

Disconnected Interferometry at 81.5 MHz

Graham Woan
Clare College

A dissertation submitted for
the degree of Doctor of Philosophy
at the University of Cambridge.

February 1989

Preface

This dissertation describes research carried out at the Mullard Radio Astronomy Observatory between October 1985 and September 1988. It contains less than sixty thousand words and, except where clear reference is made, is my own work and includes nothing which is the result of work done in collaboration. It has not, nor has any part of it been submitted for any other degree, diploma or qualification at this or any other University.

The last three years have been a more or less completely pleasurable experience, mostly due to the friendship, support and advice of those around me, both in and out of the Physics department. No one is more responsible for this state of affairs than my supervisor, Peter Duffett-Smith, whose good humour and practical help have been quite indispensable. For that (and Trisha's Sunday lunches) I am very grateful. Further thanks are due to those who eased the execution of the project, in particular to Richard Davis for over-seeing the Jodrell Bank end of the 3C48 observations, Olivia Judson and John Aldington for their early morning excursions to remote caravan sites, Chris Vincent-Smith and the boys of Christ's Hospital School who faultlessly looked after the mobile station for three weeks, David Robinson for help with the Maxent processing in Chapter 6 and the gentlemen of Lords Bridge who dutifully repaired all my dynamic modifications to the motorcaravan. Finally I'd like to thank David Buscher for some fearless adventures in data and food processing, my parents for their encouragement and support over many years, and Katie for making it all a bit more worthwhile.

“Another damned, thick, square book. Always
scribble, scribble, scribble. Eh: Mr Gibbon?”

William Henry (Duke of Gloucester), quoted in a Note to Boswell’s *Life of Johnson*.

Contents

Preface	i
Contents	iii
Abstract	vii
1 Low frequency VLBI – an introduction	1
1.1 What is interesting at low frequencies?	1
1.2 Synthesis mapping	3
1.3 VLBI at 81.5 MHz	4
1.4 VLBI correlation system	7
2 System Enhancements	9
2.1 The theory of interferometry	9
2.2 Digital fringe rotation	12
2.2.1 Digital systems	12
2.2.2 The fringe rotator	13
2.2.3 Fringe function correction	17
2.3 A portable phased array	20
2.4 Closure phase	21
2.5 Sensitivity	21

3	Internal phase calibration	25
3.1	Adaptive calibration in VLBI	25
3.2	Use in low frequency VLBI	27
3.3	Applicability of the ‘thin-triangle’ method	28
3.4	Determination of closure phase in noisy conditions	29
3.4.1	Phase noise in a single baseline	30
3.4.2	Phase noise in a closure relation	33
3.4.3	The effect of phase slopes	38
3.4.4	Conclusions	41
4	The megaparsec component of 3C48	43
4.1	Introduction	43
4.2	Recent observations	46
4.3	Results	48
4.4	Confusion models	50
4.4.1	Simulating the recent data	51
4.4.2	Simulating the data of Hartas <i>et al.</i>	54
4.5	Discussion	57
5	External phase calibration	61
5.1	Phase calibration using VHF transmitters	63
5.1.1	The choice of reference transmitter	63
5.1.2	Implementation and results	65
5.1.3	How good is the calibration?	68
5.2	Accuracy and stability	71
5.2.1	Phase tracking	71
5.2.2	Consequences of 1-bit sampling	73
5.2.3	Effects from the atmosphere	81
5.2.4	Receiver systematics	83
5.3	Review	83

6	The structure of Cassiopeia A at low frequency – short baseline observations	85
6.1	Introduction	85
6.2	Current impressions of Cas A	86
6.3	LBI observations	88
6.3.1	Ionospheric effects	90
6.4	Data reduction	92
6.4.1	Combining the frequency channels	92
6.4.2	Yagi alignment	93
6.4.3	Amplitude calibration	95
6.5	Results	96
6.6	Conclusions	115
7	The compact structure of Cas A at metre wavelengths	121
7.1	Introduction	121
7.2	81.5 MHz LBI observations	123
7.2.1	Data reduction	123
7.2.2	Interpretation	124
7.2.3	Discussion	130
7.2.4	Conclusions	132
7.3	Pulsar Search	133
7.3.1	Effects of dispersion	134
7.3.2	Sensitivity	136
7.3.3	Method	137
7.3.4	Results	142
7.3.5	Conclusions	145
7.4	IPS evidence	146
7.4.1	Simulations	149
	APPENDIX:	152

A	The fringe rotator transfer function	153
B	Portable phased array	155
C	Automatic tape switching	157
D	Gaussian phase noise	159
E	Theoretical beam patterns	163
F	Simulation source list	167
G	Co-ordinate transformations	171
H	Pulsar data acquisition system	173
I	Pulsar search sensitivity	175
I.1	Flux calibration	175
I.2	Minimum detectable signal-to-noise ratio	177
	References	179

Abstract

Long Baseline Interferometry is a well established technique that enables high angular resolution measurements to be made with a radio interferometer containing independent local oscillators and signal recorders. This Thesis addresses the problems encountered when this technique is applied to low frequency (81.5 MHz) astronomy, and specifically methods of improving its phase stability. The technique is used to reassess existing evidence for the existence of large-scale structure associated with the quasar 3C48, and to give a much better understanding of the structure and behaviour of the supernova remnant Cassiopeia A at low frequencies.

Much of the instrumentation used for this work already existed, but in a form unsuitable for measuring phase. The first section of the Thesis shows how these data collection and digital correlator systems could be modified and extended to improve sensitivity and generate true interferometric phase. Two phase calibration schemes are also considered. Firstly, the ‘thin triangle’ method, which uses the closure phase generated by a three-element interferometer to determine the true astronomical phase of a source comprising a single, bright point source and an associated resolvable component. This includes an analysis of how best to determine closure phase when the signal-to-noise ratio on one or more of the baselines is low, and shows techniques based on the ‘triple-product’ method, proposed by Cornwell, to perform best. Secondly, a new method of instrumental stabilisation is introduced and demonstrated, in which the interferometer is calibrated by broadband signals from VHF public broadcast transmitters, which are processed along with the astronomical data. The technique is shown to be capable of removing practically all the instrumental phase drifts usually associated with a long baseline interferometer on baselines up to 100 km.

A number of new observations are also presented, mostly made with the stabilised mobile interferometer described above. Firstly of the bright quasar 3C48 to investigate earlier reports of an associated steep-spectrum extended component. These show the importance of considering the effects of confusion in the earlier observations and prompt a critical reassessment of the existing evidence, which is shown to be flawed. Detailed observations of the supernova remnant Cassiopeia A are also presented, with a resolution and sensitivity previously unattained at metre

wavelengths. Comparisons with 20 cm VLA data of the same epoch show the remnant to look remarkably similar at the two frequencies, though tentative spectral index measurements indicate that its younger features have slightly steeper spectra than the rest. Furthermore, sensitive observations made on a 128 km baseline show that Cassiopeia A does *not* contain the steep-spectrum compact component with a flux density of more than ~ 5 Jy reported by many earlier authors. An even deeper search at 408 MHz, this time for pulsed emission from the remnant, puts an upper limit of 80 mJy on the flux density due to any associated pulsar with a period $4 > P > 0.02$ sec.

Chapter 1

Low frequency VLBI – an introduction

In this thesis I present techniques and observations resulting from my work in low frequency radio astronomy at the Mullard Radio Astronomy Observatory Cambridge. With the exception of Chapter 7 all the work was carried out at frequencies between 81 and 82 MHz. Local agreements have made this a semi-protected band in the region of Cambridge, although the growing amount electromagnetic pollution in the area is gradually reducing its usefulness. In this chapter I give a brief review of the relevance of low frequency data in astronomy together with the work already carried out at MRAO on the problems of long baseline interferometry at low frequencies.

1.1 What is interesting at low frequencies?

Any purely observational science rests on the quality and completeness of its data. In astronomy we attempt to understand how features of the Universe work by fitting what we see to mathematical models obeying fixed laws. When radio astronomy began as a science, there was an abundance of observed but unexplained phenomena. Progress was determined by the rate at which theories could be devised capable of going at least some way to explaining them. Today the situation is rather different. Although much, if not most, of the phenomena that we observe do not have a fully satisfactory explanation, it is usually the case that we have many more models providing explanations than data points providing constraints. It is important therefore to explore as great a range in frequency, resolution and sensitivity as possible to provide these constraints. The low frequency (~ 100 MHz), sub-arcminute resolution measurements with which this thesis is concerned are of particular importance to

the determination of the evolution and spectral structures of both galactic and extragalactic radio sources. Most sources exhibit some form of power law spectrum (i.e. $S(\nu) \propto \nu^{-\alpha}$) over a range of frequencies, with the constant α being determined largely by the nature of the emission mechanism and the emitting material. However, it is frequently the departures from a simple power law that tell us the most about a source. It is widely accepted that the synchrotron mechanism is the dominant generator of non-thermal radiation in the Universe. Natural turnovers from a simple power law can be identified at both ‘low’ and ‘high’ frequencies with properties of the synchrotron source. A low frequency turnover, which may or may not be at a physically low radio frequency (depending on the properties of the source), could be generated in a number of ways:

- There may simply be no electrons around of sufficiently low energy to emit at this frequency.
- The frequency may be below the cyclotron frequency of the source.
- The effective optical depth of the source becomes sufficient to block the non-thermal radiation when its apparent brightness temperature approaches the kinetic temperature of the relativistic plasma (i.e. the source becomes ‘self-absorbed’). The self-absorbed spectrum will have the form $S \propto B^{-0.5} \Theta^2 \nu^{2.5}$, where Θ is the angular size of the source, and B the magnetic flux density.
- If the refractive index of the plasma has dropped sufficiently to lessen the effects of synchrotron beaming, the source spectrum will show a sharp cutoff below a frequency of about $20n_e/B$ MHz, where n_e is the electron density. (The Razin-Tsytovič effect.)
- Both external and internal free-free absorption (Bremsstrahlung) become more important at low frequencies. Measurements of the optical depth can then give a value for the emission measure along a path to the source.

At the ‘high’ frequency end of the nominally power law spectrum, another turnover can be expected. As the electrons lose energy they radiate at predominantly lower and lower frequencies, so this turnover is determined by the balance between the rate of generation of high energy particles and their rate of energy loss by processes such as synchrotron radiation, inverse Compton scattering or adiabatic expansion. For a ‘dead’ radio source, that is a source in which the accelerating mechanism has ceased, this turnover occurs at progressively lower frequencies giving a very steep spectrum. Low frequency observations therefore frequently select old, extended structures not

seen at higher frequencies. Each of the mechanisms described above may act either on the entire source or just in selected regions. The need for good resolution together with a range of observing frequencies becomes clear.

On a slightly more pragmatic level, there is a good deal of low frequency radio data on the Universe that might not tally with higher frequency observations not because things are different at low frequencies, but because the measurements are just wrong. Low frequency astronomy is greatly hampered by large primary beams, calibration difficulties, interference, poor resolution and badly behaved plasmas between the telescope and the radio source. Cross-referencing between low frequency techniques is vital if their results are to be taken seriously.

1.2 Synthesis mapping

For the most part, synthesis mapping at low frequencies has been confined to the generation of sky surveys. The 6C survey, to take a specific example, has identified tens of thousands of sources down to a flux density of a few hundred mJy at 151 MHz. The survey demanded only modest resolution (about 4x4 arcminutes half power beam width) at which few sources were resolved to any great degree. Even so, the maximum baseline was several km.

It is clear from this example that low frequency synthesis mapping of extended structure needs baselines of several tens of km. Long distances imply high cost, both due to the cost of equipment and the necessary logistical support. This problem has been tackled with particular success by MERLIN, using microwave links over baselines of up to 233 km. However the aperture plane coverage, especially at the short spacings, is far from complete.

With these points in mind, it is interesting to investigate how much can be done using portable VLBI equipment that could generate adequate aperture plane coverage at a reasonable cost. A portable antenna implies a cheap antenna, as it would need to be quite small. Simple receiving apparatus is all that is required, the system temperature being dominated by the galaxy. A low frequency suggests a narrow bandwidth, because the radio spectrum is quite cluttered around 100 MHz, and this in turn means that data rates and synchronisation requirements are modest. So what would the sensitivity of such a system be? To quite a good approximation, the equivalent noise flux density per sample per baseline of a 2-element interferometer is

$$S_n = \frac{kT_s}{\sqrt{A_1 A_2 \eta}} \sqrt{\frac{2}{\Delta\nu\tau}}, \quad (1.1)$$

where k is Boltzmann's constant, T_s the system temperature, A_1 and A_2 the effective areas of the antennas, η an efficiency factor for the processing (≤ 1), $\Delta\nu$ the RF bandwidth and τ the integration time. At a frequency of 100 MHz the system temperature is about 2000 K and the available bandwidth about 1MHz. If the map is to be small, say 16x16 pixels, then the maximum integration time at the edge of the aperture plane could be as much as 15 minutes. Taking η as 0.5 this gives a noise of

$$S_n \simeq \frac{250}{\sqrt{A_1 A_2}} \text{Jy} . \quad (1.2)$$

It is usual in VLBI to have to search for fringes in delay space, as synchronisation is never exact. This difficulty could be overcome if hydrogen masers, with a stability of 1 part in 10^{15} , were used as the clocks. However their price (around \$50k each) is usually prohibitive. The maximum useable baseline is therefore the one at which the signal-to-noise ratio is about 3, corresponding here to a minimum correlated flux of $3S_n$ Jy. To maximise the total effective area while still retaining the concept of a portable system, A_1 must be large (and at a fixed position) and A_2 must be as large as can be made portable, say an array of 16 Yagis, giving an effective area of about 100 m² at 100 MHz. Taking A_1 as 5000 m², (the effective area of a ~ 100 m dish), this gives a sensitivity of about 1/3 Jy per synthesised beam. However, if the map is to contain about 16² beams, the integrated flux density of the source over the map may have to be much greater than this, depending on the fraction of flux that is resolved. If the source contains no bright compact component, its integrated flux density might have to be as high as 80 Jy before the making of such a map could be justified. In reality the limiting flux will be somewhere between the two, so there are quite a number of sources, (the order of 100), which could be mapped.

1.3 VLBI at 81.5 MHz

To investigate just how far the the ideas in the previous section could be pushed, a research programme into low frequency very long baseline interferometry was begun in Cambridge. Much of the pioneering work was carried out by Duffett-Smith and Hartas [41], a detailed description of which can be found in Hartas's thesis [40]. This prototype VLBI system, operating at 81.5 MHz, was used to make transit observations on a number of 3C sources over baselines of between a few hundred and a few hundred thousand wavelengths (Figure 1.1). At low frequencies the bandwidth available is limited and the system temperature is dominated by the Galaxy. Sensitivity is therefore determined solely by the collecting area of the antennas and the integration time. In order to get good aperture plane coverage a portable inter-

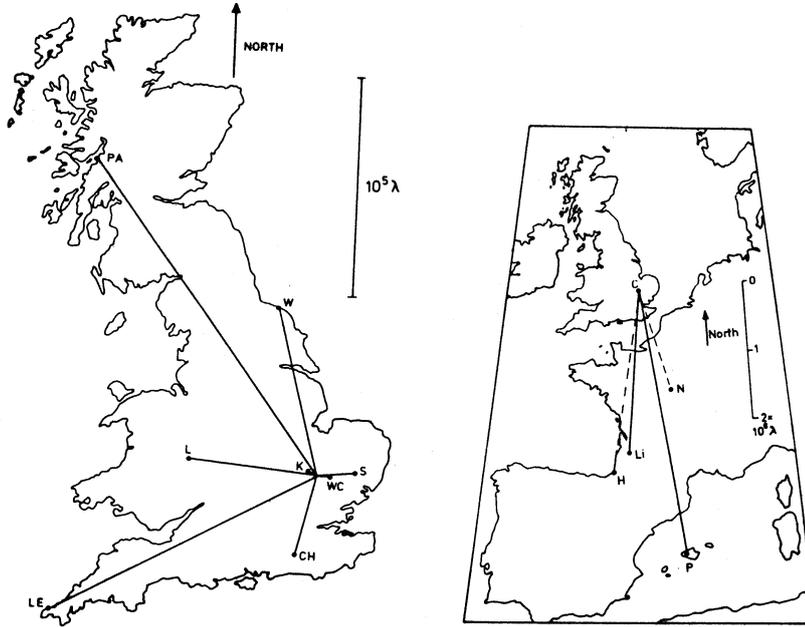


Figure 1.1: Baselines used by Hartas *et al.* [41]

ferometer was constructed, consisting of a converted motorcaravan containing the receivers, frequency standard and recording equipment, and a trailer that carried 16 Yagi antennas. Once a site had been found these antennas were erected to make a transit array, consisting of two East-West rows pointing at the zenith. The rows received radiation polarised in both the North-South and East-West directions so as to take account of the effects of differential Faraday rotation on the longer baselines. The base station, sited permanently at Lords Bridge near Cambridge, used identical electronics but was connected to the 3.6-hectare array, a transit instrument usually used for IPS measurements. This gave a total effective area of about 1500m^2 in each polarisation. Synchronisation between the two stations was achieved using rubidium standards and, when available, the 60 kHz transmissions from MSF Rugby. The data was 1-bit digitised (i.e. assigned a value of 1 if positive and 0 if negative) and recorded onto floppy discs, which were transferred to a microcomputer for correlation and integration at a later date.

This work showed that the system sensitivity did not change appreciably from site to site and that low frequency VLBI with a portable system was therefore possible. Fringes were obtained on nearly all baselines, and ionospheric decorrelation was never very significant. Anomalously high short baseline fluxes for a number of the twenty 3C sources studied led Hartas and Duffett-Smith to conclude that these

had steep spectrum extended structure, contributing up to 30% of their total flux at 81.5 MHz. This effect was especially prominent for 3C48 and 3C216, both of which appeared to be surrounded by haloes more than an arcminute in extent. Although confusion from neighbouring radio sources was considered to be a major problem by Hartas, it was never seriously investigated or accounted for. Chapter 4 of this Thesis goes into the problem in more detail and offers an alternative explanation for the measurements.

Rees and Duffett-Smith [74] used identical apparatus to that of Hartas *et al.* to investigate the effects of the various irregular plasmas along the propagation path of the radio signals. In particular they measured the dependences of the correlated signals from two quasars on their solar elongations, and hence derived a measure for the power spectrum of the solar wind. They showed that decorrelation was negligible for elongations of greater than about 40° . They also put limits on the amount of differential Faraday rotation expected on their baseline of 130 km, and found that it was usually less than 20° .

Further work by Rees and Duffett-Smith [73] resulted in the conversion of the large fixed parabolic trough element of the ‘4C’ antenna [77] into a phased array. The array, measuring 470m East-West by 20m North-South, could be steered electrically by means of 256 phase rotators (one per dipole), each capable of introducing phase in steps of 22.5° . The instrument could track the position of any source north of dec -5° for at least six hours, and was intended for use as a sensitive base station antenna for further long baseline experiments at 81.5 MHz.

The next step, completed by Spinks and Duffett-Smith [85], was to construct recording, playback and correlation equipment capable of exploiting the six hours of continuous observation time made available by the new antenna. A motorcaravan was again used as the mobile station, containing receivers which were connected to a single Yagi antenna erected a few yards away from the vehicle. The observing band (81 to 82 MHz) was divided up into eight 125 kHz sub-channels. Each channel was mixed to baseband by two local oscillators in phase quadrature, so generating a ‘complex’ representation of the signal – a procedure also carried out at the base station. Coarse timing information (to within a few milliseconds) was obtained from MSF Rugby, and this was augmented by signals from the LORAN C transmitter network, based in Sylt, Germany. This combination enabled the time to be determined to an accuracy of a few microseconds, which was well within the delay ‘window’ of the correlator. The local oscillators were phase locked to rubidium standards, one at each station, with a nominal frequency of 10 MHz and a coherence time of about 15 minutes. Coherence between the two stations was further improved

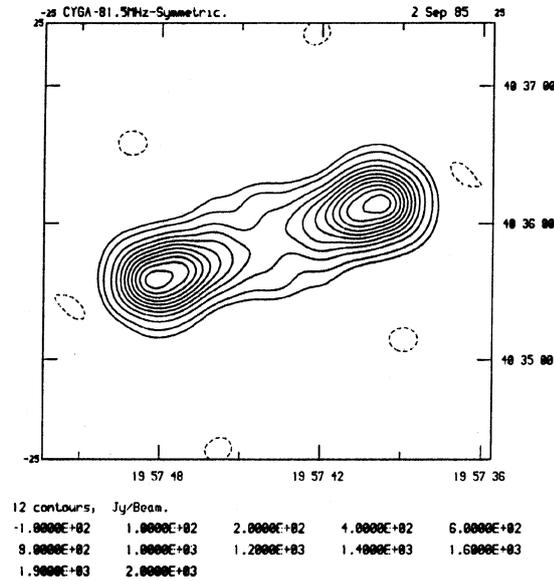


Figure 1.2: Map of Cygnus A, made at 81.5 MHz by Spinks *et al.* [86].

by phase locking the two frequency standards to the LORAN C transmissions. This technique succeeded in removing the long term frequency drifts so that synchronisation could be maintained indefinitely, but did not significantly improve the short term coherence of 10 to 20 minutes. The received signals were 1-bit digitised and multiplexed before recording onto magnetic tape. Modified video recorders were employed so that a six hour observation could be contained on a pair of three hour cassettes. Each run therefore generated four cassettes (two at each station) which were brought back to the laboratory for correlation.

This equipment was used by Spinks, Rees and Duffett-Smith to observe Cygnus A, a classical double source, on 20 baselines out to 45 km [86]. Although they did not have any phase information, the high degree of symmetry in the source enabled them to make a simple map (Figure 1.2) and infer a spectral age of about 3 million years for the centre of the bridge between the hotspots.

1.4 VLBI correlation system

To carry out these observations, it was necessary to construct a correlator capable of processing the tapes generated at the two stations. As this correlator, in a modified form, has been of central importance in the continuation of the project, it will be described here in some detail. A fuller description can be found in Spinks's thesis [85].

Once collected from the stations, the cassettes were inserted into modified video playback machines, from which the data emerged at a rate of 2.5×10^6 bits per second. Data were held in ‘frames’, similar to conventional video frames, each 20 ms long. The frames were numbered, and to aid synchronisation every 20th frame began at the start of a UT second. The audio track on the tape contained information to identify precisely which UT second this was. Data within the frame was held in a self-clocking code, from which both the clock and the data could be recovered with phase locked loops. The data from each tape then needed to be deformatted to split up the signals from the different frequency channels before being transferred to a buffer for storage. The data were then ready for correlation, which was achieved using EXCLUSIVE-OR gates. Shift registers provided 15 simultaneous delays in one of the data channels so that a search could be made for the correlated signal in delay space. Double sideband mixing meant that the data in each of the ‘sine’ and ‘cos’ components of a frequency channel had a bandwidth of 62.5 kHz and these were sampled at the Nyquist frequency (125 kHz). One delay therefore corresponded to a time interval of $8 \mu\text{s}$, so the whole window was $120 \mu\text{s}$ wide.

