPL–Coml, L A, Calif