The correlator therefore processed eight sets of complex products (32 real products in all) over a range of 15 delays, and produced a new set of correlated products every $8 \mu\text{s}$. In order to reduce the data flow rate to manageable proportions, the correlator integrated these results for one second before transferring them to a pdp11 for storage. Clearly any signal with a natural fringe rate approaching, or greater than, 1 Hz would be washed out by this process, and so some degree of fringe rotation had to be carried out in the correlator itself. This was done by multiplying one of the data streams by a 3-level approximation to a cosine prior to correlation (Chapter 2 of this Thesis describes an improved fringe rotation scheme).

The correlated, partially fringe-rotated data were stored by the pdp11 on conventional magnetic tape at a rate of one block per second. The bandwidth constraints of the playback machines meant that correlation had to be done in ‘real time’, i.e. each cassette took three hours to correlate. Once correlation had been completed, the computer tape was rewound and reprocessed to remove the residual fringe rate and to integrate coherently for 60 seconds. A program then searched the eight delays to find the fringes and extracted an interpolated signal which was taken as the ‘output’ of the interferometer.

Chapter 2

System Enhancements

The theory and practice of constructing interferometers and correlators for radio astronomy has been well explored by many authors. A particularly comprehensive account of the subject can be found in ‘*Interferometry and Synthesis in Radio Astronomy*’ by Thompson Moran and Swenson [91]. Here I will present only the essentials necessary to highlight points of particular interest to long baseline interferometry and to clarify future notation.

2.1 The theory of interferometry

Consider a simple two-element interferometer consisting of two antennas separated by a baseline vector \mathbf{D} (Figure 2.1). Both antennas pick up a radio signal from a distant point source in a direction defined by the unit vector \mathbf{s}_0 . The signal received at antenna 1 is just that received at antenna 2, but delayed by $\tau_0 = \mathbf{D} \cdot \mathbf{s}_0 / c$, where c is the speed of propagation of the wavefront. For most practical purposes, the two signals must be mixed to baseband before they can be processed. This can be achieved in a number of ways, but in this example they will each be multiplied by two local oscillators in phase quadrature, to generate what can be regarded as two complex baseband voltages. If the RF signal has just a single spectral component at a frequency $\omega_0 + \omega$ and with amplitude $A(\omega)$, the RF voltages at antennas 1 and 2, also considered complex, are

$$V_1(t) = A(\omega) \exp[i(\omega_0 + \omega)t], \quad (2.1)$$

$$V_2(t) = A(\omega) \exp[i(\omega_0 + \omega)(t - \tau_0)]. \quad (2.2)$$

Usually the local oscillators used to mix these signals to baseband have the same frequency, being fed by the same master oscillator. In VLBI however, this is not

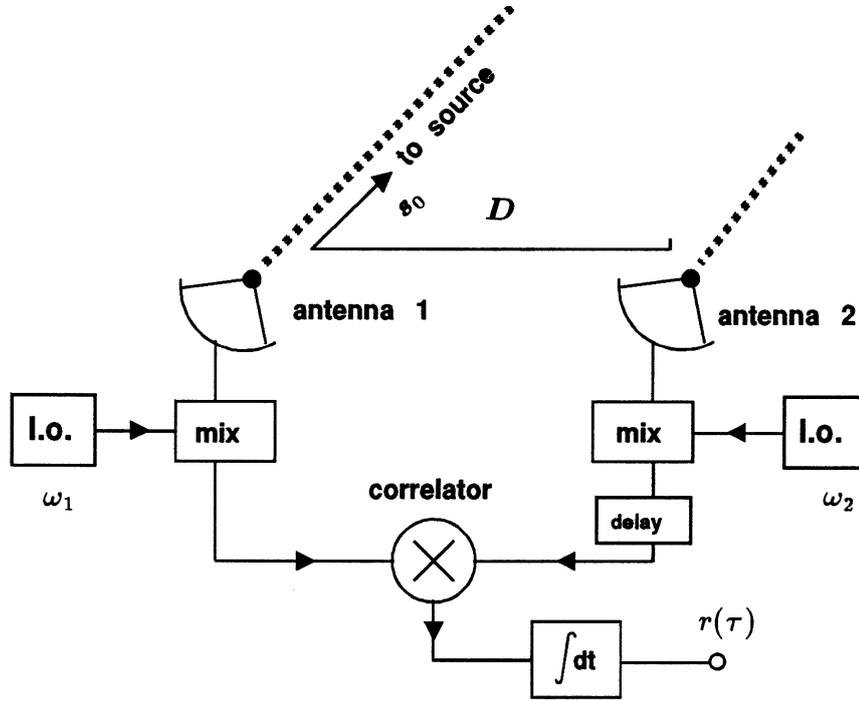


Figure 2.1: A simple two-element interferometer using independent local oscillators.

the case so that although they have the same nominal frequency ω_0 , their actual frequencies ω_1 and ω_2 will be slightly different. Consequently the two baseband signals will be

$$V_{1b}(t) = A(\omega) \exp[i(\omega_0 - \omega_1)t + i\omega t], \quad (2.3)$$

$$V_{2b}(t) = A(\omega) \exp[i(\omega_0 - \omega_2)t - i\omega_0\tau_0 + i\omega(t - \tau_0)]. \quad (2.4)$$

It is these signals which, in VLBI, are recorded onto magnetic tape and transported to the correlator. The correlator produces the complex cross-correlation between the two signals over a number of delays, τ , to give

$$g(\tau) = \langle V_{1b}(t)V_{2b}^*(t + \tau) \rangle_t, \quad (2.5)$$

where $\langle \dots \rangle_t$ denotes a time average. If $(\omega_2 - \omega_1)/\omega_0$ is small, this reduces to

$$g(\tau) = A(\omega)A^*(\omega) \langle \exp[i(\omega_2 - \omega_1)t] \rangle \exp[i\omega_0\tau_0 + i\omega(\tau_0 - \tau)]. \quad (2.6)$$

The coherence time of the signal is clearly defined by $(\omega_2 - \omega_1)$, although Chapter 5 describes a method for extending it indefinitely. For the remainder of this analysis it will be assumed that the coherence time is much longer than the integration time, so that the angled brackets can be dropped.

This result can now be generalised to an extended source emitting broadband noise signals. Although cross correlation is a fundamentally non-linear process, the condition that the source is spacially and temporally incoherent is sufficient for the non-linear terms to be removed by the averaging process. It is therefore quite valid to identify the correlation coefficient of radiation from such a source, $r(\tau)$, with a linear sum of coefficients due to its spectral and spacial components, i.e.

$$r(\tau) = \int_{4\pi} \int_{\omega} g(\tau, \mathbf{s}, \omega) d\omega d\Omega, \quad (2.7)$$

where $d\Omega$ is a small solid angle in the direction of \mathbf{s} . As the field of interest on the sky is usually quite small, it is useful to put $\mathbf{s} = \mathbf{s}_0 + \boldsymbol{\theta}$ where $\boldsymbol{\theta}$ is a small angle. $\mathbf{D} \cdot \mathbf{s}_0/c$ is therefore the gross geometrical delay between signals from the centre of this field being received at the two antennas. Equation 2.7 can then be written as

$$r(\tau) = \exp[i(\omega_2 - \omega_1)t] \times \int_{4\pi} \int_{\omega} B(\omega, \mathbf{s}_0 + \boldsymbol{\theta}) \exp[i\omega_0\tau_0 + 2\pi\mathbf{D}\boldsymbol{\lambda} \cdot \boldsymbol{\theta} + \omega(\tau_0 - \tau + \mathbf{D} \cdot \boldsymbol{\theta}/c)] d\omega d\Omega, \quad (2.8)$$

where $B(\omega, \mathbf{s}_0 + \boldsymbol{\theta})$ is the power received from the sky per unit frequency per unit solid angle in the direction $\mathbf{s}_0 + \boldsymbol{\theta}$, and $\mathbf{D}\boldsymbol{\lambda}$ is the baseline vector measured in wavelengths ($\lambda = 2\pi c/\omega_0$). This double integral can be split into a product of spacial and spectral integrals if $\omega\mathbf{D} \cdot \boldsymbol{\theta}/c$ is taken as small. This is the case if the resolution of the maximum baseline divided by the field of view is a lot less than the fractional bandwidth of the system, i.e. if there is no ‘bandwidth smearing’.

Assuming a square passband for the receivers of width $\Delta\omega$ centred on ω_0 , and a white noise spectrum for the source, the expression further reduces to

$$r(\tau) = \exp[i(\omega_2 - \omega_1)t] \times \exp[i\omega_0\tau_g] \times \text{sinc}[\Delta\omega(\tau - \tau_g)/2] \times \int B(\boldsymbol{\theta}) \exp[2\pi i\mathbf{D}\boldsymbol{\lambda} \cdot \boldsymbol{\theta}] d\Omega. \quad (2.9)$$

It is important to note here that by placing the mixer frequency at the centre of the RF passband, that is by double sideband mixing, the spectral integral in Equation 2.9 is purely real. This means that, at least in continuum work, the phase of the correlated signal does not depend on the correlator delay τ – a property of particular advantage in an LBI system.

Equation 2.9 contains four main components. The first is an unknown and changing phase factor determined by the difference in frequency between the local oscillators at the two ends of the baseline. This is only significant in LBI work,

and is similar to the phase noise from atmospheric effects which even a ‘connected’ interferometer is prone to. However, it is often of much greater magnitude than any atmospheric term, and therefore determines the coherence time of the instrument. The second phase term only depends on the nominal observing frequency, ω_0 , and the geometric delay, τ_g , and is therefore calculable. As the Earth rotates τ_g changes, so this phase also changes. Its instantaneous frequency, ω_g , the ‘fringe rate’ of the signal, must be accounted for in the data processing before any integration of the correlator output can take place. This operation is known as ‘fringe rotating’, or ‘fringe stopping’. The ‘sinc’ term in Equation 2.9 defines the range of correlator delays over which the signal can be detected and corresponds to a time of about the reciprocal of the receiver bandwidth. The constraint that $\tau - \tau_g$ must be within this interval is often referred to as the ‘white light’ fringe condition. In a digital correlator τ can only take on discrete values, determined by the sampling rate. For Nyquist sampling, the interval between the allowed values of τ is the same as the periodicity of the sinc function, so that when the sinc function is exactly centred on a delay, only that delay contains any signal. The final integral term in Equation 2.9 represents the fringe visibility Γ , which is of course the quantity the interferometer was designed to measure. It is usually convenient to decompose the baseline vector \mathbf{D}_λ into vectors parallel and perpendicular to the incoming wavefront. Over small angles the integral then becomes the simple two dimensional Fourier transform

$$\Gamma(\mathbf{u}) = \int B(\boldsymbol{\theta}) \exp[2\pi i \mathbf{u} \cdot \boldsymbol{\theta}] d\Omega, \quad (2.10)$$

where \mathbf{u} is the component of \mathbf{D}_λ parallel to the wavefront. Accordingly, the sky brightness distribution $B(\boldsymbol{\theta})$ can be determined from the inverse transform of $\Gamma(\mathbf{u})$.

2.2 Digital fringe rotation

2.2.1 Digital systems

Although in principle VLBI could be carried out with a totally analogue recording and processing system, digital techniques are usually preferable. At first this may seem surprising. Digital sampling requires that the waveform be measured at intervals of at least $(2\Delta\nu)^{-1}$ seconds, where $\Delta\nu$ is the bandwidth of the signal, if information is not to be lost. These samples must then be quantised so they are represented by a stream of bits, and the bit stream then laid down on magnetic tape. The bandwidth of this data stream is always greater than the that of the original signal, so, as the recording bandwidth is limited, the *information* rate in a digital

system is always less. A digital system throws away the information-carrying capability in the dynamic range of the recording, and so limits the maximum bandwidth of the original signal to one that, when sampled and quantised, will fit on the tape. The real attraction of this method is that once digitised, the signal becomes very resilient, and any subsequent processing can be controlled with great precision. This contrasts sharply with an analogue system which will in general contain components that are dispersive, drift with temperature and have a restricted dynamic range. Delay compensation and correlation are two essential functions which, although easily implemented with digital circuitry, have no simple analogue form.

The question therefore arises of what the best sampling scheme might be. Both the sampling rate and the number of quantisation levels determine how much information is lost in the digitisation process – a quantity to be minimised. However they also determine the accessible RF bandwidth for the observations, $\Delta\omega$, – a quantity to be maximised. The overall sensitivity will therefore be of the form

$$\gamma \propto \eta\sqrt{\Delta\omega}, \quad (2.11)$$

where γ is the correlated signal-to-noise ratio and η is a degradation factor introduced by the sampling procedure. If the bandwidth of the recorder is fixed at b , then 1-bit (i.e. 2-level) quantisation of the signal at the Nyquist rate degrades γ by $2/\pi$ ($\simeq 0.64$) [99] so that

$$\gamma \propto 0.64\sqrt{b/2}. \quad (2.12)$$

If 2-bit (4-level) quantisation were employed, also with Nyquist sampling and the same bit rate, the accessible RF bandwidth would be halved. However the finer quantisation levels would increase η to about 0.88 [11], giving

$$\gamma \propto 0.88\sqrt{b/4}, \quad (2.13)$$

which is down by 2% on the 1-bit case, and therefore corresponds to an overall *reduction* in sensitivity. Although 3-level quantisation can be shown to be in some sense ‘ideal’, the attraction of a simple encoding system combined with very reasonable performance has made 1-bit systems a natural choice in VLBI. This is also the scheme employed in the present system.

2.2.2 The fringe rotator

It is clear from the discussion in Chapter 1 that the system sensitivity of a mobile, low frequency interferometer is quite low, so it is vital that the signal-to-noise ratio is not degraded significantly at the processing stage. The correlator developed by

Spinks had been used to observe Cygnus A, and to collect amplitude data alone. Fringe rotation was therefore rather basic, consisting of a single $\cos \omega_g t$ factor used to multiply one of the two data streams. The other Fourier component of the signal was therefore lost, so reducing the signal-to-noise ratio by $\sqrt{2}$. Little regard had been given to the phase of the signal, and both the hardware and software generated mysterious jumps and ramps in the correlated phase.

Any further progress would rely on two major improvements being made. Firstly the fringe rotation had to be carried out properly. This could be done by installing a fully complex fringe rotator and would increase the overall signal-to-noise ratio by $\sqrt{2}$. Secondly the corrupting effects of the correlator had to be corrected for, so making the phase a more meaningful quantity.

The uncorrelated baseband signals can be thought of as having real and imaginary components, which will be denoted here by c and s . A complex correlator generates the product of two such signals from the real products of their components, i.e.

$$V_{1b}V_{2b}^* = c_1c_2 + s_1s_2 + i(s_1c_2 - c_1s_2) . \quad (2.14)$$

This signal has a geometrical fringe rate ω_g due to the Earth's rotation. Integrating such a signal for t_0 seconds therefore produces a drop in fringe visibility of $\text{sinc}(\omega_g t_0/2)$ and introduces a phase bias of $\omega_g t_0/2$. Ideally to avoid these effects, the signal would be rotated backwards by $e^{i\omega_g t}$ prior to integration. In a digital system this fringe unwinding function must be partially realised in hardware, as the initial data rates are too high for a computer to cope with. For the moment the digital approximation to this unwinding function will simply be denoted by

$$\mathcal{F}(\omega_g, t) = \mathcal{S}(\omega_g, t) + i\mathcal{C}(\omega_g, t) , \quad (2.15)$$

where \mathcal{C} and \mathcal{S} are the digital approximations to \cos and \sin functions.

The original correlator contained eight separate frequency channels. On each of these channels it generated the complex product of two factors over a range of 15 discrete delays. The correlation was done on boards that could compute a real product over eight delays. A complex product can be assembled from four real products, so each channel needed eight boards, and the whole correlator contained 64. These boards also integrated the products for one second before passing them, as 16-bit words, to the host computer via a direct memory access link. The computer, a pdp11, was unable to access more than about 70 of the boards in the time available between integration periods, so that if major modifications were to be avoided the number of boards in the new system would also have to be about 64. Unfortunately, complex fringe rotation generates eight real products per frequency channel rather

than four. To keep the data rate the same therefore, the number of delays had to be reduced from 15 to 7, giving a window of $56\mu\text{s}$ in which to find the fringes. On no occasion has the correlated signal been more than one delay ($8\mu\text{s}$) away from its expected position, so this reduction has not been a problem.

Once reconfigured, the correlator generated the products of all the components of V_{1b} , V_{2b} and \mathcal{F} over the range of seven delays and integrated each for one second. A schematic diagram for one delay is shown in Figure 2.2. The eight channels were divided between four frequencies (labelled CH 0,2,4 and 6) and two polarisations (L and R). With the notation

$$\begin{aligned} a &= \langle c_1 c_2 \mathcal{C} \rangle & b &= \langle c_1 s_2 \mathcal{S} \rangle & c &= \langle c_1 c_2 \mathcal{S} \rangle & d &= \langle c_1 s_2 \mathcal{C} \rangle \\ a' &= \langle s_1 s_2 \mathcal{C} \rangle & b' &= \langle s_1 c_2 \mathcal{S} \rangle & c' &= \langle s_1 s_2 \mathcal{S} \rangle & d' &= \langle s_1 c_2 \mathcal{C} \rangle \end{aligned} \quad (2.16)$$

the crudely fringe rotated output can be identified as

$$\langle \mathcal{F} V_{1b} V_{2b}^* \rangle_{t_0} = [(a + a') - (b - b')] + i[(c + c') - (d - d')]. \quad (2.17)$$

If the function \mathcal{F} acts on a 1-bit data stream, its components can have only three useful values: 1, 0 and -1. This makes \mathcal{C} and \mathcal{S} 3-level approximations to cosine and sine waves, their transition points being chosen to minimise the loss in signal-to-noise ratio introduced by the approximation. Only the first Fourier harmonic of $\mathcal{F}(t)$ correctly unwinds the correlated signal, so the degradation in the signal-to-noise ratio introduced by the unwinding process is ([91] pp 296)

$$\eta_f = \left[\frac{\text{power in the 1st harmonic}}{\text{power in } \mathcal{F}(t)} \right]^{\frac{1}{2}}, \quad (2.18)$$

which, for a function whose components have the form shown in Figure 2.3 becomes

$$\eta_f = \frac{4 \cos a}{\sqrt{2\pi(\pi - 2a)}}. \quad (2.19)$$

This factor is maximised when $a \simeq 2\pi/15.505$. In most practical systems \mathcal{F} is sampled discretely, with an integer number of samples per period, and with a equalling the width of one segment. Under these circumstances, the optimum number of segments is 16, giving a loss factor of 0.96. Due to constraints imposed by the existing electronics, the new system has a fringe function \mathcal{F} which is divided into 12 segments. This drops the signal-to-noise ratio by a further 0.5%, an amount too small to be of any concern. It is interesting to note that a 2-level fringe rotator can be generated by reducing a to zero. This would degrade the signal-to-noise ratio by a factor of 0.900.

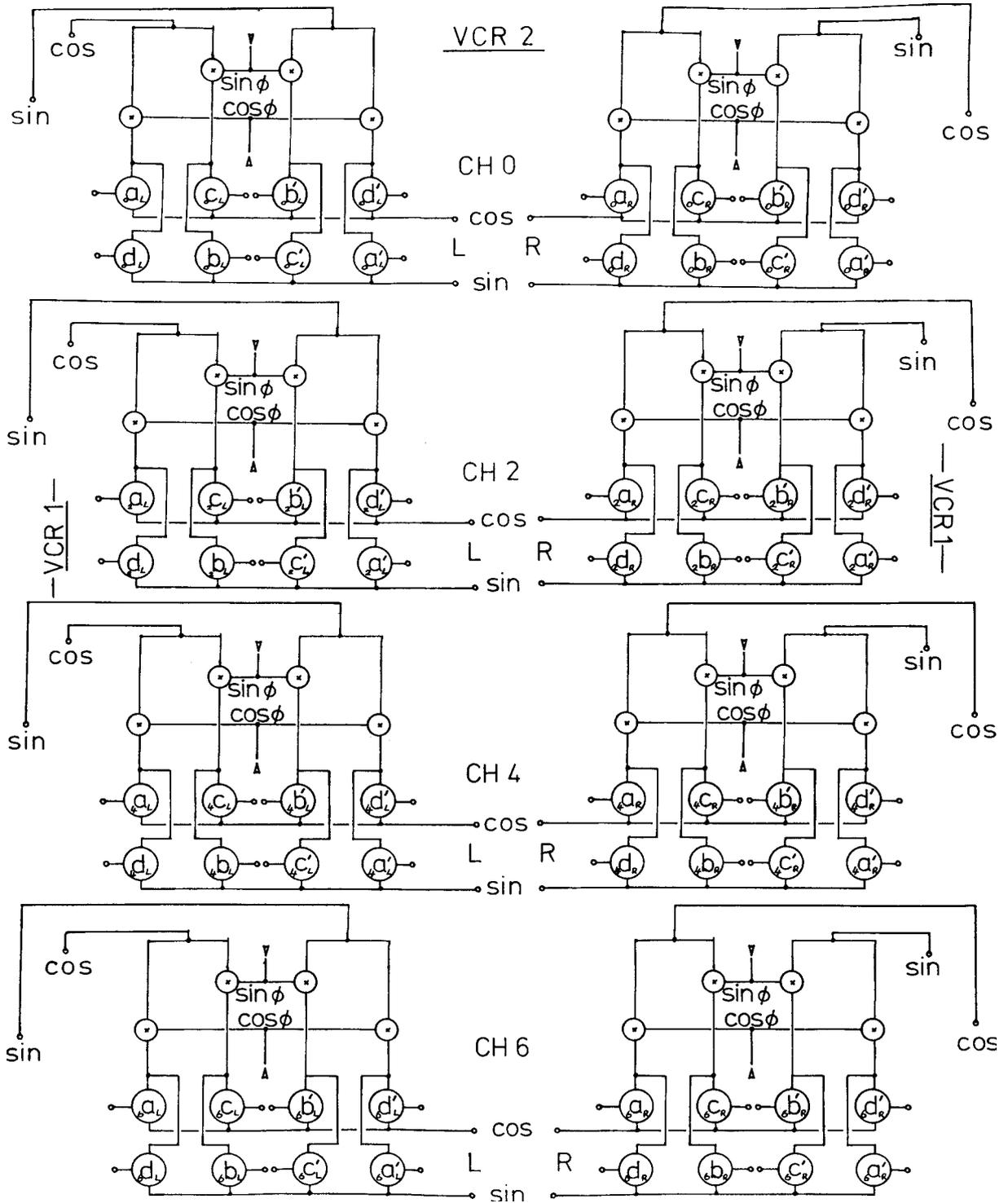


Figure 2.2: Schematic diagram showing how data on one of the seven correlator delays is fringe rotated and processed (see text for key).

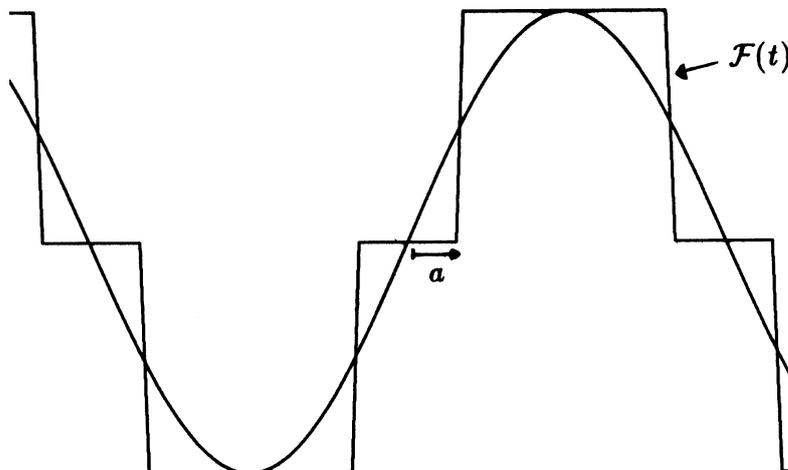


Figure 2.3: A 3-level approximation (labelled $\mathcal{F}(t)$) to a true sinusoid for use in digital fringe rotation. The parameter a is chosen to minimise signal loss.

2.2.3 Fringe function correction

The above analysis is only true for an infinite integration time, and a form for $\mathcal{F}(t)$ with exactly the right period to match the geometrical fringe rate. Neither of these are realised in practice, so the result must be modified. To see how to do this, consider a correlated measurement R consisting of a true signal Γ with fringe rate ω_g and an associated noise term, i.e.

$$R(t) = \Gamma \exp[i\omega_g t] + n(t) . \quad (2.20)$$

If the complex noise n is Gaussian, white and with a standard deviation of σ_0 (the root mean power of n), the signal-to-noise ratio of R is simply

$$\gamma = \Gamma/\sigma_0 . \quad (2.21)$$

Fringe rotating $R(t)$ with $\mathcal{F}(t)$ and averaging for t_0 seconds increases the signal-to-noise ratio. The process is most efficient when $\mathcal{F}(t) = \exp[-i\omega_g t]$. Any other form for \mathcal{F} will reduce the final signal-to-noise ratio by

$$\eta_f = \frac{|\langle \mathcal{F}(t) \exp[i\omega_g t] \rangle_{t_0}|}{|\langle n(t) \mathcal{F}(t) \rangle_{t_0}|} \cdot \frac{|\langle n(t) \exp[i\omega_g t] \rangle_{t_0}|}{1} . \quad (2.22)$$

As long as the integration time t_0 is a lot greater than the noise coherence time, this reduces to

$$\eta_f = \frac{|\langle \mathcal{F}(t) \exp[i\omega_g t] \rangle_{t_0}|}{\sqrt{\langle \mathcal{F} \mathcal{F}^* \rangle_{t_0}}} . \quad (2.23)$$

As well as \mathcal{F} degrading the signal-to-noise ratio, it also corrupts the amplitude and phase of the averaged signal. These effects can be included in a complex transfer

function for the digital fringe rotator, $G(\omega, t)$, so that a true estimate for the fringe visibility becomes

$$\Gamma_e(t) = \frac{1}{t_0 G(\omega_g, t)} \int_t^{t+t_0} \mathcal{F}(t') R(t') dt' . \quad (2.24)$$

G is generally a complicated function of both the fringe rate and the time at which the integration began. It contains factors from three sources:

1. The fringe unwinding frequency is generally quantised, so that there is usually a residual fringe rate present during integration.
2. A digital system introduces phase in steps. In a 3-level system these steps are 45° apart.
3. The number of phase steps introduced depends to a degree on the time at which fringe rotation begins.

Item (1) introduces sinc-like smearing effects for all fringe rates. Items (2) and (3) are both only important at low fringe rates, when the integration time is similar to the reciprocal of the fringe frequency. In this system, the hardware integration time is one second and fringe rates are usually about one Hz, so all three factors are important. The transfer function can be readily evaluated from

$$G(\omega, t) = (1/t_0) \int_t^{t+t_0} \mathcal{F}(t') \exp[i\omega t'] dt' . \quad (2.25)$$

Its form may be simplified if the phase of the fringe rotator is reset to zero at the start of the integration period, so that

$$G(\omega, t) = \frac{\exp[i\omega t]}{t_0} \int_0^{t_0} \mathcal{F}(t') \exp[i\omega t'] dt' . \quad (2.26)$$

The integral term is now just a function of the fringe rate, and can be conveniently tabulated for the chosen value of integration time. The phase factor outside the integral represents the fringe rate at the start of the integration, which can of course be calculated.

In this system, \mathcal{F} takes a straightforward, if rather awkward, form (see Appendix A). The amplitude and phase of $G(\omega)$, neglecting the phase factor at the front of Equation 2.26, are shown in Figure 2.4. This function was tabulated, and used to correct the data as it emerged from the correlator. For fringe rates greater than about 1.5 Hz the function becomes very periodic, so the table needed only to stretch to rates of about 2 Hz. The associated drop in signal-to-noise ratio, as calculated in Equation 2.23, is also shown in Figure 2.4. At the higher fringe rates η_f approaches the simple theoretical value of 0.955 .

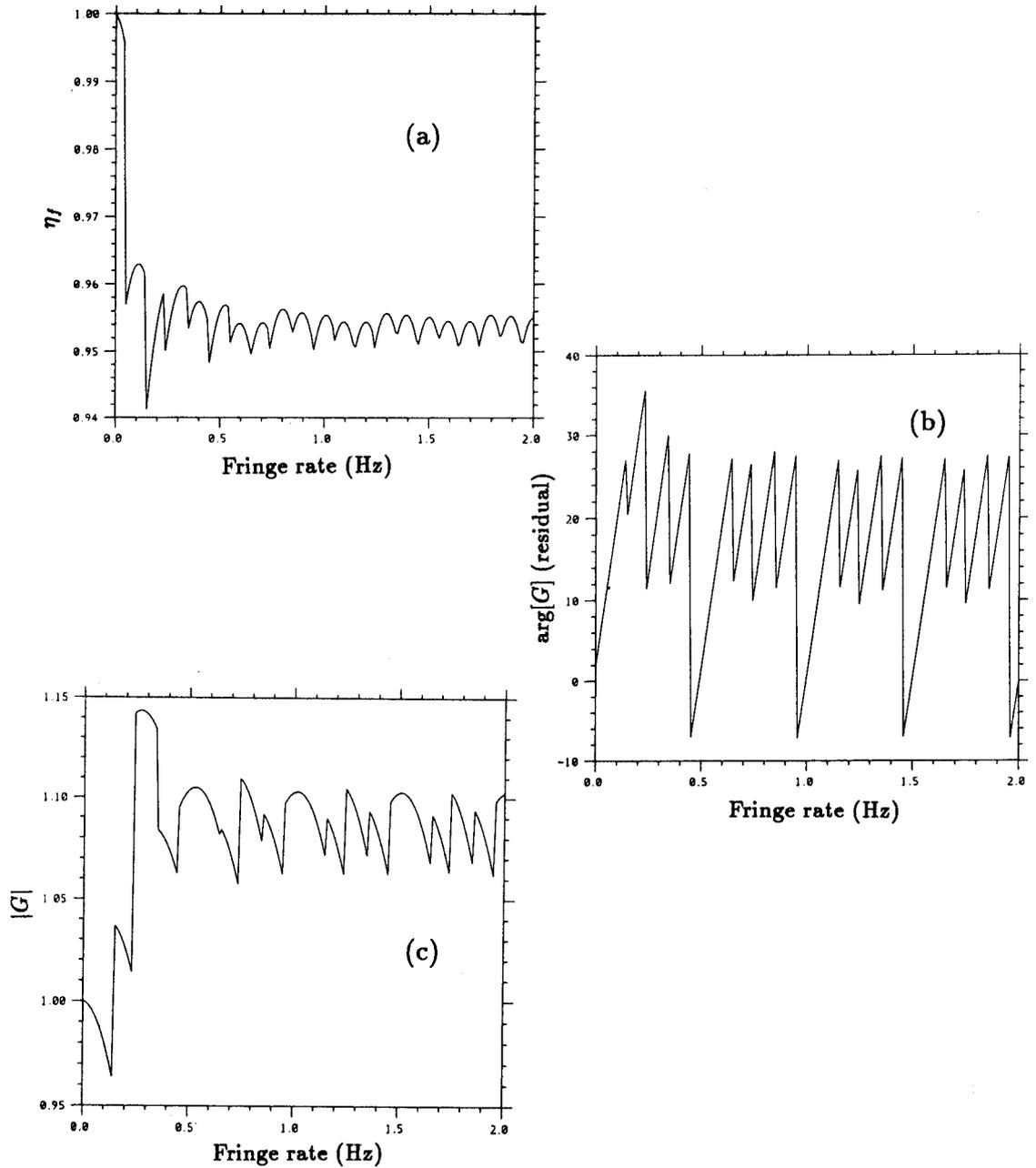


Figure 2.4: The amplitude (c) and phase (b) (in degrees) of the fringe rotator transfer function, $G(\omega)$, as a function of the fringe rate. (a) shows the drop in signal-to-noise ratio, η_f , introduced by the rotator.

2.3 A portable phased array

The system sensitivity has been further improved with the construction of a ‘phased array’ that can be used at the mobile station. The array consists of 16 Yagi antennas pointing at the zenith, arranged in two rows of eight (Figure B.1) and covering a physical area of 30m x 7.5m. The primary beam of each antenna is sufficiently wide for the array to be able to track a source for three hours either side of transit without major changes in the overall gain. For any particular source, beam effects can be calibrated out from zero baseline observations – a procedure already required to correct for the beam of the 4C antenna used at the base station. The antennas are connected, via cables of equal length, to a weatherproof box containing preamplifiers, phase rotators and matched splitter/combiners. The phase rotators are controlled by the microcomputer in the motorcaravan which manages the data recording.

The array can be used in two modes. In circumstances when differential Faraday rotation (DFR) can be neglected, the mobile station need only observe one (linear) polarisation, and all 16 Yagis can be phased together to give an effective area of about 100 m². However, if DFR is considered significant, the array can also be used to observe circular polarisation. In this mode the antennas in the North and South rows are aligned to be sensitive to N-S and E-W polarisations respectively. The software controlling the phase rotators maintains a 90° phase shift between the phase centres of the two rows throughout the observation (The details of this are presented in Appendix B). Left and right-handed circular polarisations can therefore be generated by either adding or subtracting the signals from the two rows. The effect of DFR on a circularly polarised component is to introduce a phase shift into the correlated signal equal (or opposite, depending on the polarisation) to the rotation angle. The correlated amplitude is not affected, a point of particular importance in LBI where a strong signal is required at all times if the fringes are to be tracked.

The phase rotators and, in a modified form, the preamplifiers used in the phased array were of a design employed by Rees and Duffett-Smith in the conversion of the 4C [72, 73]. Phase can be introduced by each phase rotator in steps of $\lambda/16$, using a series of π -networks which are switched in and out by diodes. This quantisation was found to have little effect on the forward gain of the array.

2.4 Closure phase

Possibly the most troublesome aspect of VLBI is its inability to deliver good phase information. This is mainly due to a combination of semi-coherent local oscillators and the atmosphere, the effects of which become more important over long baselines. As is explained more fully in Chapter 3, any attempt to rectify this with adaptive calibration techniques relies on each antenna being used in more than one baseline simultaneously. If an investigation is to be made of how these techniques can be used at low frequencies, at least three stations are required. It was therefore decided that another motorcaravan, identical to the first, should be built in an attempt to measure closure phases. This was done, and modifications were made to the correlator (see Appendix C) enabling the three resulting tapes to be correlated automatically. Tests made on Cygnus, over baselines of ~ 1 km, gave closure errors of less than a degree. It can be concluded that the correlator has been stabilised, and can now generate meaningful phase data.

2.5 Sensitivity

The noise on the output of the 1-bit correlator can be calculated quite easily, provided the correlated signal is weak. Each component (cos or sin) of each 125 kHz channel generates 125,000 bits per second. If there is no signal at all, and no systematic bias is present, these bits will each be randomly set at either 0 or 1. The correlator integrates these for one second, and (in this system) divides the result by eight to keep the numbers manageable. The result of this process is a number with a mean value of

$$\mu = \begin{array}{ccccccc} 125000 & \times & 0.5 & \div & 8 & = & 7812.5 \\ \uparrow & & \uparrow & & \uparrow & & \\ \text{No. samples} & & \text{mean of 0, 1} & & \text{division factor} & & \end{array}$$

and a standard deviation of

$$\sigma_{1\text{sec}} = \sqrt{125000} \times 0.5 \div 8 \simeq 22.1 .$$

The d.c. offset of μ (7812.5) is then subtracted, and the result averaged for 60 seconds, reducing σ by $\sqrt{60}$. In a fully complex system there are always two independent estimates of the same visibility component. If these are simply added, the noise goes up by $\sqrt{2}$, giving a final noise figure of $\sigma = 4.03$ in both the real and imaginary components of the visibility measurement. The RMS amplitude of this complex noise is $\sqrt{2}\sigma$, corresponding to 5.7 amplitude units on this scale.

The number of janskys of flux that this corresponds to depends on the sky and antenna temperatures for the observation. As long as the correlation coefficients are small, the noise level will remain at 5.7 units. Sensitivity calculations in conditions of high signal-to-noise ratio are considered in Section 6.4.3.

It was an important consideration in the design of the receiving equipment that the it should be kept simple. As a result, when the local oscillators phase lock to the rubidium standards they do so with unknown phase offsets. This greatly simplifies the design of the receivers, but means that there is no direct procedure for determining the relative phase differences between the correlated frequency channels. Coherent integration across the full 1 MHz bandwidth is therefore very difficult when the signal-to-noise ratio is low. An incoherent summation of the eight channels will increase the sensitivity by only $8^{1/4} = 1.68$. Under such conditions the sensitivity of the system is therefore only 1.68 times the sensitivity of a single channel.

If the system temperature is a lot greater than the antenna temperature, the correlated signal-to-noise ratio from a source of flux $S \text{ WHz}^{-1}\text{m}^{-2}$ is approximately

$$\gamma = \frac{\sqrt{A_1 A_2} S \eta_q \eta_f}{k \sqrt{T_{s1} T_{s2}}} \sqrt{\frac{\Delta \nu \tau}{2}}, \quad (2.27)$$

where η_q is the quantisation efficiency (0.64 in a 1-bit system) and η_f is the fringe rotation loss (~ 0.96 , see [56]). At 81.5 MHz, the system temperature equals the sky background temperature for all but the brightest of sources. The galactic background has been mapped with great care at this frequency by Bridle [15], and ranges between about 700 and 4000 K. The present experiments use instruments with large primary beams, so it would not be unreasonable to assign a figure of 2000 K on the system temperatures of both stations. When the mobile phased array is being used in circular polarisation mode, its effective area when correlated against the 4C antenna (E-W polarisation) is only the effective area of its East-West polarised dipoles. Furthermore, the North-South polarised dipoles add in uncorrelated noise, so the effective system temperature goes up by a factor of $\sqrt{2}$. The reasons for having this circular polarisation option will be described in Chapter 4. For the moment it should just be noted that when correlating between the mobile array and the 4C, the system sensitivity in circular mode is a factor of two less than in linear mode.

The linear mode sensitivity per frequency channel can now be calculated. Taking the effective areas of one arm of the 4C antenna and the 16-Yagi phased array as 2000 m^2 and 100 m^2 respectively, the signal-to-noise ratio in one frequency channel after 60s integration is

$$\gamma_{\text{lin}} \simeq 0.2S/\text{Jy}. \quad (2.28)$$

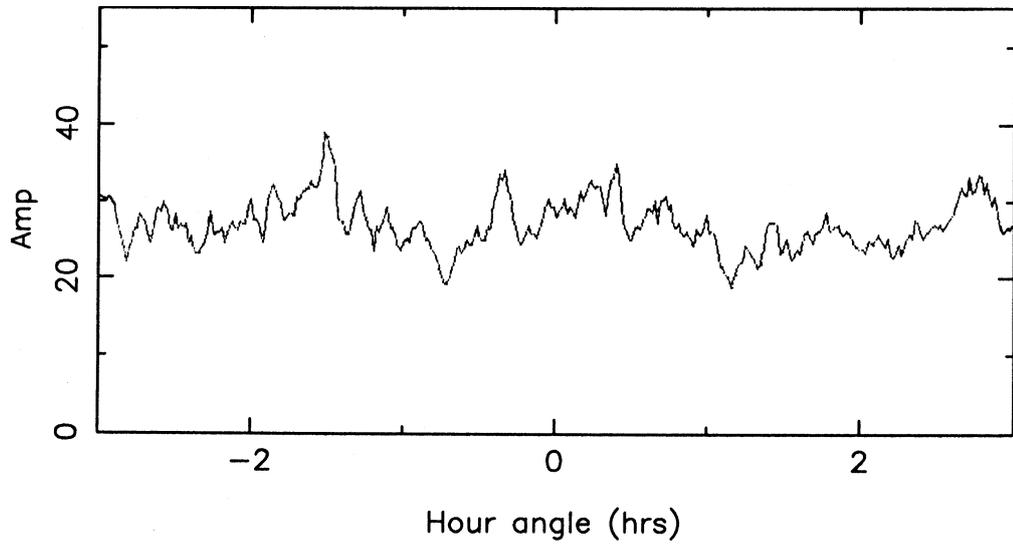


Figure 2.5: *Data collected on 3C48 using the phase array in circular polarisation mode. Although the data is the result of averaging a number of channels, the amplitude is displayed in units corresponding to the signal from a single channel.*

where S is the source flux in Janskys. Figure 2.5 shows signals from 3C48 (84 Jy) obtained using the array in circular polarisation mode. the theoretical signal-to-noise ratio is $0.2 \times 84 \div 2 = 8.4$, corresponding to a mean amplitude of about 45 units. The system sensitivity is therefore approximately 75% of its theoretical maximum value.

Chapter 3

Internal phase calibration

3.1 Adaptive calibration in VLBI

Radio interferometry, in common with most other forms of interferometry, suffers from what has come to be known as the ‘phase problem’; that is, although the intensity of a correlated signal may be readily measured, the phase of the signal is often corrupted or lost. These effects are of particular significance in long baseline interferometry due to two corrupting influences. Firstly, VLBI suffers the consequences of density fluctuations in the Earth’s troposphere and ionosphere which introduce unpredictable phase errors into the signals collected from widely separated elements of the interferometer, and lead to severe decorrelation. Secondly, the technique relies on local oscillators whose phase coherence is limited. These introduce both slowly varying offsets in the fringe rate and noise in the correlated phase. Although the fringe rate offset may be taken into account in processing, the phase noise cannot be dealt with so easily, and it is the limiting factor in determining the coherence time, and hence the sensitivity, of the interferometer.

The key to overcoming these difficulties is to realise that they can all be modelled as ‘antenna based’, that is to say they can be attributed to unknown and varying phase offsets at each antenna. In these circumstances, the complex spacial coherence function, R_{jk} , measured on the baseline comprising the j th and k th elements of an array is

$$R_{jk} = \exp[i(\varphi_j(t) - \varphi_k(t))]r_{jk} + n_{jk} , \quad (3.1)$$

where φ_j and φ_k are the unknown phase errors, r_{jk} the true correlation coefficient and n_{jk} an associated (complex) noise term. It was pointed out by Jennison [48] that phase errors of this kind were cancelled by summing the measured phases on

baselines in closed loops, giving a ‘true’ loop phase plus a noise term. This sum was named the ‘closure phase’ by Rogers [76] who was the first to apply Jennison’s idea to VLBI measurements.

The closure phase contains information about the *true* source structure, and is an observable unaffected by phase errors introduced by tropospheric and ionospheric refraction, local oscillator errors and inaccurate determination of both the antenna and source positions. It can therefore be used as a constraint in the determination of the source structure. Phase errors that do not factorise to the form of Equation 3.1 are called closure errors. These are dominated by the effects of mismatches in the signal paths and correlators on the different baselines. It has been shown by a number of authors (e.g. [103], [90], [71]) that these instrumental effects largely determine the dynamic range of maps generated with closure constraints.

Since the pioneering days of Jennison and later Rogers, there has been considerable progress in the search for the best strategy for applying the closure relations. Early attempts involved the use of trial maps in an iterative scheme, in which the phase of the map transform was adjusted to make it consistent with the closure data. This adjusted data was then used to generate a *hybrid map* [4] by Fourier inversion, and this hybrid map converted to a new trial map by applying a non-linear constraint (e.g. by CLEANing, applying positivity or windowing). This process was repeated until the hybrid and trial maps were indistinguishable. The technique, later to be known as ‘hybrid mapping’, was first used by Fort and Yee [32], but it was Readhead and Wilkinson [70] who devised the first practical scheme.

Although the closure phase formalism is of great historical significance in the development of new calibration techniques for quasi-stable synthesis arrays, it becomes clumsy when applied to multi-element telescopes. It is also difficult to take account of differences in the reliability of the individual visibility phases when they are explicitly combined in a closure relation. For these reasons the related technique of ‘self calibration’ was devised [20] [78], in which the complex antenna gains were considered as free parameters to be determined along with the ‘true’ visibilities. For an array of N antennas this reduces the problem to one of estimating $(N-1)$ antenna phases from $N(N-1)/2$ constraints, giving each constraint a weight determined by the noise in its measurement. In this way, Cornwell and Wilkinson [20] were able to specify the expected errors on each telescope of the MERLIN array as complex noises and avoided the explicit determination of closure phases. A fine review of the progress in these and related techniques up to 1983 has been made by Pearson and Readhead [60]. Recent work by Sivia [82] has produced a mapping package that includes the self calibration principle within a maximum entropy algorithm which

in turn uses the optimal gridding and degriding methods of Tan [88]. This package can be regarded as one of the best available for VLBI work.

3.2 Use in low frequency VLBI

It is important to note that neither the self calibration nor closure phase principles provide a linear constraint on the map transform. Because of this the result from, for example, a maximum entropy algorithm may not be unique, i.e., there may be many *local* maxima to which the algorithm could converge. This effect is especially prominent when there are only a few antennas observing simultaneously. By contrast, a calibrated aperture plane will always produce a unique solution, no matter how noisy or incomplete its sampling because it *does* apply a linear constraint on the final map. It is interesting that early applications of closure phase frequently consisted of attempts to derive true phases from redundant baselines (e.g. Twiss [96]) to simplify the interpretation of the data. However, as noted by Pearson and Readhead [60], this is not a good reason for building a redundant array today. With a reasonable number of antennas, it is usually better to arrange them to optimise the aperture plane coverage rather than to generate redundancy.

The best deployment strategy for the Cambridge VLBI system needs further thought. With only three antennas ('stations') operating at any one time, at least one station has to be fully mobile if reasonable aperture coverage is to be achieved. Another station can be fixed without losing any coverage capability. Furthermore this antenna can be large, thus increasing the sensitivity on two of the three baselines. The properties of the third antenna are not as easy to determine. Another mobile station would maximise the final aperture coverage, but would generate an insensitive mobile-to-mobile baseline. It would also double the manpower requirements for *each* observation, as these are dominated by the movement of the mobile stations and the collection of tapes from remote locations. If the third station were fixed, the third antenna could be large, and sensitivity would not be compromised. The problem then would be that the *rate* of aperture plane coverage would drop by a factor of three so that the total manpower needs would be greater than in the previous case. The siting of such a large antenna poses further problems. It is clear that the choice of which configuration to use depends heavily on the nature of the source being mapped.

The three-element system described in Chapter 2 was first used to investigate reports of arcminute structure around the quasar 3C48, as made by Hartas some three years earlier [41]. The mapping of such large-scale diffuse structure around a

strong point source allows for an alternative calibration scheme. If the structure is *totally* resolved out on two of the three baselines, then the closure phase around the triangle will equal the true astronomical phase on the third baseline – a variant on the redundant baseline idea. Clearly, to meet this condition the baseline triangle must be ‘thin’, i.e. have one side much shorter than the other two. For such a geometry, the closure phase represents a linear (Fourier) constraint on the map. This phase, together with the amplitude from the shortest baseline, could generate a dirty map through a simple Fourier inversion, and would remove the need for the non-linear techniques described in the previous section.

This ‘thin-triangle’ scheme requires a single large antenna a considerable distance from the other two, and it was fortunate that in the summer of 1986 such an antenna was available in the form of the ‘Wardle’ station of MERLIN, freed during a period of maintenance. The method has two further attractions. Firstly, it removes the need to construct a large phased array to serve the third station, and secondly, the measurements yield the true visibility function directly and so are in principle uniquely invertible. This second point was of particular importance considering the small number of simultaneous baselines available.

The results of these observations will be presented in Chapter 4. It is at this point that two major limitations of the method should be considered.

3.3 Applicability of the ‘thin-triangle’ method

If this technique is to be appropriate to the mapping of extended structure around a bright compact source, it is vitally important that the extended emission be totally resolved out by two of the three baselines. Consider as an example a 1-dimensional source with a brightness distribution as represented in Figure 3.1, containing an unresolved central component contributing a fraction $(1 - f)$ to the total flux and an extended component providing the remainder. If the extended structure has a spacial profile $H(\theta)$, where $\int H(\theta)d\theta = 1$, the normalised brightness of the source is

$$B(\theta) = (1 - f)\delta(\theta) + fH(\theta) , \quad (3.2)$$

and the corresponding visibility as a function of baseline (u) is

$$\Gamma(u) = 1 - f + f\tilde{H}(u) , \quad (3.3)$$

where \tilde{H} is the Fourier transform of H . Assuming that the extended emission contributes only a small fraction to the total flux (i.e. $f \ll 1$), the phase fluctuations

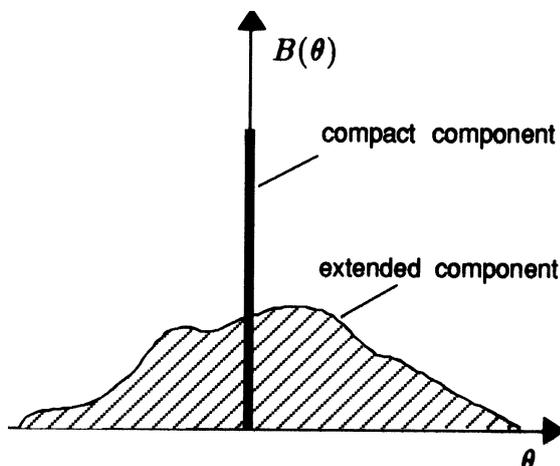


Figure 3.1: Simple one dimensional source with both a compact and extended component.

characteristic of the shorter baselines will be approximately $f\tilde{H}(0)$. On the longest baseline, u_0 , on which the extended emission is assumed to be mainly resolved, the fluctuations will be approximately $f\tilde{H}(u_0)$. If this long baseline is to provide a good phase reference, these fluctuations must be much less than those on the shorter baselines, i.e. $\tilde{H}(u_0) \ll \tilde{H}(0)$. In the case of 3C48, the Hartas evidence is that $|\tilde{H}(u_0)| \ll |\tilde{H}(0)|$ on baselines of more than a few thousand wavelengths. As Wardle is over 50,000 λ away from Cambridge, it will be assumed that the condition is met, and that the astronomical phase derived from both the Wardle-Cambridge and Wardle-mobile baselines is entirely due to the central unresolved component.

3.4 Determination of closure phase in noisy conditions

The arguments presented in section 3.2 demonstrate the advantages, under certain conditions, of evaluating closure phases around thin baseline triangles. However, the presence of the noise term in Equation 3.1 leads to the question of how best to evaluate the closure phase when the signal-to-noise ratio on one or more of the contributing baselines is low. It has already been pointed out that the sensitivity of a low frequency instrument using a portable antenna is necessarily limited, so this point needs particular attention.

The subject of phase averaging is a tricky one. The following discussion will consider two ‘robust’ methods as applied to the processing of both conventional and closure phases. One involves the assignment of the phases or closure phases to unit

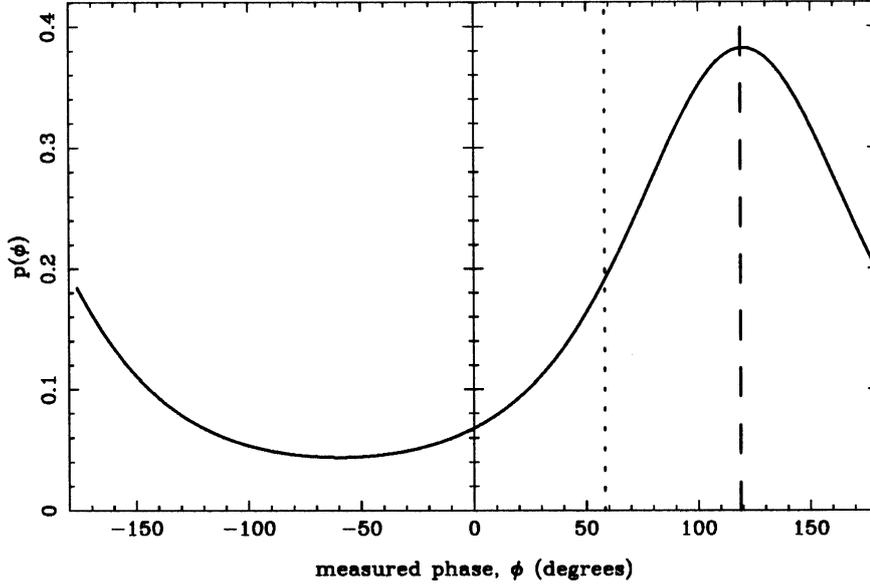


Figure 3.2: Probability distribution of the phase derived from a weak signal (with a phase of 120°) embedded in Gaussian noise ($\gamma = 0.6$). The most probable phase is marked by a dashed line and the mean phase by a dotted line.

vectors before averaging (the ‘unit vector’ method) whilst the other preserves a link with the original visibility vectors from which the phases are derived. It will be shown that the second method is as much as a factor of two better than the first, under very noisy conditions.

3.4.1 Phase noise in a single baseline

Consider a baseline on which the correlated signal-to-noise ratio is γ . If the noise is Gaussian and white, the probability of a phase error, θ , in the measured phase, ϕ , is

$$p(\theta, \gamma) = \frac{1}{2\pi} e^{-\gamma^2} [1 + \gamma\sqrt{\pi} \cos\theta e^{\gamma^2 \cos^2\theta} (1 + \operatorname{erf}(\gamma \cos\theta))] \quad (3.4)$$

(see Appendix D). This error cannot be reduced by averaging the phase over a number of samples as the periodic nature of $p(\theta)$ results in the mean value of ϕ not equalling its most probable ϕ_0 except in the case when $\phi_0 = 0$ (Figure 3.2). Instead an average of a vector representing the phase must be made. The most natural choice for the amplitude of this vector is, of course, the amplitude of the original visibility vector from which the phase was derived $|R_{ij}|$. The averaged phase will then be the phase of the visibility vectors averaged over N samples, $\arg[\langle R_{ij} \rangle_N]$. This vector average can be considered as two separate arithmetic averages of its

real and imaginary components, each of which will be normally distributed. The complex noise in the result will again be Gaussian, so that the phase noise will have the form of Equation 3.4 but with $\gamma = \gamma_0\sqrt{N}$ where γ_0 is the signal-to-noise ratio in a single sample. This result is not surprising, it is the equivalent of averaging the complex visibility before extracting the phase.

A similar improvement can be made by working with the phases alone, that is without any further reference to the original visibility, if they are assigned to unit vectors and averaged as before. If the signal lies along the x-axis, then for a single unit vector

$$\begin{aligned}\langle x \rangle_0 &= \int p(\theta) \cos \theta d\theta, & \langle y \rangle_0 &= 0, \\ \sigma_{0x}^2 &= \int p(\theta) \cos^2 \theta d\theta - \langle x \rangle_0^2, & \sigma_{0y}^2 &= \int p(\theta) \sin^2 \theta d\theta,\end{aligned}\quad (3.5)$$

where $\langle \dots \rangle$ denotes the taking of a mean, so that $\langle x \rangle_0$ and $\langle y \rangle_0$ are the expected real and imaginary components of the unit vector, and σ^2 is the corresponding variance.

If N of these vectors are now added, and if N is large, the probability density functions (pdfs) of the components x, y , in the sum will be

$$p(x) = \frac{1}{\sqrt{2\pi N}\sigma_{0x}} \exp \left[\frac{-(x - N\langle x \rangle_0)^2}{2N\sigma_{0x}^2} \right], \quad (3.6)$$

and

$$p(y) = \frac{1}{\sqrt{2\pi N}\sigma_{0y}} \exp \left[\frac{-y^2}{2N\sigma_{0y}^2} \right]. \quad (3.7)$$

As long as $\sigma_{0x} \simeq \sigma_{0y} (= \sigma_0)$, the phase noise in this ‘unit vector’ sum will again have a distribution with the form $p(\theta, \gamma')$ (Equation 3.4), where

$$\gamma' = \frac{\langle x \rangle_0 \sqrt{N}}{\sigma_0 \sqrt{2}}. \quad (3.8)$$

From this exercise it can be concluded that both methods of vector averaging decrease the phase noise in the result, and have similar forms of phase pdf. However, it is better to average the visibility vectors themselves than to average unit vectors with the phases of the visibility vector. Indeed, the unit vector average is worse by a factor

$$\eta = \frac{\langle x \rangle_0}{\gamma_0 \sigma_0 \sqrt{2}}, \quad (3.9)$$

in the final value of γ (see Figure 3.3).

In the limiting case when the signal-to-noise ratio tends to zero, Equation 3.4 reduces to

$$p(\theta, \gamma) \simeq \frac{1}{2\pi} (1 + \gamma\sqrt{\pi} \cos \theta). \quad (3.10)$$

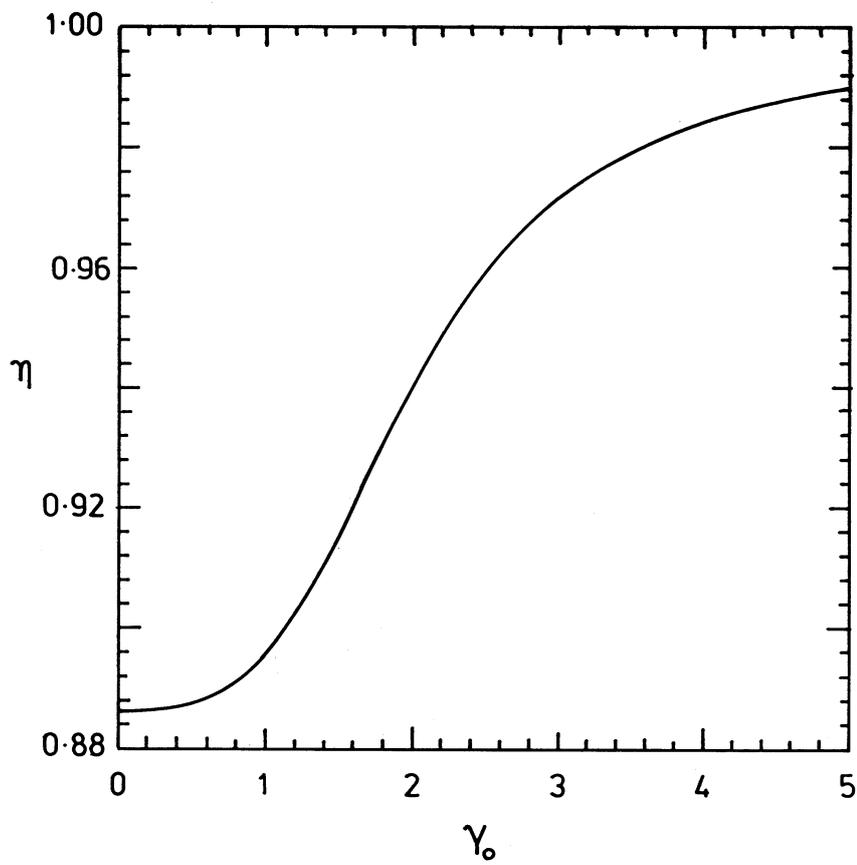


Figure 3.3: The efficiency factor, η , relating unit vector phase averaging techniques to those using the full visibility vector.

Using this distribution, the efficiency factor can be calculated from Equation 3.9 to give

$$\eta = \frac{\sqrt{\pi}}{2} \simeq 0.89 . \quad (3.11)$$

When the signal-to-noise ratio is very high, Equation 3.4 approaches

$$p(\theta, \gamma) \simeq \frac{1}{\sqrt{\pi}} \gamma e^{-\gamma^2 \theta^2} \quad (3.12)$$

and so $\eta \simeq 1$. Under these conditions the two methods are indistinguishable.

3.4.2 Phase noise in a closure relation

Having investigated the behaviour of phase noise on a single baseline, we can now approach the problem of how best to evaluate a closure phase given noisy visibility data around a loop. It should be pointed out here that the following methods extend to any process that involves the summing of phases. The only ‘special’ property of closure phases in this context is that they are observables that change slowly with time, and are therefore suitable for averaging.

The closure phase around the loop is evaluated once the phases on each baseline have been averaged for as long as experimental conditions permit. The next step is of central importance, as it is the point at which the phase ambiguities due to the noise are compounded. We must add together the phases derived from each of the three baselines, with the knowledge that this process is carried out modulo 2π . It is a standard result in modulo arithmetic that the modulo operator is distributive. That is to say that if

$$\Phi = \phi_1 + \phi_2 + \phi_3 \quad (3.13)$$

then

$$[\Phi]_{\text{mod } 2\pi} = [\phi_1]_{\text{mod } 2\pi} + [\phi_2]_{\text{mod } 2\pi} + [\phi_3]_{\text{mod } 2\pi} . \quad (3.14)$$

Phase ambiguities cannot therefore be introduced into the *signal* phase by modulo addition. However the phase *noise* is ‘wrapped around’ by the summing process, and this can generate a significant drop in the overall signal-to-noise ratio. It will be shown in this section that the periodic nature of $p(\theta)$ means that the variance in a sum of independent phase measurements only approaches the sum of the individual variances under conditions of high signal-to-noise ratio. Indeed, when the signal-to-noise ratio is very low ($\ll 1$) the variance of the sum goes as the product of the variances.

To see this more clearly, consider the phase noise in the sum of just two phases, ϕ_1 and ϕ_2 , such that

$$\begin{aligned}\phi_1 &= \phi_{1,0} + \theta_1 \\ \phi_2 &= \phi_{2,0} + \theta_2 ,\end{aligned}\tag{3.15}$$

where θ_1 and θ_2 represent phase noises with zero mean. If the probability distributions of θ_1 and θ_2 are $p_1(\theta_1)$ and $p_2(\theta_2)$, then the pdf for the noise in the sum of ϕ_1 and ϕ_2 is

$$p_+(\theta) = \int_{-\pi}^{\pi} p_1(\theta') p_2(\theta - \theta') d\theta' ,\tag{3.16}$$

where $\theta = \theta_1 + \theta_2$, and $p(\theta + 2\pi n) = p(\theta)$. With suitable normalisation, this can be regarded as the convolution of two periodic functions, p_1 and p_2 . These functions can be decomposed into their Fourier series:

$$p(\theta) = \frac{1}{2\pi} \sum_{-\infty}^{\infty} A_n e^{in\theta} ,\tag{3.17}$$

with

$$A_n = \int p(\theta) e^{in\theta} d\theta .\tag{3.18}$$

The first two Fourier coefficients can then be identified as

$$A_1 = \langle x \rangle\tag{3.19}$$

$$A_2 = 1 - 2\langle y^2 \rangle ,\tag{3.20}$$

where

$$\langle x \rangle = \int p(\theta) \cos \theta d\theta\tag{3.21}$$

$$\langle y^2 \rangle = \int p(\theta) \sin^2 \theta d\theta .\tag{3.22}$$

Taking p to be the distribution given by Equation 3.4 we can, with reference to Equation 3.9, use the result just derived for unit-vector averaging of a single baseline's phase to express the value of γ that characterises the phase noise on the baseline as

$$\gamma = \frac{1}{\eta} \frac{\langle x \rangle}{\sqrt{2\langle y^2 \rangle}} = \left[\frac{A_1^2}{1 - A_2} \right]^{1/2} \frac{1}{\eta} .\tag{3.23}$$

As has been pointed out, η ranges from 1 when γ is large to $2/\sqrt{\pi}$ when γ is small. Equation 3.23 is a useful tool for evaluating the phase noise in the sum of the phase angles. Application of the convolution theorem to Equation 3.16 gives

$$p_+(\theta) = \frac{1}{2\pi} \sum A_n^{(1)} A_n^{(2)} e^{in\theta}\tag{3.24}$$

where the bracketed superscripts denote the two phase pdfs. This equation can now be used to relate the ‘variance’ of $p_+(\theta)$, that is $2\langle y^2 \rangle_+$, and the square of the ‘mean’, $\langle x \rangle_+^2$, to the signal-to-noise ratios characteristic of the two constituent phases. The ratio of these reduces to

$$\frac{2\langle y^2 \rangle_+}{\langle x \rangle_+^2} = \frac{1}{\eta_1^2 \langle x \rangle_2^2 \gamma_1^2} + \frac{1}{\eta_2^2 \langle x \rangle_1^2 \gamma_2^2} - \frac{1}{\eta_1^2 \eta_2^2 \gamma_1^2 \gamma_2^2}. \quad (3.25)$$

In general, if the phase sum is the result of adding M phases, this ratio will be

$$\frac{2\langle y^2 \rangle_+}{\langle x \rangle_+^2} = \frac{\prod \gamma_i^2 - \prod (\gamma_i^2 - \langle x \rangle_i^2 / \eta_i^2)}{\prod \gamma_i^2 \langle x \rangle_i^2}, \quad (3.26)$$

where the symbol \prod represents a product over i ($i = 1 \rightarrow M$). When dealing with closure phases, we usually have the sum of just three phases, i.e., $M = 3$. The following arguments will be greatly simplified if it is assumed that all three baselines have the same sensitivity, and that the noise on the underlying visibility is Gaussian. Taking the characteristic signal-to-noise ratio of a single baseline as γ_s , Equation 3.26 now becomes

$$\frac{2\langle y^2 \rangle_+}{\langle x \rangle_+^2} = \frac{3}{\gamma_s^2 \langle x \rangle_s^4 \eta^2} - \frac{3}{\gamma_s^4 \langle x \rangle_s^2 \eta^4} + \frac{1}{\gamma_s^6 \eta^6}, \quad (3.27)$$

where $\langle x \rangle_s = \int p(\theta) \cos \theta d\theta$, and again η is between .89 and 1.

This ratio can be loosely associated with the reciprocal of the square of the ‘signal-to-noise ratio’ for the sum of the three phases. The precise interpretation will become clear later. Now that the closure phase has been evaluated, it must be integrated. The methods for averaging phases have been discussed in the previous section. There is a similar choice now of the type of vector with which to associate the closure phases as there was then. One option, as before, is to assign them to unit vectors and average these. The argument proceeds as before. The central limit theorem tells us that provided sufficient vectors are averaged, the resulting complex noise will again be Gaussian, and the signal-to-noise ratio after an average of N samples will be

$$\gamma_+ = \frac{\langle x \rangle_+ \sqrt{N}}{\sqrt{2\langle y^2 \rangle_+}}. \quad (3.28)$$

The subscript $+$ has been retained as a reminder that this is the result of an explicit evaluation of the closure phases prior to assignment to unit vectors and averaging. In terms of the properties of the constituent baseline phases, the signal-to-noise ratio of the vector whose argument is our integrated estimate of the closure phase is

$$\frac{1}{\gamma_+^2} = \frac{1}{N} \left[\frac{3}{\gamma_s^2 \langle x \rangle_s^4 \eta^2} - \frac{3}{\gamma_s^4 \langle x \rangle_s^2 \eta^4} + \frac{1}{\gamma_s^6 \eta^6} \right]. \quad (3.29)$$

The noise in the estimate of the closure phase will be have the distribution of Equation 3.4, i.e. $p(\theta, \gamma_+)$.

So how does this method perform? When the baseline signal-to-noise ratio is low ($\gamma_s \ll 1$), it is appropriate to take η as $\sqrt{\pi}/2$. Under these conditions the above expression for γ_+ reduces to

$$\gamma_+ \simeq \frac{\pi^{3/2}}{8} \gamma_s^3 \sqrt{N}. \quad (3.30)$$

The point to notice here is that the closure noise is proportional to the *cube* of the single baseline noise. In weak signal conditions, closure phase is clearly a very noisy quantity, and it is very important that the measurements from the baselines are integrated individually for as long as possible. Only in the limit of large γ_s is the loss of information from summing the phases sufficiently small for γ_+ to vary as

$$\gamma_+ \simeq \sqrt{N} \frac{\gamma_s}{\sqrt{3}}. \quad (3.31)$$

The second method of integrating the closure phases is to assign them to vectors whose amplitudes equal the product of the three baseline amplitudes. This is entirely equivalent to the method suggested by Cornwell [19] in which, by analogy with optical speckle-imaging techniques, the ‘bi-spectrum’ or ‘triple product’ (C_{123}) of the three visibility vectors in the baseline loop is evaluated and averaged. If φ_j is the phase error associated with the j th antenna then

$$\begin{aligned} C_{123} &= R_{12}R_{23}R_{31} \\ &= (r_{12}e^{i(\varphi_1-\varphi_2)} + n_{12})(r_{23}e^{i(\varphi_2-\varphi_3)} + n_{23})(r_{31}e^{i(\varphi_3-\varphi_1)} + n_{31}), \end{aligned} \quad (3.32)$$

which can be written as

$$\begin{aligned} C_{123} &= r_{12}r_{23}r_{31} + n_{12}n_{23}n_{31} \\ &+ r_{12}n_{23}n_{31}e^{i(\varphi_1-\varphi_2)} + r_{12}r_{23}n_{31}e^{i(\varphi_1-\varphi_3)} \\ &+ r_{23}r_{31}n_{12}e^{i(\varphi_2-\varphi_1)} + r_{23}n_{31}n_{12}e^{i(\varphi_2-\varphi_3)} \\ &+ r_{31}r_{12}n_{23}e^{i(\varphi_3-\varphi_2)} + r_{31}n_{12}n_{23}e^{i(\varphi_3-\varphi_1)}. \end{aligned} \quad (3.33)$$

If the noise processes have zero mean, all but the first of the above terms vanish and the expectation value of C_{123} is just $r_{12}r_{23}r_{31}$. Hence the argument of C_{123} equals the closure phase. The whole process is analogous to full visibility averaging for a single baseline, as described in Section 3.4.1. The variance of the triple product can be readily evaluated using the general result: $\text{var}[uv] = \langle u^2 \rangle \langle v^2 \rangle - \langle u \rangle^2 \langle v \rangle^2$, if u and v are uncorrelated variables. Hence if $\sigma_{ij}^2 = \text{var}[r_{ij}]$

$$\begin{aligned} \text{var}[C_{123}] &= r_{12}^2 r_{23}^2 \sigma_{31}^2 + r_{12}^2 \sigma_{23}^2 \sigma_{31}^2 + \sigma_{12}^2 \sigma_{23}^2 \sigma_{31}^2 \\ &+ r_{23}^2 r_{31}^2 \sigma_{12}^2 + r_{23}^2 \sigma_{31}^2 \sigma_{12}^2 \\ &+ r_{31}^2 r_{12}^2 \sigma_{23}^2 + r_{31}^2 \sigma_{12}^2 \sigma_{23}^2. \end{aligned} \quad (3.34)$$

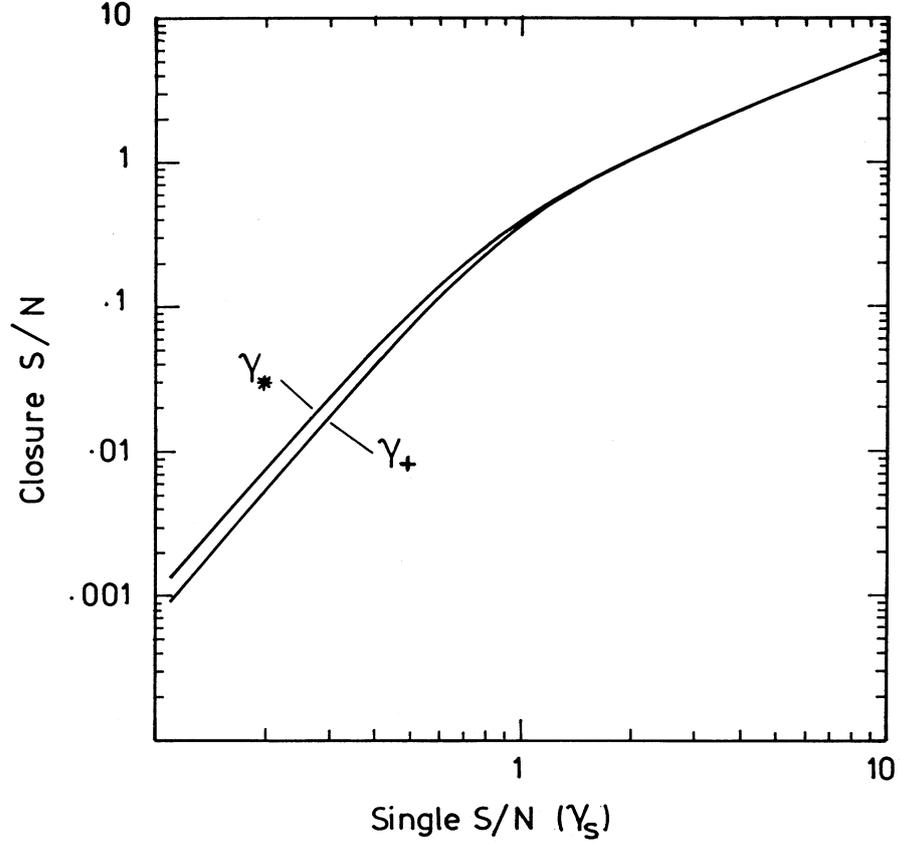


Figure 3.4: A comparison of the ‘triple product’ method of evaluating closure phase with the ‘unit vector’ method. Note that for single-baseline signal-to-noise ratios $\gtrsim 1$, the two methods are equally good.

Again, this expression can be greatly simplified by setting variances and signals on the three baselines equal, so they each have a signal-to-noise ratio of γ_s . If such a triple product is vector-averaged over N samples, the signal-to-noise ratio in the result (γ_*) will be related to γ_s by

$$\frac{1}{\gamma_*^2} = \frac{1}{N} \left[\frac{3}{\gamma_s^2} + \frac{3}{\gamma_s^4} + \frac{1}{\gamma_s^6} \right]. \quad (3.35)$$

This ‘triple product’ result should be compared with Equation 3.29, which is the equivalent result for the unit vector method. Equations 3.29 and 3.35 are compared graphically in Figure 3.4. The general expression for the signal-to-noise ratio of the product of M complex measurements can be calculated by a similar approach to that used above. If the signal-to-noise ratio of the i th measurement is γ_i , the

signal-to-noise ratio of the product is specified by

$$\left(\frac{1}{\gamma_{*,M}^2} + 1\right) = \prod_{i=1}^M \left(\frac{1}{\gamma_i^2} + 1\right). \quad (3.36)$$

When the signal is strong, $\gamma_* \simeq \gamma_s \sqrt{N/3}$, which is the same result as in the ‘unit vector’ case. For low signal-to-noise ratios Equation 3.35 reduces to

$$\gamma_* = \gamma_s^3 \sqrt{N}. \quad (3.37)$$

Under such weak signal conditions the triple product method can be seen to be an improvement on the unit vector method by a factor of $8/(\pi^{3/2}) \simeq 1.44$. Furthermore, if the contributing baseline phases were themselves averaged by the unit vector method prior to addition, the resulting signal-to-noise ratio could be lower by a factor of as much as $(2/\sqrt{\pi})^6 \simeq 2.1$ than that generated by the full visibility and triple product methods.

As well as triple product averaging performing better in noisy conditions, it is also computationally quicker than the unit vector method. The arguments above clearly show that if baseline amplitudes are available, as they almost always are, they should be included in the determination of closure phase. Furthermore, the triple product method provides a natural way in which this may be achieved.

3.4.3 The effect of phase slopes

The previous section has shown how important it is that the signal-to-noise ratio of the three baseline visibilities is greater than one before evaluating a triple product. The length of time over which these visibilities can be integrated is usually determined by the limited coherence of the interferometer rather than the astronomy. However, it is also important that the closure phase relation is preserved through the integration, so that

$$\arg[r_{12}r_{23}r_{31}] = \arg \left[\int r_{12} e^{i(\varphi_1 - \varphi_2)} dt. \int r_{23} e^{i(\varphi_2 - \varphi_3)} dt. \int r_{31} e^{i(\varphi_3 - \varphi_1)} dt \right]. \quad (3.38)$$

Assuming that the astronomy (r_{jk}) does not change in the integration period, this is equivalent to constraining the antenna phases so that

$$\arg \left[\int e^{i(\varphi_1 - \varphi_2)} dt. \int e^{i(\varphi_2 - \varphi_3)} dt. \int e^{i(\varphi_3 - \varphi_1)} dt \right] = 0. \quad (3.39)$$

If the antenna phases only depend to first order on t (i.e. $\varphi = a_0 + a_1 t$) over the integration period, this expression reduces to a product of three real sinc functions,

and the condition is met. If however the antenna phases are quadratic in t the expression becomes the product of three Fresnel integrals. Such a product is generally complex, and its phase represents a closure error. Although this appears to be a severe limitation, it turns out that closure errors generated by this mechanism are quite small. If for example the three baseline phases increase quadratically to maximum values of 20, 5 and -25 degrees over the integration period, the resulting closure error is still < 0.02 degrees.

Therefore the precombinational integration period is determined not by how much closure error it generates but on the coherence times of the signals. For the sake of simplicity, we will only consider antenna phase errors that generate unknown but *constant* fringe rates in the baseline visibilities over the integration period. The result of integrating a constant signal S with a fringe rate of α radians per second can be readily calculated. Taking the noise as Gaussian the pdf for the n th sample, with a sampling period of τ , is

$$p(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x - S \cos n\alpha\tau)^2}{2\sigma^2}\right) \exp\left(-\frac{(y - S \sin n\alpha\tau)^2}{2\sigma^2}\right). \quad (3.40)$$

If N of these samples are summed, the polar distribution of the result is

$$p_N(R, \phi) = \frac{1}{\sqrt{2\pi N}\sigma} \exp\left[-\frac{R^2 + \varepsilon^2 S^2 - 2RS\varepsilon \cos(\phi - \beta)}{2N\sigma^2}\right], \quad (3.41)$$

where

$$\varepsilon = \frac{\sin \frac{N\alpha\tau}{2}}{\sin \frac{\alpha\tau}{2}}, \quad \beta = \frac{N\alpha\tau}{2}. \quad (3.42)$$

A straight comparison with Appendix D shows that this result has the same phase noise distribution described by Equation 3.4 but with

$$\gamma = \frac{\gamma_s}{\sqrt{N}} \frac{\sin(N\alpha\tau/2)}{\sin(\alpha\tau/2)}, \quad (3.43)$$

where γ_s is the signal-to-noise ratio in a single visibility sample. Although the phase of the signal has been rotated by β in the integration process, the sum of the β s on the three baselines is zero, so the closure phase is unaffected. However, since the single baseline signal-to-noise ratio is no longer proportional to \sqrt{N} , the length of integration period employed before forming the triple product needs to be reconsidered. Suppose there are M visibility samples on each of three baselines from which we want to evaluate one closure phase by the triple product method. Suppose further that the signal-to-noise ratio in any single visibility sample can be characterised by γ_s and that its unknown fringe rate is α . The closure constraint of $\alpha_1 + \alpha_2 + \alpha_3 = 0$ will be neglected as it does not affect the quality of the result.

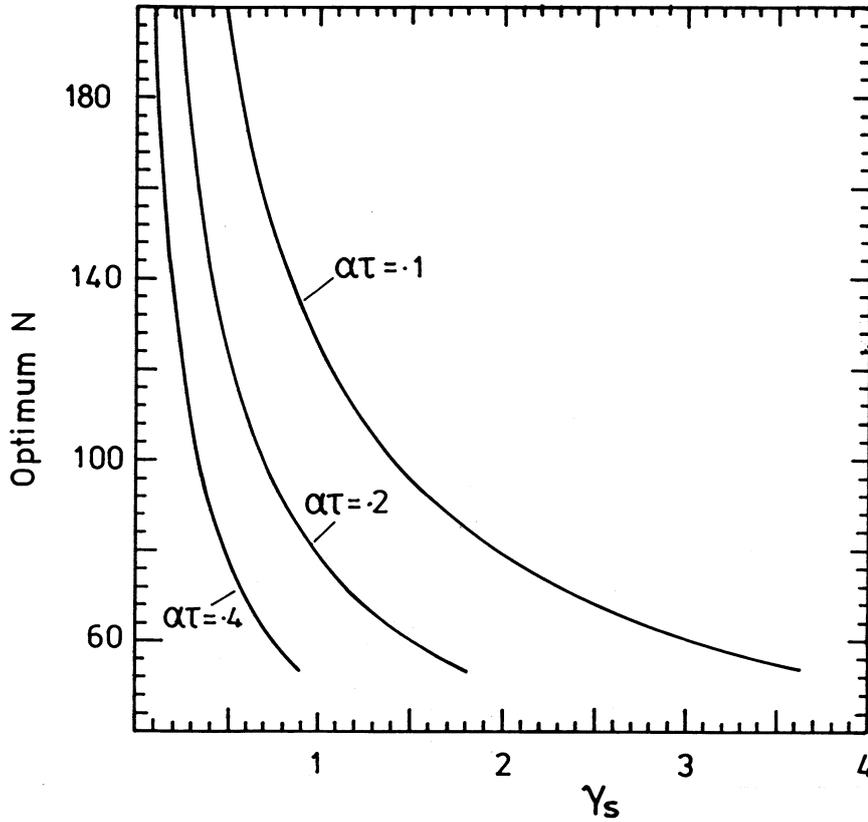


Figure 3.5: The optimum number of samples over which the single-baseline visibilities ought to be averaged before combining in a triple product, when phase slopes are present. Results for slopes of 0.1, 0.2 and 0.4 degrees per sample are shown.

Each baseline visibility is averaged over N samples and then combined to form M/N triple products which are themselves averaged. The value of N that minimises the noise in the final closure phase will depend on both α and γ_s . Figure 3.5 shows solutions to the equation

$$\frac{d}{dN} \left(\frac{1}{\gamma_*^2} \right) = 0, \quad (3.44)$$

for phase slopes of 0.1, 0.2 and 0.4 degrees per sample. If both the approximate phase slopes and the signal-to-noise ratios on the baselines are known, Figure 3.5 shows the optimum number of samples over which the single baselines should be integrated before forming the triple product.

3.4.4 Conclusions

Closure phase is sensitive to the signal-to-noise ratios of the three baselines in a loop, and if all are less than one the closure phase becomes a very noisy quantity. Under such conditions the individual baseline visibilities should be integrated for as long as possible, combined by the ‘triple product’ method and then averaged further. The argument of the result will then be the best estimate of the closure phase. Working with the phases alone, suitably assigned to unit vectors, is not recommended as this can increase the final phase noise by as much as a factor of two when the signals are weak ($\gamma_s < 1$).

Chapter 4

The megaparsec component of 3C48

4.1 Introduction

A number of observations made at low frequencies of the quasar 3C48 have suggested that it possesses structure on the scale of a few arcminutes that has low surface brightness and is not detected at higher frequencies. Two of these observations were made in Cambridge at 81.5 MHz. The first comes from the IPS survey of Readhead and Hewish [68] which, with the revised model of the IPM due to Readhead, Kemp and Hewish [69], suggests that a large fraction (perhaps up to 40%) of the total flux from 3C48 cannot be attributed to a component more compact than 2 or 3 arcseconds. The later IPS observations of Purvis, Tappin, Rees, Hewish and Duffett-Smith [65] suggest a 15% anomaly between its total and compact ($< 2''$) fluxes, though this measurement is thought to be less accurate. Further observations were made with a long baseline interferometer by Hartas, Rees Scott and Duffett-Smith [41], and shows a dramatic 10% drop in the visibility of 3C48 between baselines of 130λ and 1350λ . This clearly points to a component of low surface brightness and with an extent of more than 2.5 arcminutes. Observations by Vasiljev *et al.* [100] at 85 MHz using a radio linked interferometer also seemed to detect the extended structure. These were of a lower quality than those of Hartas *et al.* but showed a similar downward trend in the visibility on short spacings. The two interferometric observations are compared in Figure 4.1. Assuming that the redshift of the quasar is cosmological, this extended structure would have a linear extent of about a megaparsec (taking $z = 0.367$, $H_0 = 60$ km/sec/Mpc and $\Omega_0 = 1$, one arcminute corresponds to about 640 kpc).

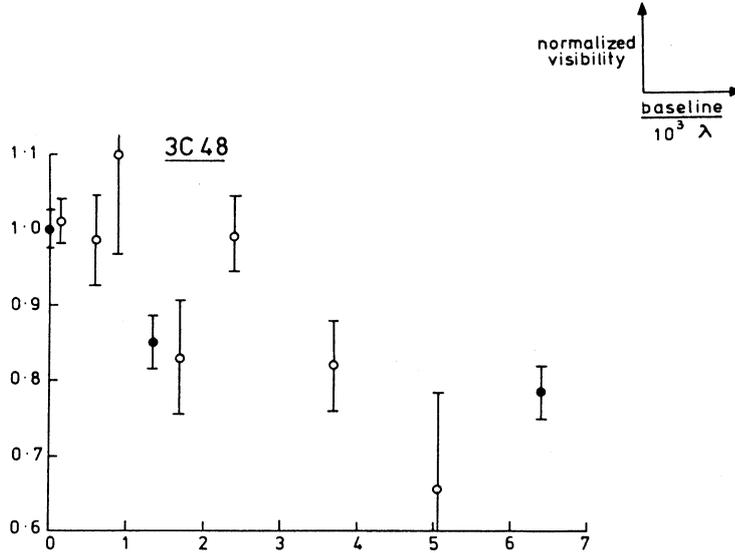


Figure 4.1: Comparison of visibilities measured by Vasiljev *et al.* (empty circles) and Hartas *et al.* (filled circles) for 3C48.

There are only two other examples of quasars known to be associated with giant radio sources, 4C 34.47 ($z = 0.206$) [6] and 4C 74.26 ($z = 0.104$) [75], each of which have a projected linear dimension of more than a megaparsec. Both can be classified as giant FR II sources, having very well defined double-lobed radio structures with associated hotspots; in fact their morphologies are very similar (see Figure 4.2). In contrast, no multiple structures or hotspots have been associated with 3C48 on the scale of an arcminute. The most striking features of 3C48 are its unusually steep core spectrum ($\alpha = 0.7$) and its powerful far-IR flux as detected by IRAS. The MERLIN map of the core at 327 MHz [104] reveals a badly distorted steep spectrum jet about a kiloparsec in extent (Figure 4.3). It is thought that the distortion has been caused by a strong interaction between the beam and the interstellar matter within the host galaxy, and the steep spectrum from premature ageing of the ejecta in the galactic magnetic field. Infrared observations [58] have revealed strong ($L \sim 10^{12}L_{\odot}$) emissions that are most probably not an extension of the radio spectrum of the core but rather thermal re-radiation from heated dust in the surrounding galaxy. This supports the notion that the jet from the core is close to the galactic plane and has been ‘muffled’ by its interaction with the galactic medium. The impact of the jet could then be responsible for the induction of rapid

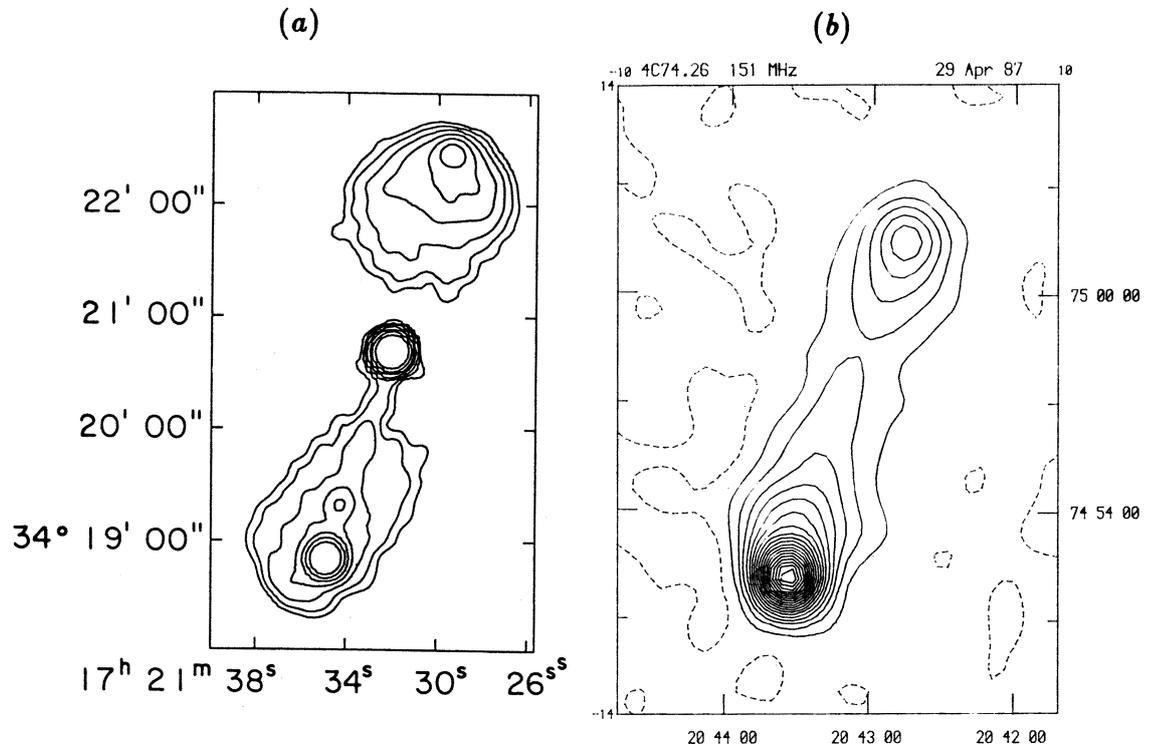


Figure 4.2: (a) 4C 34.47 mapped with the VLA at 1.4 GHz ($11''$ beam); (b) 4C 74.26 mapped with the CLFST at 151 MHz.

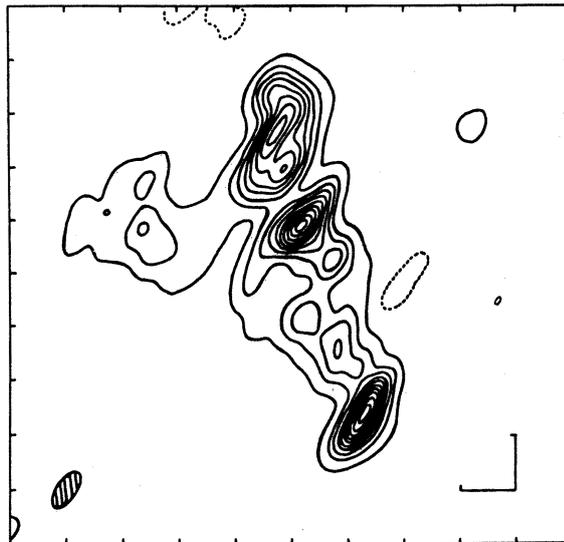


Figure 4.3: MERLIN map of 3C48 at 327 MHz. The scale indicates $0.05''$.

star formation, and hence the anomalous IR luminosity. Nebulosity from the host galaxy, an optically bright spiral [10], has been traced out to 160 kpc ($\sim 15''$) [35], but no structure on larger scales than this has been reported, with the exception of the three radio observations described above. Indeed, it would be difficult to explain such structure as a consequence of a *blocked* jet.

It might be expected that any large scale extended structure would have a steep spectrum due to synchrotron losses, and there are a number of possible explanations for its presence that do not rely directly on beaming arguments. Firstly, although the large redshift has prevented any definite observation of a cluster of galaxies around the quasar, it is quite possible that extended emission could be associated with the smoothed remains of adjacent radiogalaxies [18]. Secondly, the flux could be due to ‘quiet’ galactic emission, attributable to energetic ejecta from an ensemble of supernova remnants in the field. However, the ‘halo’ reported by Hartas *et al.* contributed $\sim 10\%$ of the total flux density of the quasar at 81.5 MHz. This is about 1000 times the extended flux that would be expected from arguments based on the observed correlation between far infrared and radio emissions from quiet galaxies [31], and it is unlikely that this mechanism could generate emissions on the observed scale. Thirdly, the structure could be a low frequency remnant of an earlier active phase in the evolution of 3C48 which stopped before, or was interrupted by, the formation of the host galaxy.

With all these mechanisms in mind, it is equally important to consider the possibility that the apparent structure is totally spurious, and generated by systematic errors in the observations. The most likely cause of systematics in results derived directly from visibility measurements is confusion from other sources in the primary beam of the telescope. This point will be considered in greater detail later.

4.2 Recent observations

I made further investigations (with Duffett-Smith) of the reported arcminute structure in August 1986, using VLBI equipment and the phase calibration scheme described in Section 3.3. These involved the use of the Wardle antenna of the Jodrell Bank MERLIN network, which was converted to operate at 81.5 MHz by fitting a crossed dipole feed in the focus box. Left and right circularly polarised signals from the antenna were fed to a motorcaravan, which was permanently stationed by the antenna and contained the rest of the receiving and recording equipment. Cassettes were collected regularly by local staff and posted to Cambridge, where they were correlated with the tapes from the other two sites. In this way it was hoped that the

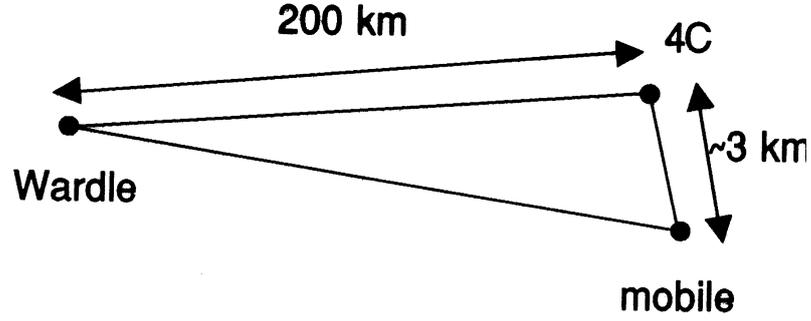


Figure 4.4: The ‘thin triangle’ baseline geometry used for the observations. The distant (Wardle) station was for phase calibration only.

name	l (λ)	l (km)	hr. ang (rad)	dec. (rad)
5 km	261	0.960	-2.307167	0.480460
Rose Cottage	352	1.296	-1.577755	0.0048290
Haslingfield	437	1.609	-0.897964	-0.451227
Comberton	946	3.484	-3.031907	0.657149
Wimpole Hall	1603	5.898	1.217509	-0.258550
Fowlmere	2461	9.059	-0.409832	-0.6189264

Table 4.1: The short baselines used for VLBI of 3C48.

phase on the shorter baselines could be determined and used to map any structure that was detected.

The geometry of the closure triangles is shown in Figure 4.4. Six short baselines were used in all, ranging in length from 0.96km to 9.06km, bracketing the relevant measurements made by Hartas *et al.*

If 30% of the flux from 3C48 is extended, the *maximum* astronomical phase deviation expected on the shorter baselines is about a third of a radian. With the antennas in the configuration described above, this phase is equal to the closure phase, so the closure phase itself must be determined to ~ 0.1 radians if it is to be useful. Therefore the signal-to-noise ratio in the triple product (γ_*) needs to be about $10/\sqrt{2} \simeq 7$. The sensitivity of the three baseline measurements can be readily calculated from

$$\gamma_{ij} = \frac{S\eta_q\eta_f}{k} \sqrt{\frac{A_i A_j \Delta\nu\tau}{2T_i T_j}}. \quad (4.1)$$

A further term must be included to take into account the fact that the 4C station only observed linear (E-W) polarisation, whilst the other two received and recorded

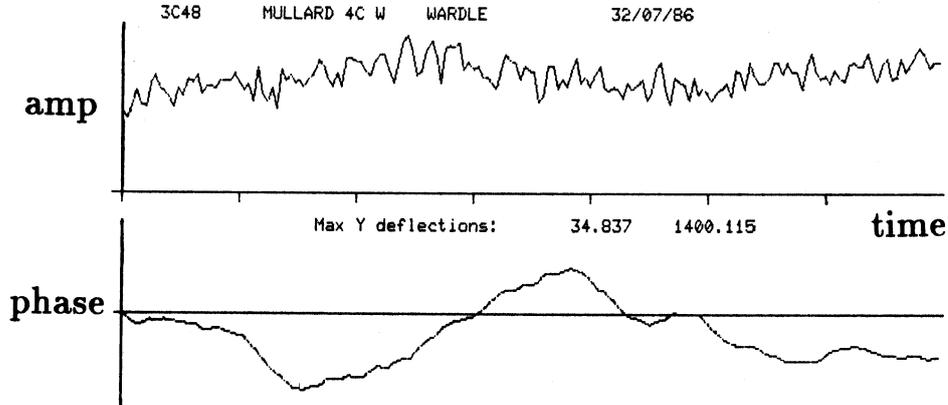


Figure 4.5: A 3 hour visibility track from the 4C-Wardle baseline. Note that although the maximum phase deflection is 1400 degrees, the changes are smooth and there is no short-term decorrelation.

circular polarisation. The effective areas of the three antennas were estimated to be: 4C, $A \simeq 2000\text{m}^2$; the mobile array of 16 Yagis, $A \simeq 100\text{m}^2$ and the Wardle dish, $A \simeq 500\text{m}^2$. Taking $S = 83\text{Jy}$, $\eta_q\eta_f = 0.5$, $\Delta\nu = 125\text{kHz}$, $\tau = 60\text{s}$ and the system temperatures as 2000 K, the signal-to-noise ratio in the triple product was about 4 (see Equation 3.36). Integration of the triple product for about 5 min would therefore have given a closure phase of sufficient accuracy for mapping.

4.3 Results

After some considerable effort, both in the servicing of equipment and in correlation, a total of 15 useful runs were made within the time that the telescope was available, including examples from all six positions of the mobile station. However, the results were singularly disappointing. First of all, the signal-to-noise ratios on the Wardle-4C and mobile-4C baselines were almost the same, indicating that the Wardle antenna had an effective area similar to that of eight phased Yagis! The phase record shows that this cannot be attributed to decorrelation (Figure 4.5), either from the ionosphere or from the equipment. It was more likely to be due to focussing problems, or the disturbing amount of interference generated by a nearby factory which would have increased the effective antenna temperature. Either way it resulted in an extremely weak Wardle-mobile baseline which, even after global fringe-fitting [79], gave a signal-to-noise ratio on that baseline of about one in a

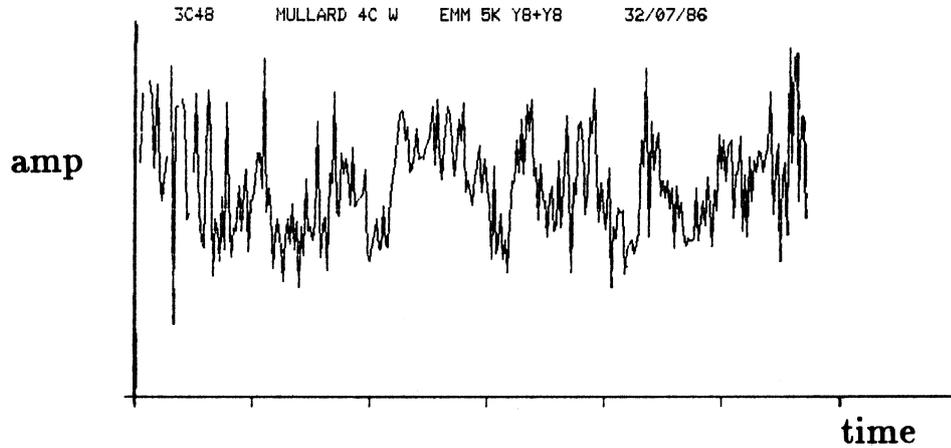


Figure 4.6: A 6 hour visibility track from the shortest baseline. The end of the run was dominated by Cas passing through the beam.

minute. Although closure phase could have been obtained using the methods described in Chapter 3 the accuracy would have been insufficient, and this largely precluded the generation of a map. All was not lost however, as the short spacings yielded visibilities of approximately the expected magnitude, and the amplitude records could be searched for evidence of extended structure.

The visibility tracks from the 4C-mobile baselines clearly showed a further problem (Figure 4.6). The shortest baseline, for example, displayed large fluctuations in its amplitude on a timescale of about 20 minutes. This corresponded to features a few degrees away from the phase centre which could not be associated with 3C48. Instead, the fluctuations must have been due to other sources in the primary beam confusing the record. Such effects are usually not a fundamental problem, as the contributions from these sources can be isolated by making a big map. What is important is that the confusion might also explain the anomalous short-spacing fluxes measured by Hartas *et al.* who used an instrument with a similarly large beam size but simply took a spot height in the aperture plane. On the two longest baselines, bandwidth smearing and differential fringe rate effects would have been sufficient to remove most of this confusion, both in my observations and those made by Hartas *et al.*

4.4 Confusion models

It appears, therefore, that the effects of confusion might account for the apparent large scale structure seen by Hartas *et al.* Of the three pieces of evidence cited, that of Hartas is by far the most convincing, and the case for such structure is severely weakened without it. It is difficult to pursue this topic further without a better understanding of both the instruments used in the observations and the radio source distribution in the vicinity of 3C48. To this end detailed theoretical models were constructed of both the 4C antenna used in the present observations and the 3.6 hectare array used by Hartas. These were applied to a model of the radio sky containing 103 point sources, including 3C48, in and around the primary beams of the telescopes (Appendix F, Tables F.1 and F.2). The list includes Cas A (item 102) and Cygnus A (item 103) as their massive low frequency fluxes can be expected to make a contribution, even though they lie well beyond the central lobe of the primary beam. The 81.5 MHz fluxes of the other sources were largely obtained from the compilation made by Purvis [64] and compared with results from the most recent IPS survey [65]. Their positions were taken from preliminary results of the 6C and 7C surveys and from the composite catalogue of Kühr *et al.* [50], and were precessed to appropriate epochs for the simulations. Some of these sources may be partially resolved on the longer baselines (notably Cas and Cygnus), but for the shorter baselines, with resolutions of many arcminutes, they can all be regarded as points.

The contribution of a particular source to the total visibility depends on a number of factors. In the following simulations it is assumed that the primary beam of the interferometer is defined by the beam of the larger antenna. In the case of the present observations this is the beam of the 4C, and for the Hartas observations it was that of the 3.6 ha array. This is reasonable, as the beams of the two mobile stations were broad, and could be taken as constant over the region of sky considered. The response of the interferometer to a point source off axis is modified by two further terms. Firstly, the difference in fringe rate between a source and the phase centre of the sky results in a smearing of the integrated visibility. Secondly, the finite bandwidth of the instrument reduces its response to sources which have a geometrical phase (in radians, with respect to the phase centre again) comparable with the fractional receiver bandwidth – this is just chromatic aberration. The response of the interferometer to the source list can be calculated from

$$V(H_0) \propto \sum_{j=1}^N P(H_0, \delta_0, H_j, \delta_j) S_j \quad (4.2)$$

$$\begin{aligned} & \cdot e^{i(\phi_0 - \phi_j)} \operatorname{sinc} \left(\frac{\Delta\omega}{2\omega} (\phi_0 - \phi_j) \right) \\ & \cdot e^{i(\dot{\phi}_0 - \dot{\phi}_j)\tau/2} \operatorname{sinc} \left(\frac{\tau}{2} (\dot{\phi}_0 - \dot{\phi}_j) \right), \end{aligned}$$

where $P(H_0, \delta_0, H_j, \delta_j)$ is the beam response to a source at hour angle H_j and declination δ_j when the antennas are phased to (H_0, δ_0) , ϕ_0 and ϕ_j are the geometric phases of the phase centre and the j th source, S_j is the flux of the source, $(\Delta\omega)/\omega$ is the fractional RF bandwidth and τ is the integration time.

4.4.1 Simulating the recent data

In order to model the observations presented in Section 4.2 by Equation 4.3 a form for P is needed that accurately represents the beam of the 4C. This is derived in Appendix E and shown graphically in Figure 4.7. As can be seen, the 4C has a banana shaped beam when phased away from transit, and the curvature of the banana changes as the observation progresses. This is a very great problem when using the 4C as an aperture synthesis instrument because it means that confusing sources that are further from the beam centre than the minor axis of the banana ($\sim 1^\circ$) will fade in and out. Under these conditions the visibility plane no longer faithfully represents the transform of the sky brightness and the effects of confusion from such sources are very difficult to remove.

The confusion simulation for the shortest baseline (with the mobile station near the 5km building at Lords Bridge) is shown on Figure 4.8, together with the corresponding observational data, three hours either side of transit. The large deflection at the end of the run is due to Cas passing through the beam. There is a striking similarity between the two records, showing that most, if not all of the crinkles in the observed visibility tracks are due to confusion. It is not surprising that the two do not fit exactly. The beam model is simple, and assumes that the trough of the 4C acts as a perfect rectangular aperture. The true beam will probably have reduced side lobes and a broader central peak. Furthermore, the low frequency flux densities of a number of the brighter sources are only known to a low precision. Nevertheless the *characters* of the two records are the same, showing that the *degree* of confusion has been accurately modelled. As the baseline length increases, the effects of confusion gradually wash out. Figure 4.9 shows the same comparison for the longest baseline. Again the characters of the two records are the same, indicating that the model is sufficient to predict the degree of confusion on a given baseline.

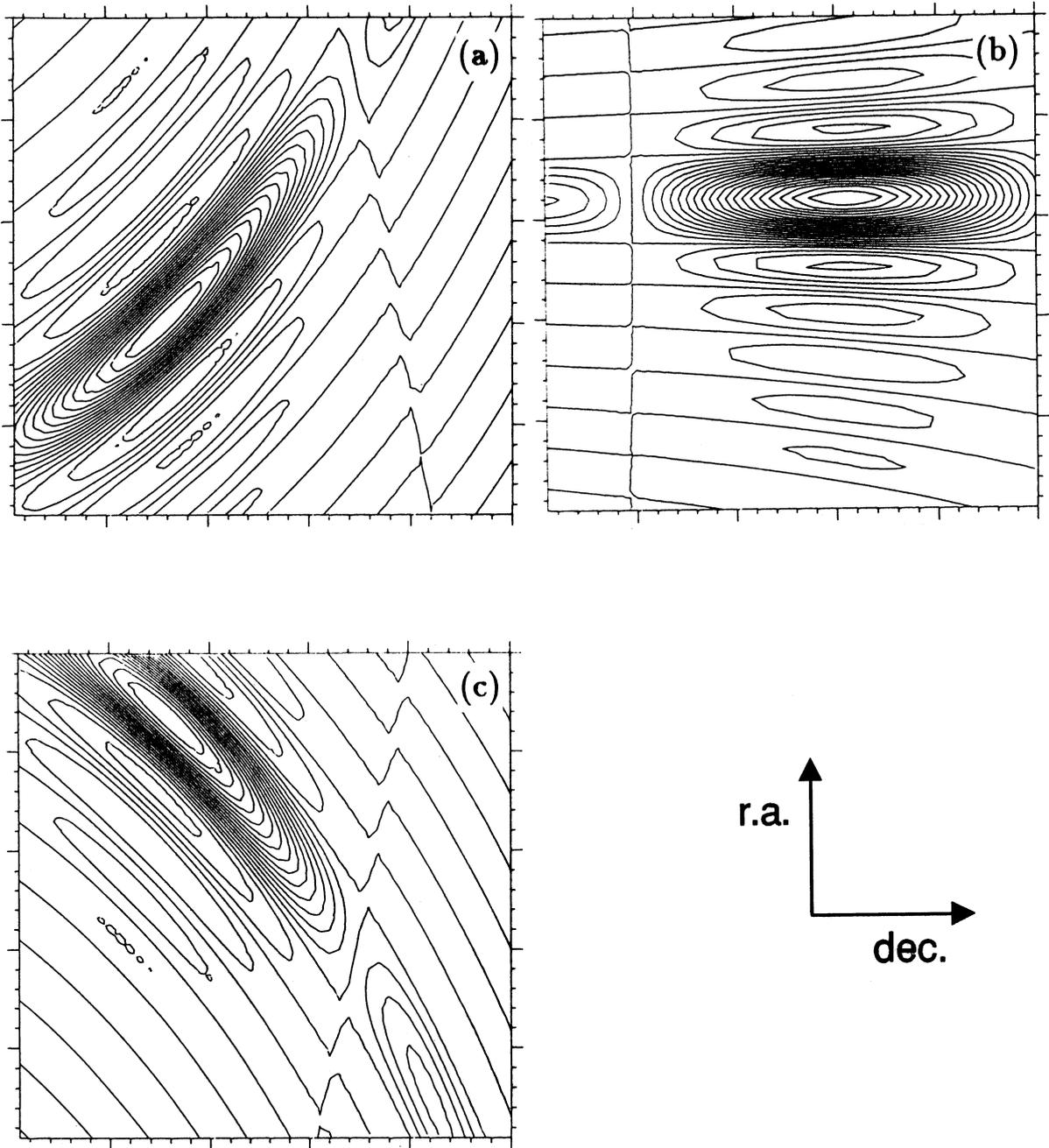


Figure 4.7: The beam of the 4C array (a) -3 hr, (b) 0 hr and (c) +3 hr from transit. The field is $50''$ in R.A. by 20° in dec.

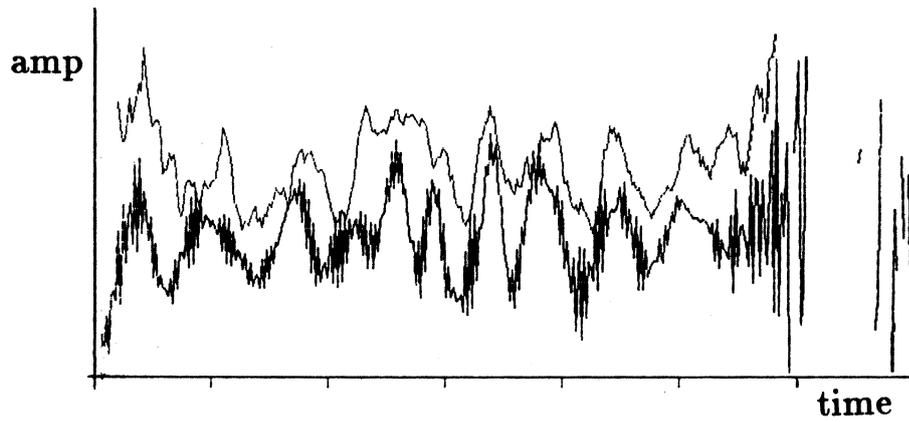


Figure 4.8: Comparison of measured (upper trace) and simulated (lower trace) visibility tracks for the shortest baseline over 6 hrs. The simulation has been displaced vertically to make the correlations clearer.

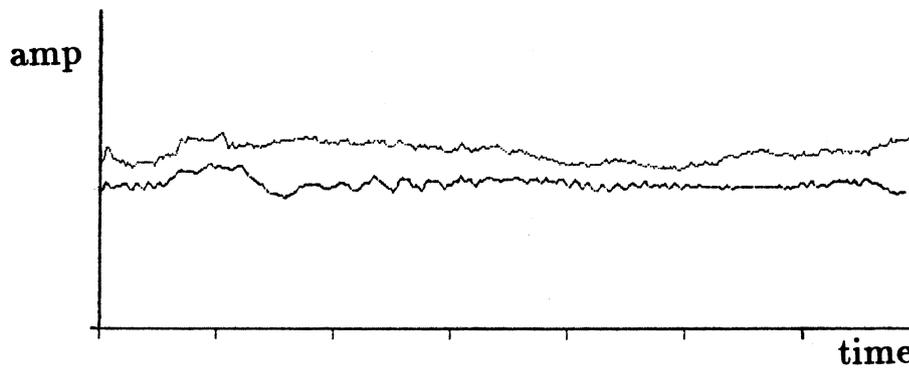


Figure 4.9: As Figure 4.8 but for the longest baseline (Fowlmere). Again, the lower (simulation) trace has been displaced slightly for clarity.

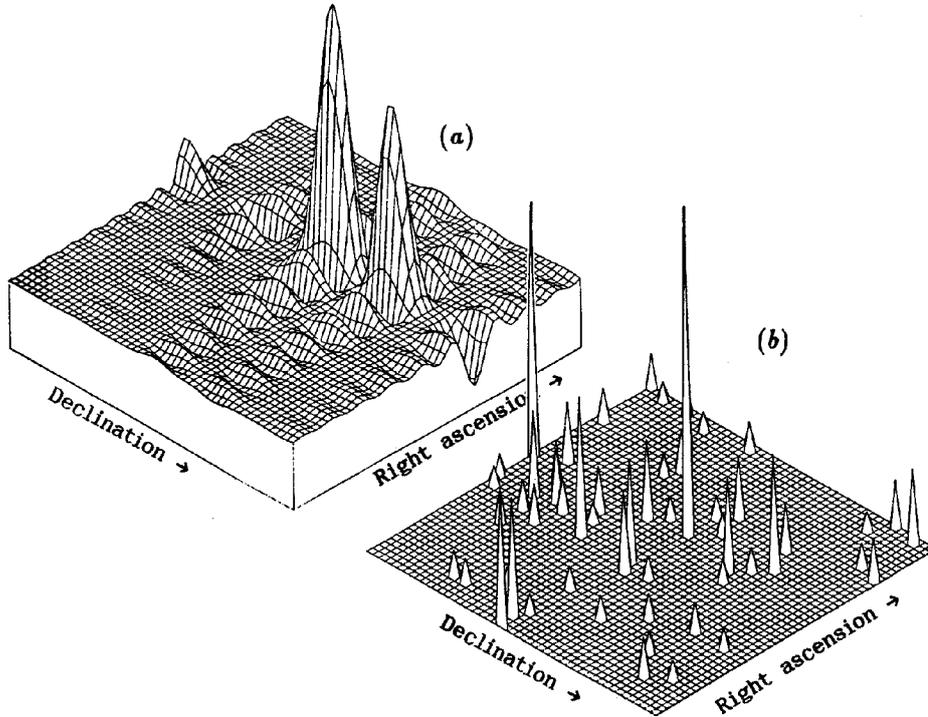


Figure 4.10: Models of the telescope beam (a) and surrounding sky (b) used to simulate the observations of Hartas *et al.* . The field size is 25° in dec by 50^m in R.A. The large peak near the centre of (b) corresponds to 3C48.

4.4.2 Simulating the data of Hartas *et al.*

It is clear from the previous section that the effects of confusion dominated my short-spacing measurements. If this were also the case for the measurements made by Hartas *et al.* , their report of extended structure could easily be spurious. Again, it is necessary to model the beam pattern of the antenna, which in this case is the 3.6-ha array. This is a transit instrument, designed for IPS work, and is electrically divided into two arrays, each consisting of 128 dipoles in 16 rows. Each row is fed to a Butler matrix which generates a number of static beams on the sky. Hartas made his observations by choosing the beam closest to the transit position of 3C48 and then phasing the two halves of the array together. The composite beam that results has a rather curious shape (Figure 4.10), including two large maxima at and near the centre. This theoretical beamshape is also derived in Appendix E. Visibilities were measured every second for 100 seconds centred about transit and then averaged.

The simulations mimic this procedure closely, and evaluate the theoretical system response to the list of point sources at closely spaced intervals, $\pm 300\text{m}$ from the sam-

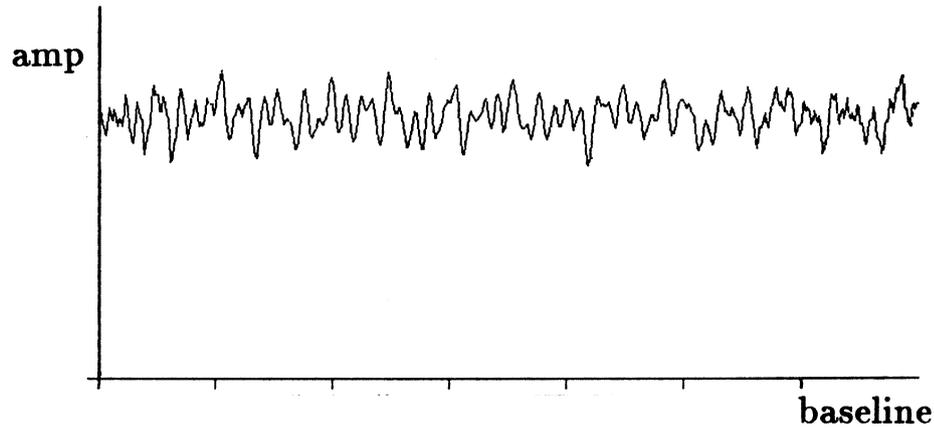
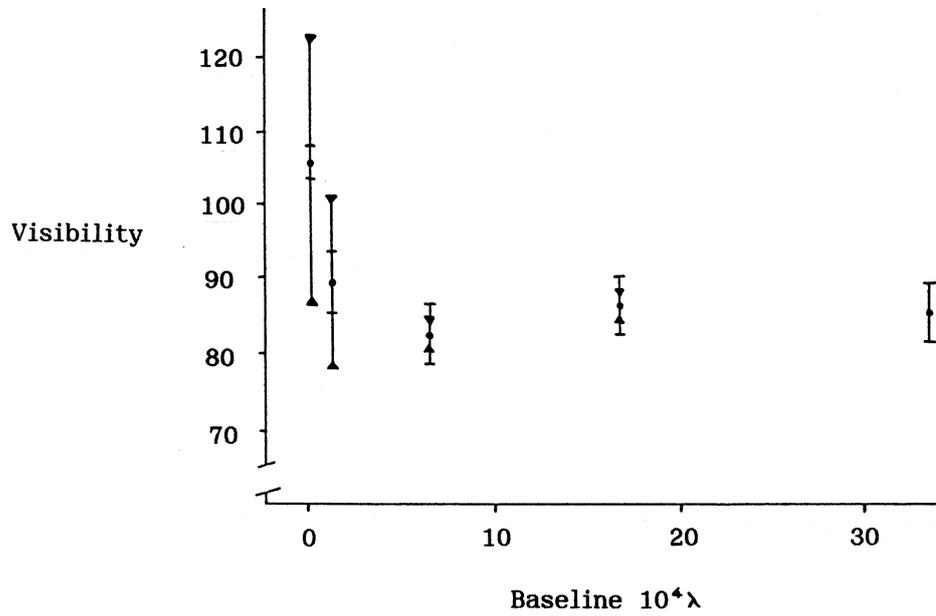


Figure 4.11: A 1000λ E-W cut through the simulated transform plane. All features are due to confusion.

ple points along vectors corresponding to the five shortest baselines used in the Hartas experiment. These were the baselines to Lords Bridge (130λ), Kingston (1350λ), Weston Colville (6360λ), Stowmarket (16830λ), and Christ's Hospital (34970λ). The aim is to get a feel for the character of the transform space in the vicinity of the sample points, and thus assign a figure for the expected component of confusion in the data. At short spacings, the expected amplitude pattern is just the transformed autocorrelation of the sky-beam product. Figure 4.11 shows an E-W cut through the pattern out to a distance of about 1000λ , and demonstrates the magnitude of the confusion problem on the shorter baselines. The fractional range in amplitudes in the vicinity of a sample point corresponds to the 'confusion bounds' at the point and defines the limits of uncertainty in the measurement. These are summarised in Table 4.2. Figure 4.12 shows the original Hartas data, but now with the computed confusion bounds superimposed. The 1σ error bars for the measurement (i.e. due to the noise) are enclosed between the horizontal marks. The limits of confusion are denoted by triangles. The measurement at Christ's Hospital is omitted as the confusion there was negligible. As can be seen, confusion considerably increases the uncertainty associated with the short spacing measurements on which the claims of extended structure were based. Although, due to the sparseness of data, it would be inappropriate to quantify a revised confidence level for structure, it is reasonable to say that the Hartas observation can no longer be regarded as sufficient grounds for its existence.

name	l (km)	l (λ)	Confusion ($\pm\%$)
Lords Bridge	0.479	130	16
Kingston	4.968	1350	13
Weston Colville	23.40	6360	2.7
Stowmarket	61.93	16830	2.4
Christ's Hospital	127.84	34739	<0.06

Table 4.2: Confusion bounds as simulated for the observations of Hartas *et al.*Figure 4.12: The data of Hartas *et al.* (error bars shown) with calculated limits of confusion between the triangles. There is negligible confusion in the most distant measurement, made at Christ's Hospital.

4.5 Discussion

Measurements made with the new tracking 81.5 MHz interferometer, together with results from simulations of the same observations, clearly identify confusion from adjacent sources as contributing a significant fraction of the measured flux of 3C48 on short baselines. By applying similar simulation techniques to measurements made by Hartas *et al.* in 1982 it becomes evident that the anomalous short-spacing increase in flux reported could be accounted for by a model in which 3C48 remains unresolved.

The other evidence for extended structure is largely circumstantial. Vasiljev *et al.* reported a similar increase in amplitude at short spacings, but their primary beam was also large, so the effects of confusion may also have been significant. The IPS evidence is not easy to interpret, and is unable to provide a scale size for structure extending more than about an arcsecond. However, it does suggest that between 10% and 40% of the integrated flux from 3C48 at 81.5 MHz is due to structure on scales greater than one or two arcseconds. This is quite consistent with the arcminute halo reported by Hartas, but it is just as consistent with features on a much smaller scale. Hartas also reports the detection of an extended component to 3C48, contributing about 10% of the total flux, but with a scale size of only 3.5 arcseconds. The MERLIN map of 3C48 at 151 MHz (Figure 4.13) also reveals structure on this scale. The features are close to the resolution of the map, and they are easier to characterise directly from the longer (~ 130 km) baseline visibilities (e.g. Figure 4.14). The relative calibration of the baseline amplitudes is poor, but the more distant tracks are modulated in amplitude at about the 5% level. This means that at least 5% of the the total flux must be resolved at 3 arcsec. Unfortunately, the MERLIN data does not extend below $\sim 2000\lambda$, so does not reveal any arcminute features. This arcsecond structure, being within the host galaxy, is quite consistent with our present understanding of the quasar.

It would be wrong to conclude from this discussion that 3C48 does *not* possess a megaparsec component, but from the evidence available there seems little reason to suppose it does.

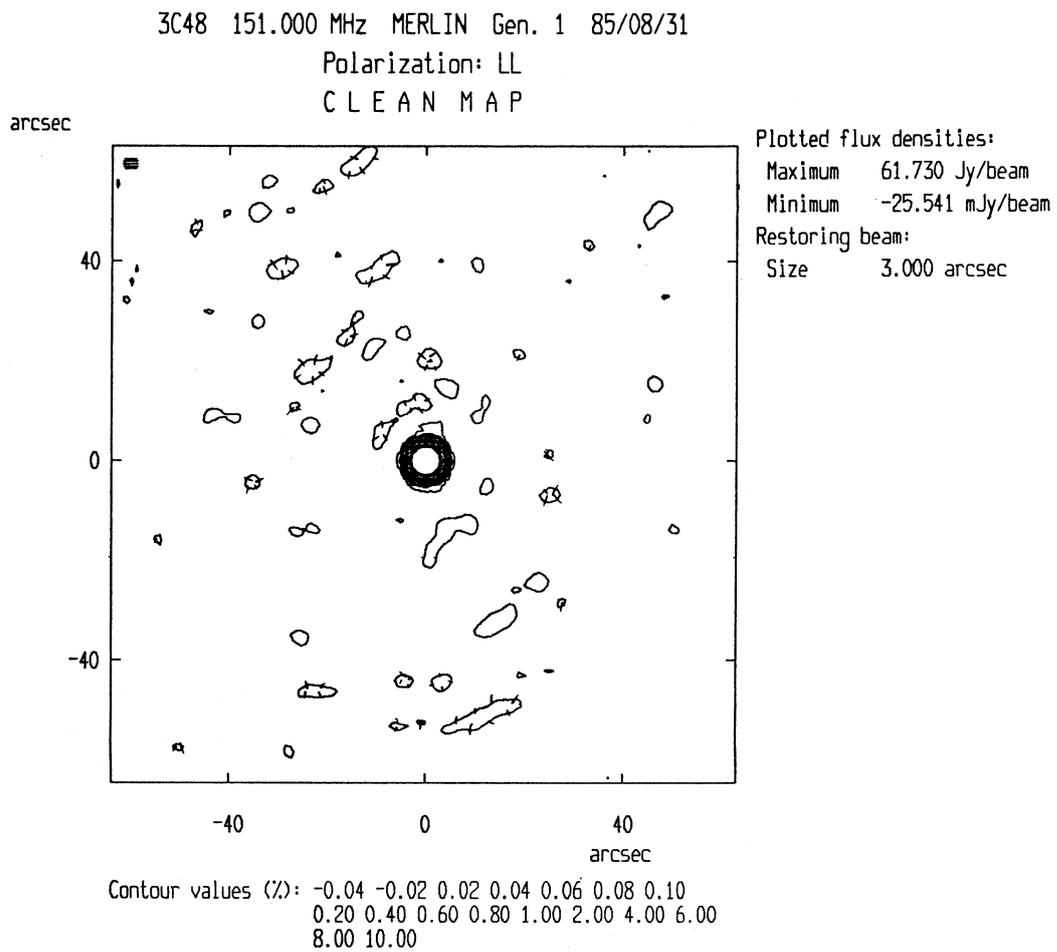


Figure 4.13: MERLIN map of 3C48 at 151 MHz. Spacings were from 6 km to 130 km.

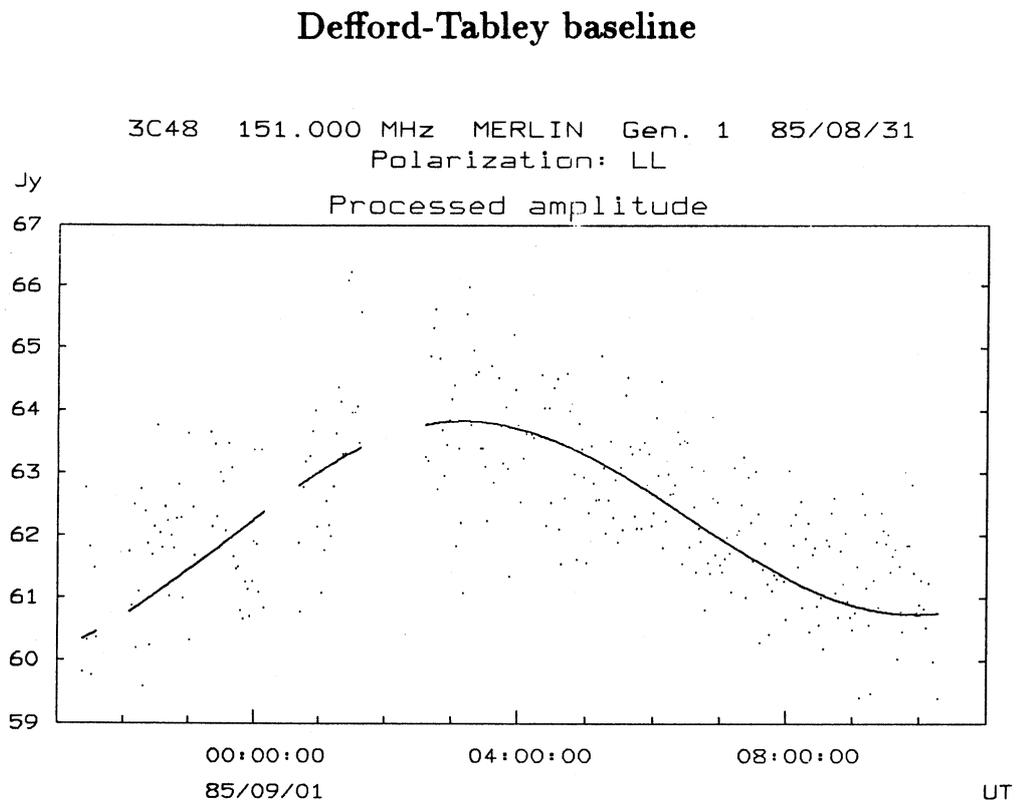


Figure 4.14: Amplitude one one of the longer baselines used to generate Figure 4.13 (note the scale).

Chapter 5

External phase calibration

The techniques of phase calibration considered in Chapter 4 can be thought of as *internal*, since the phase information is derived solely from the uncalibrated visibility measurements of the source under observation. However, calibration can also be achieved *externally* by simultaneously observing one or more calibrating sources and using these measurements, together with prior knowledge about the calibrators, to deduce the complex gain factors associated with each baseline. If this technique is to be useful for phase calibration, both the positions of the calibrators and the length and orientation of the baseline vectors need to be known with sufficient precision to compensate for the calibrator's fringe rate. Furthermore the calibrators must be bright, so that little aperture coverage is lost during calibration, and unresolved, so that phase corrections can be deduced accurately. Usually very few of these conditions can be met in VLBI, so calibration is almost universally achieved by internal means. However, the number of antennas is an important factor in the amount of phase information that can be reclaimed by these internal (closure) methods. An array of n antennas, in a non-redundant configuration, can produce $n(n-1)/2$ complex visibilities. If the antenna phases are unknown, $(n-1)(n-2)/2$ independent closure phases can be extracted to be used with the $n(n-1)/2$ amplitudes in making the map. When n is large, $(n-2) \simeq n$ and little information is lost other than the source position. On the other hand, if the array consists of only three antennas as is the case here, $2/3$ of the phase data is lost, so that the reliability of a reconstructed image may be severely compromised. Under these circumstances, external calibration schemes become attractive once more.

A particular disadvantage of external calibration methods is that they require assumptions to be made about the calibrator and the atmosphere. The calibrator must have a well known fringe rate and be within a few degrees of the source being

mapped, so that the atmosphere has similar effects on both. Furthermore, if the calibrator and the source are not within the same primary beam, the telescopes have to be driven between them during the observation and within the coherence time of the interferometer. It takes about one man-hour of manual labour to change the declination of the the 4C, so this is clearly impractical with the present LBI system.

Nonetheless, even if full external phase calibration cannot be achieved, many of the problems associated with the VLBI interferometer may be overcome if the equipment itself is calibrated. The interferometer then behaves almost as if ‘connected’, i.e. having all the desirable properties of one in which the individual antennas are linked by cables or radio. In particular the coherence time, and hence the sensitivity, increases and many of the problems associated with fringe-finding vanish. Sensitivity is also increased as another consequence of this phase stability. In an interferometer in which the frequency standards, a and b , have a relative drift, the phase of the k th correlated frequency channel can be written as

$$\phi_k = \phi_{\text{ast}} + \frac{\omega_0}{\Omega_0}(\Omega_a - \Omega_b)t + \frac{k\Delta\omega}{\Omega_0}(\Omega_a - \Omega_b)t + \epsilon_k, \quad (5.1)$$

where ω_0 is the nominal lower bound of the passband, Ω_0 is the nominal frequency of the standards, Ω_a, Ω_b are their actual frequencies, $\Delta\omega$ is the bandwidth of each frequency channel within the passband, ϕ_{ast} is the ‘true’ astronomical phase and ϵ_k is a phase offset which is constant, but can take any value. Although the constant offsets can be accounted for in software, the unknown relative phase drifts between the channels prevent the signals from each channel being added coherently. Knowledge of the relative phases of the two frequency standards enables these drifts to be calculated, so allowing coherent addition. However, the phase still suffers from atmospheric effects implicit within ϕ_{ast} which, at low frequency, are dominated by the ionosphere.

To achieve these improvements, a new method of phase calibration has been developed. In addition to receiving celestial radio signals, both elements of the interferometer receive signals from a common ground-based VHF transmitter using receivers at a different frequency from, but phase-locked to, the astronomical receivers. The ground-based transmitter acts as a phase reference; its signals are used at the processing stage to correct the phase drifts in the astronomical channels to within a few degrees over an unlimited period. This technique has a number of advantages over conventional point source calibration. Firstly, the signals are very strong, so that only small antennas and short integration times are required. Secondly, the transmitters are fixed relative to the baseline and so have zero fringe rate. Thus the phase calibrations do not depend (up to a constant offset) on accurate knowledge of the positions of either the transmitter or receivers.

5.1 Phase calibration using VHF transmitters

5.1.1 The choice of reference transmitter

The selection of a reference transmitter for use as a phase calibrator is straightforward. The original centre frequency of the interferometer, at 81.5 MHz, is so close to Band II (88-108 MHz) that the obvious candidate is one of the BBC (or ILR) VHF FM stations. There are several other advantages to this choice:

- There are many powerful transmitters to be found in this band. Figure 5.1 shows just those for the BBC. This means that only small antennas and short integration periods are needed.
- The transmitters are fixed to the ground (see above).
- The frequency is high enough for sky wave propagation modes to be negligible. Although reflections from buildings and the ground will be present, the propagation path is likely to be *stable* in the sense that it will not change throughout the run. The effects of transient reflections from moving objects such as cars and aeroplanes can be minimised by choosing integration times significantly longer than their times of action.
- Stereo encoded composite FM has a maximum frequency deviation of about ± 75 kHz, sitting comfortably within the 125 kHz receiver bandwidth available, and little power is lost during periods of high modulation.
- Many VHF stations transmit 24 hours a day, so stable interferometry does not require recourse to the *Radio Times*.

Interference from adjacent transmitters can be neglected both because transmitter frequencies are specifically allocated to minimise it in domestic use, and because any interfering transmitter must have a similar geometric delay, τ_g , to the main transmitter for its effects to show up in the correlated output. The two transmitting stations chosen as references were the local Radio Cambridgeshire relay on Madingley Hill (1kW at 96 MHz) and the more powerful Tacolneston Radio 1 & 2 transmitter (250 kW at 89.7 MHz) near Norwich. The difference in power, position and programme content between these two stations enabled systematic effects depending on these variables to be identified. Due to constraints imposed by the frequency synthesisers, the actual centre frequencies used in the receivers were 96.000000 and 89.703125 MHz.

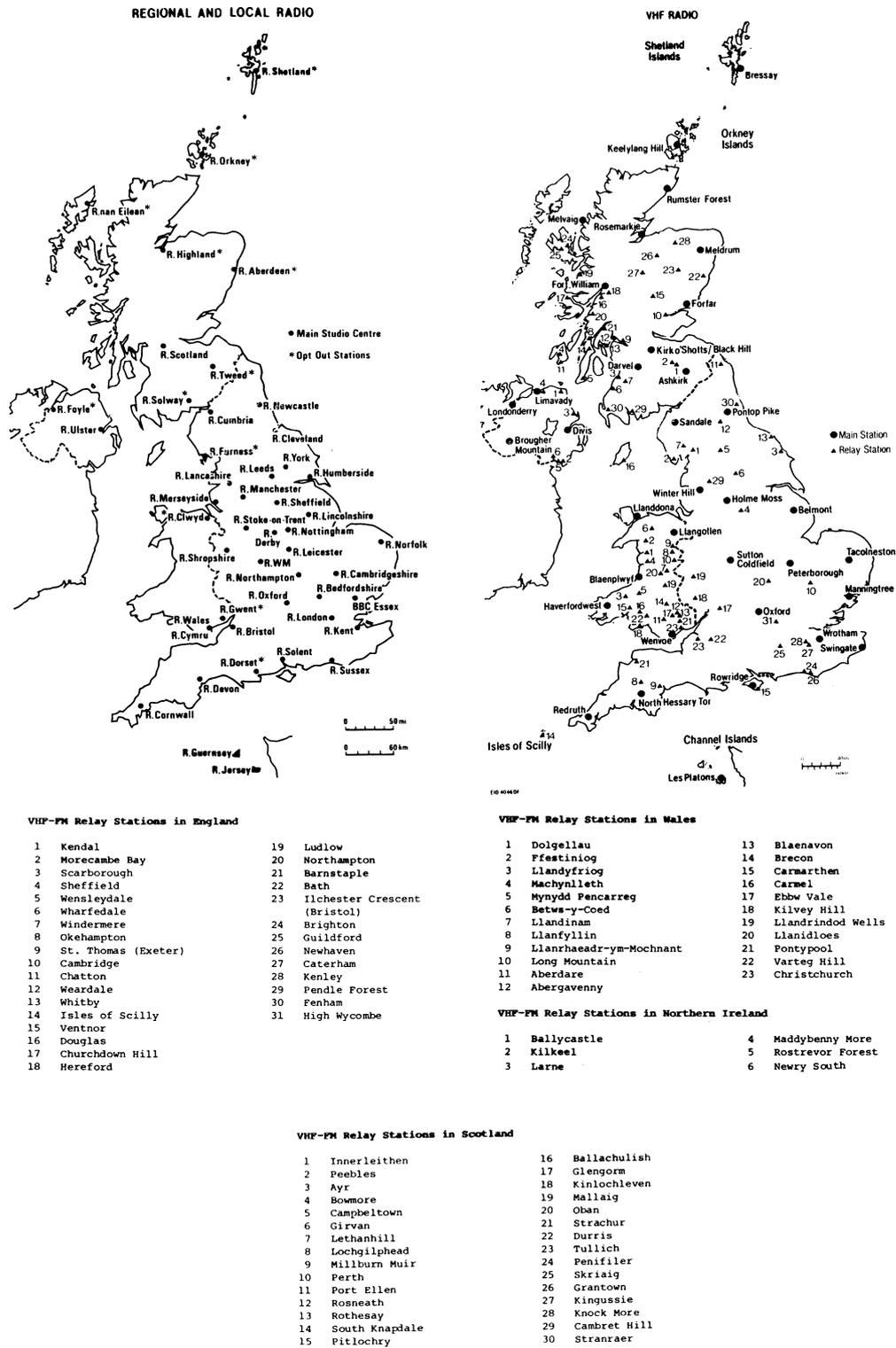


Figure 5.1: Transmitters of the BBC VHF local and national networks.

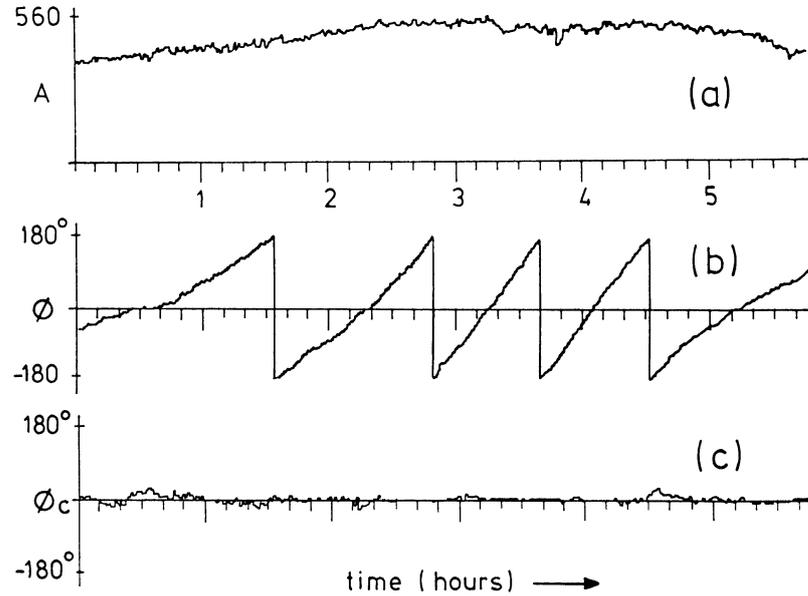


Figure 5.2: The amplitude (a) and phase (b) measured by the uncalibrated interferometer observing Cas A on a 1.321 km E-W baseline. (c) shows the phase after calibration has been applied.

5.1.2 Implementation and results

One of the most attractive features of the calibration scheme is that it requires only minor modifications to the existing equipment. As described in Chapters 1 and 2, the 1MHz passband of the interferometer is divided into eight contiguous frequency channels which are correlated independently before being combined. This basic configuration is retained, but two of the eight channels are re-tuned to the above frequencies, and are fed by independent folded dipole, rigidly mounted close to the recording equipment. Otherwise the two channels are left largely unaltered so their local oscillators are still phase-locked to the frequency standard and their data recorded on tape along with the astronomical data. The calibration itself is done retrospectively, at the correlation stage. It proceeds as follows:

1. The tapes are passed through the correlator, with the geometrical delay necessary to see a signal from the calibrator within the correlator window of $64\mu\text{s}$.
2. The 'tracked phase' of this signal is recorded on disc for the six hours of the run. The tracked phase is the *accumulated* phase, $\phi_t(t)$, up to time t , accounting for factors of $2\pi n$, i.e.

$$\phi_t(t) = \int_0^t (\omega_2(t') - \omega_1(t')) dt' + \text{const.}, \quad (5.2)$$

where ω_1 and ω_2 are the frequencies of the local oscillators used to mix the calibration signal at the two ends of the baseline. (Section 5.2.1 goes into phase tracking in more detail.)

3. The corresponding phase drift to be expected on the astronomical channels are then deduced. This is simply a matter of multiplying the tracked phase determined in the previous step by the ratio of the two frequencies ($\approx 81.5/96.0$ in the case of the Madingley Hill transmitter). It is this phase ‘scaling’, required because the calibrator and celestial signals are at different frequencies, which demands that the phase be tracked.
4. Finally, the tapes are passed through the correlator for a second time, but now with a geometrical delay and fringe rate appropriate to the astronomical signals. The results are recorded, using the data from the previous step to correct the phase or its drift (and the equations in Section 2.2.3 to correct for the fringe rotator).

Figure 5.2 shows the effect of this procedure as applied to short (1.321 km) spacing visibility data on Cas A. The amplitude measurements in Figure 5.2(a), show the source to be practically unresolved (the droops at the ends are caused by the primary beam of the 4C). The corresponding uncalibrated phase measurements in (b) ramp steadily between the principle limits of $\pm 180^\circ$ due to a difference in frequency between the rubidium standards, which in this case is more or less constant. Figure 5.2(c) shows the same phase record after calibrating with the Madingley Hill transmitter. The technique is clearly effective – the ramp has gone completely, leaving only residual wiggles due to the ionosphere. Figure 5.3 shows similar results for a 15.116 km baseline. Cas A is now well resolved, so the calibrated phase need not be flat. Most of the phase irregularities, however, are not due to the source, but are generated by travelling ionospheric disturbances passing over the baseline. These can be seen as the large undulations in the phase with a period of about 20 minutes. Another application for the phase-stabilised interferometer becomes clear. The use of celestial radio sources to investigate the ionosphere is by no means new (see for example Hewish, 1952 [43]), but an array of phase-stable receivers over a large area (~ 100 km) would have a unique ability to map ionospheric irregularities as they pass overhead. Such a device would complement other measurements made by more conventional means, such as spaced ionosondes [55], Doppler [92] and satellites [25].

Figure 5.4 shows the data derived from the reference channel on the run shown in Figure 5.3. The top trace shows the amplitude of the correlated signal from the Madingley Hill transmitter. It is reasonably constant, indicating that fading is not a

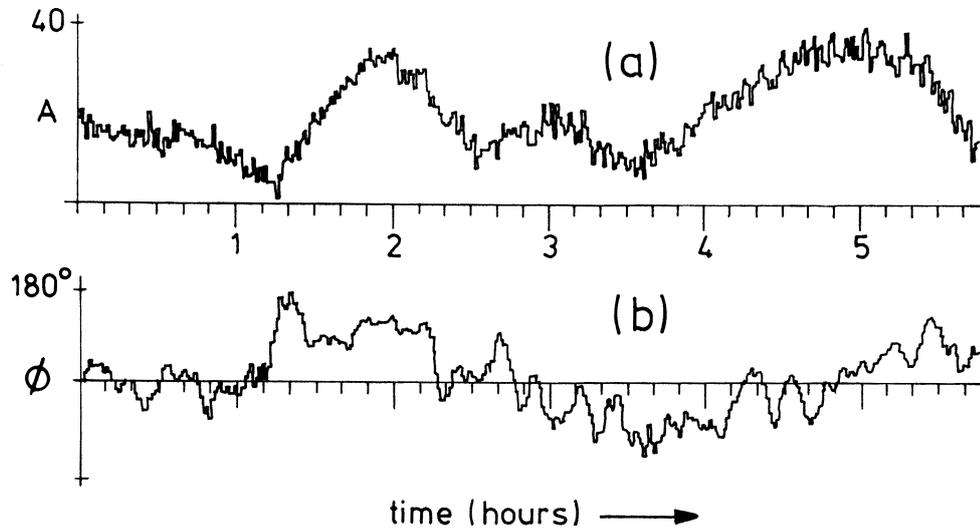


Figure 5.3: The amplitude (a) and calibrated phase (b) of Cas A, as measured on a 15.116 km N-S baseline. The undulations in (b) are caused by irregularities in the ionosphere.

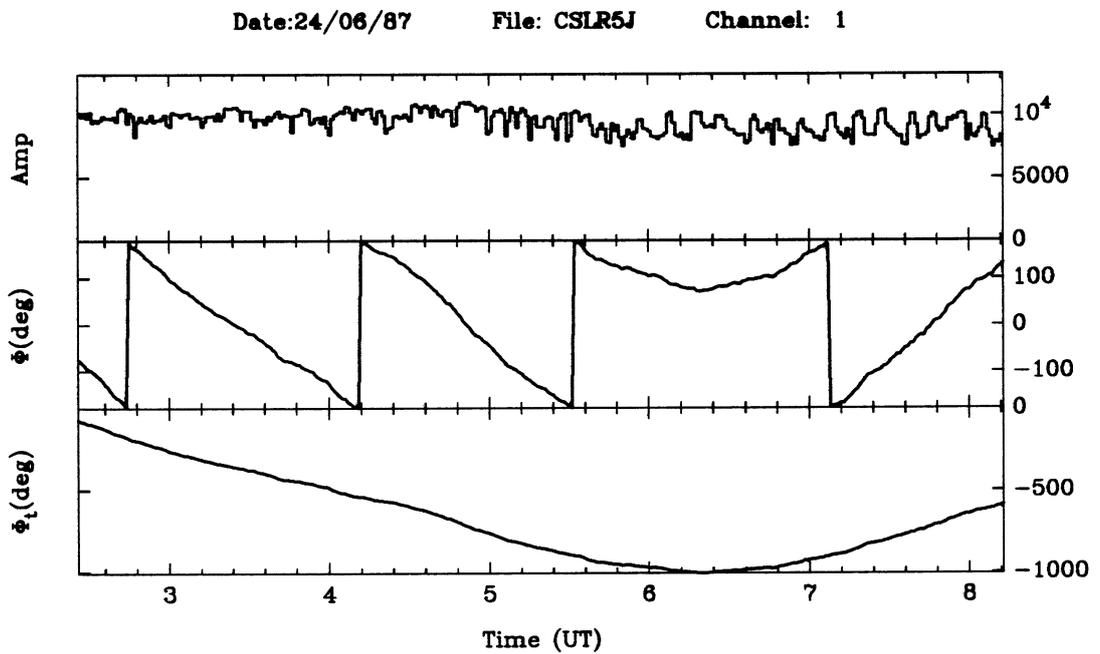


Figure 5.4: The amplitude, phase and tracked phase derived from correlating signals from Radio Cambridgeshire picked up on a 15.116 km baseline. The phase is entirely due to drifts in the local oscillators.

problem on this baseline and that the bandwidth is sufficient to capture most of the FM signal, regardless of modulation index. The middle trace shows the correlated phase in its principle range and the bottom trace the result of tracking this phase. At about 0620 UT, the two Rb standards had the same instantaneous frequency.

5.1.3 How good is the calibration?

This referencing approach clearly works well on short baselines. The phase record in Figure 5.2(c) speaks for itself. However, the real test of the technique is on the longer baselines, when propagation effects are expected to be more severe, and stability harder to maintain. Unfortunately, celestial radio sources suffer badly from the ionosphere on such baselines, so the ‘true’ phase of the interferometer is no longer constant, and the calibration is hard to verify. However, the calibrations derived simultaneously from two independent transmitters should always be the same. If the referencing technique works, the two transmitters should generate identical phase histories for the drifts between the Rb standards.

Consider therefore the following measurements of signals from the Madingley Hill and Tacolneston transmitters, made with the motorcaravan in Chatteris, on a baseline of 33.79 km (see Figure 5.5). Simultaneous recordings were made of signals from the two transmitters, and the resulting (untracked) phase records are shown in the top two traces in Figure 5.6. It is clear that the Rb standards were not well matched on this run, as the correlated phases ramp steeply. The bottom trace in Figure 5.6 shows the result of subtracting the tracked and frequency-corrected upper two records. This then represents the amount by which the path lengths from one transmitter to the ends of the baseline were changing relative to those from the other transmitter. At 100 MHz a phase change of 1° corresponds to a path length of about 1 cm, so these paths were stable to about 30 cm over a total path length of 220 km – a remarkable result! The wiggles in middle trace show that most of the phase errors were almost certainly in the signals from the more distant transmitter. This is to be expected, and a possible cause of it is discussed in Section 5.2.3. Figure 5.7 shows similar results obtained from a much shorter baseline of 3.640 km (Toft). The deviations in the bottom trace are certainly reduced on this baseline, but are again dominated by those in the signals from the distant Tacolneston transmitter.

In conclusion, the phase referencing technique appears to perform well. These tests show that the corrected phase is certainly good to 30° over baselines approaching 50 km. It is also clear that an even higher degree of stability can be attained if a local transmitter is used as the reference.

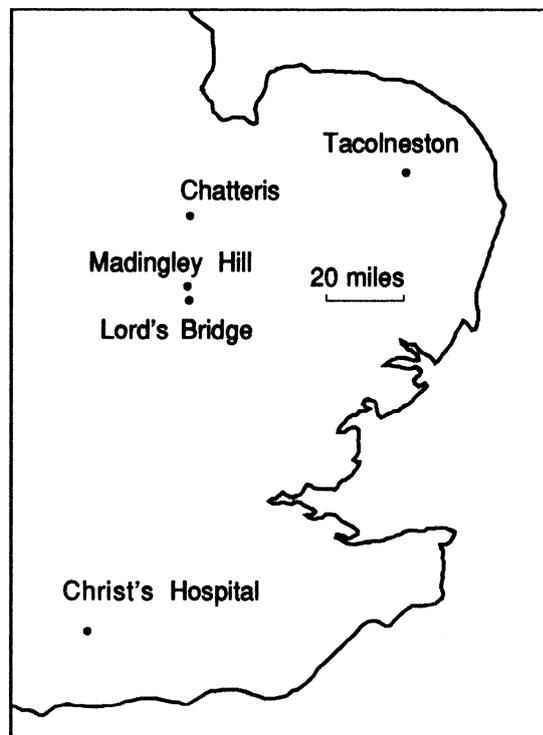


Figure 5.5: The baseline (Lord's Bridge to Chatteris) and reference transmitters (Tacolneston, Madingley Hill) used to test the calibration.

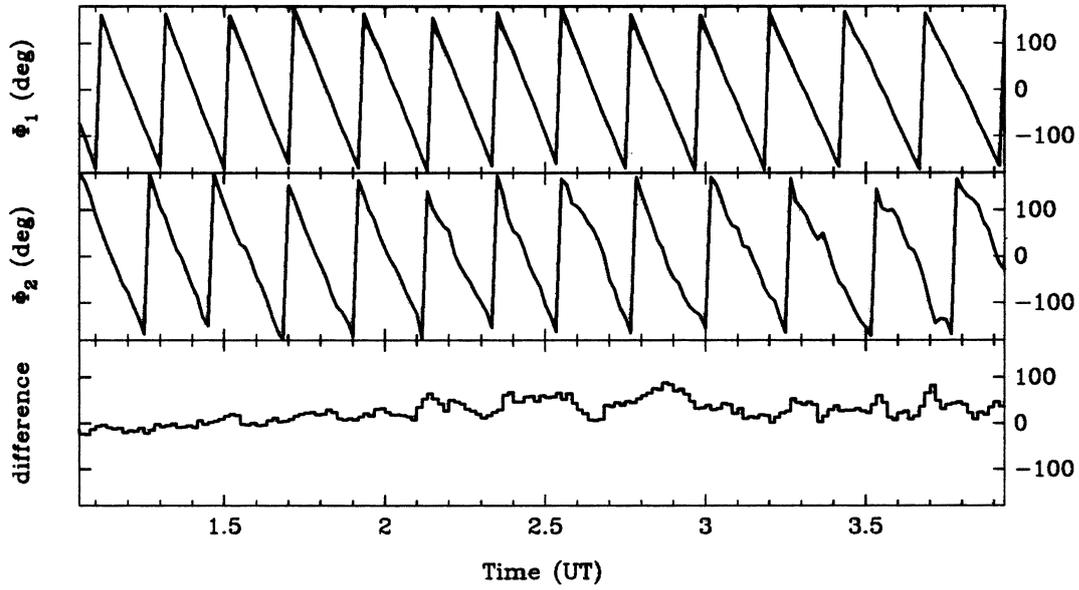


Figure 5.6: Calibrations derived on a 33.79 km baseline (Chatteris) – The top trace shows the phase of the correlated signals from Madingley Hill (ϕ_1), and the middle trace the phase from Tacolneston (ϕ_2). The bottom trace shows the frequency-corrected difference between the two.

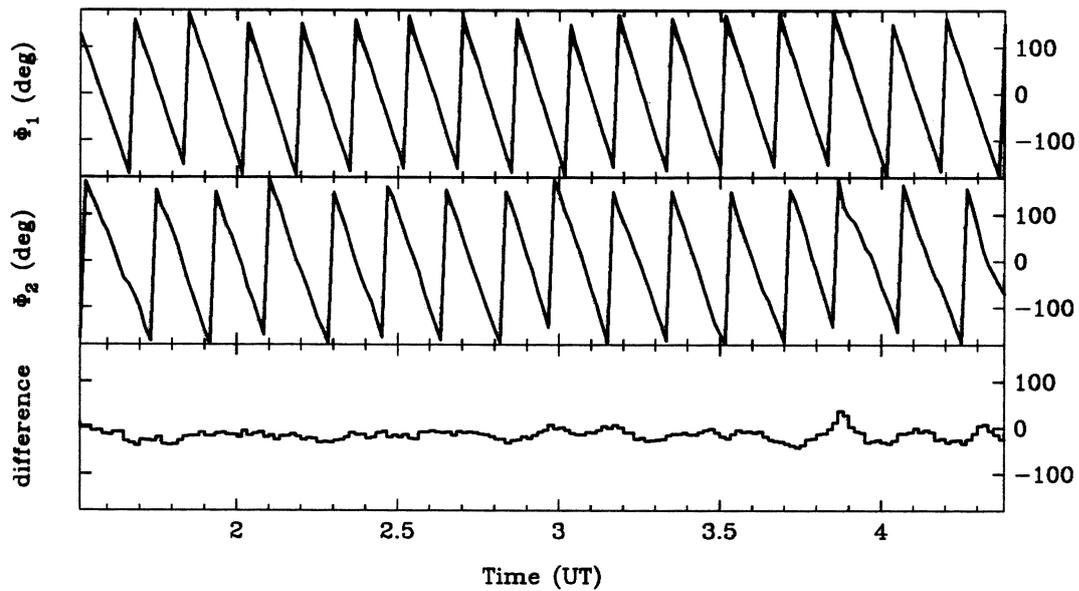


Figure 5.7: As Figure 5.6 but on a shorter, 3.640 km baseline.

5.2 Accuracy and stability

Having shown that this new method of stabilising a disconnected interferometer works well in practice, the factors that ultimately limit the method's effectiveness are now considered. One of the most obviously unpredictable features of the system is its reliance on radio signals that are frequency modulated by music or speech. A formal analysis of this is perhaps inappropriate; however it must be pointed out that the mean power spectrum of the signal should be symmetric or constant, and the centre point fixed, over periods longer than a small fraction of the coherence time. Failure to fulfil this condition will introduce a time varying phase from the frequency integral term in Equation 2.9. Spectral measurements on the BBC FM signals, together with practical experience, show this not to be a problem.

5.2.1 Phase tracking

It is a little unfortunate that, in order to calibrate signals of one frequency with those of another, the phase of the second needs to be evaluated explicitly. This is because phase is a noisy, ambiguous quantity that is difficult to work with, whereas the original complex signal from which the phase is derived is much more robust. To make matters worse, the *tracked* phase of the calibrator must be evaluated, and this can only be done with confidence when the signal is reasonably strong. Fortunately, the correlated signals from VHF transmitters are usually very strong.

The output from the correlator consists of a stream of complex numbers. In order to generate the full phase history of the frequency standards, it is necessary to take account of the times that the phase of the correlated signal from the reference transmitter passes through 180° . When the signal is strong, these moments can be easily identified as the points at which the phase record suffers a discontinuity. Such events can be seen clearly in Figure 5.4 at about 90-minute intervals. However, when the signal is weak or the phase is ramping very rapidly, the points are harder to pick out. Under these conditions, a more robust method must be used.

To show how this may be done, consider the geometrical relationships between two phases, ϕ and $\phi + \Delta\phi$ of

$$\begin{aligned}\sin \Delta\phi &= \sin(\phi + \Delta\phi) \cos \phi - \cos(\phi + \Delta\phi) \sin \phi, \\ \cos \Delta\phi &= \cos(\phi + \Delta\phi) \cos \phi + \sin(\phi + \Delta\phi) \sin \phi.\end{aligned}\tag{5.3}$$

If ϕ represents the 'tracked' phase up to some time (which may be a number $> 2\pi$), and $\phi + \Delta\phi$ the correlated phase of the next measurement (a number in the principle

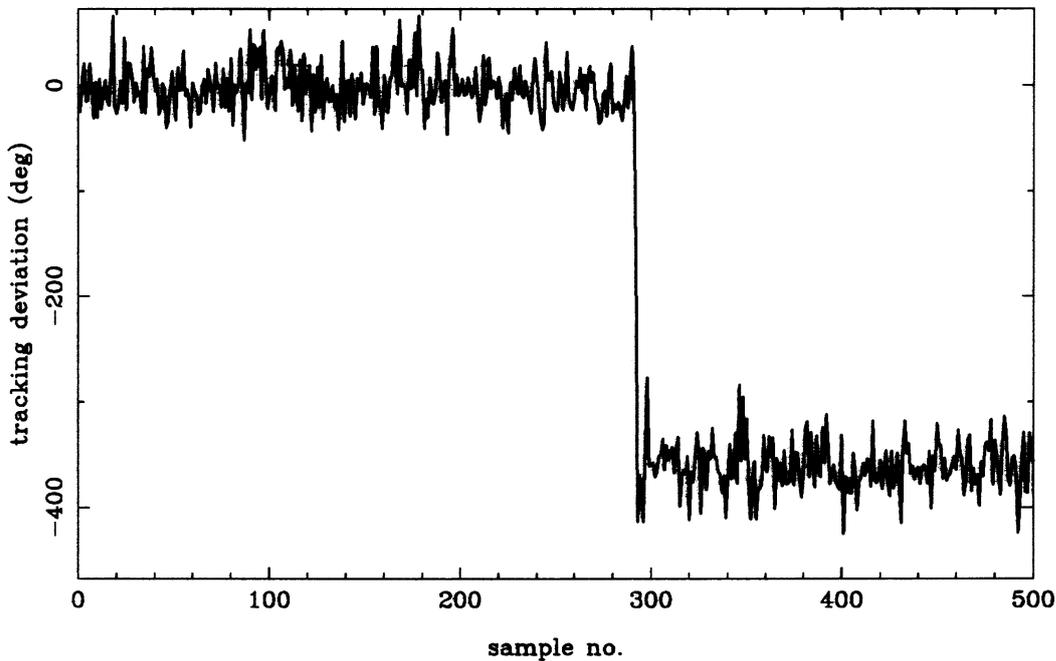


Figure 5.8: A tracking error simulation. The signal being tracked was rotating at 70° per sample point and with a S/N ratio of 2.

range), the phase difference between the two, reduced to the range $-\pi$ to π , may be determined directly from the expression above. This phase difference $\Delta\phi$ can then be added to the total tracked phase, and the result used to process the next measurement, and so on. This operation gradually builds up a phase record that can stray outside the principle range of $[-\pi, \pi]$ when required, but which, when reduced to this range, is identical to the original, untracked phase. The bottom two traces in Figure 5.4 serve as an example of this.

This algorithm is simple, but has proved perfectly adequate. Note that although a lot of incremental phases are being added together, there is no accumulated phase error. The only error that can creep in is a full revolution of 2π in the tracked phase. Such an event will be referred to as a ‘tracking error’. Tracking errors occur only in situations of low signal-to-noise ratio, or when the phase is changing by a large amount between samples. The performance of the algorithm under these conditions can be best determined by simulation. Figure 5.8 shows the result of such a simulation, in which a signal embedded in Gaussian noise ($\gamma = 2$) and rotating at 70° per sample point was tracked by the algorithm. The y-axis displays the difference between the ‘true’ tracked phase and that determined from the data – the tracking error can be seen clearly at sample 290. The algorithm’s performance

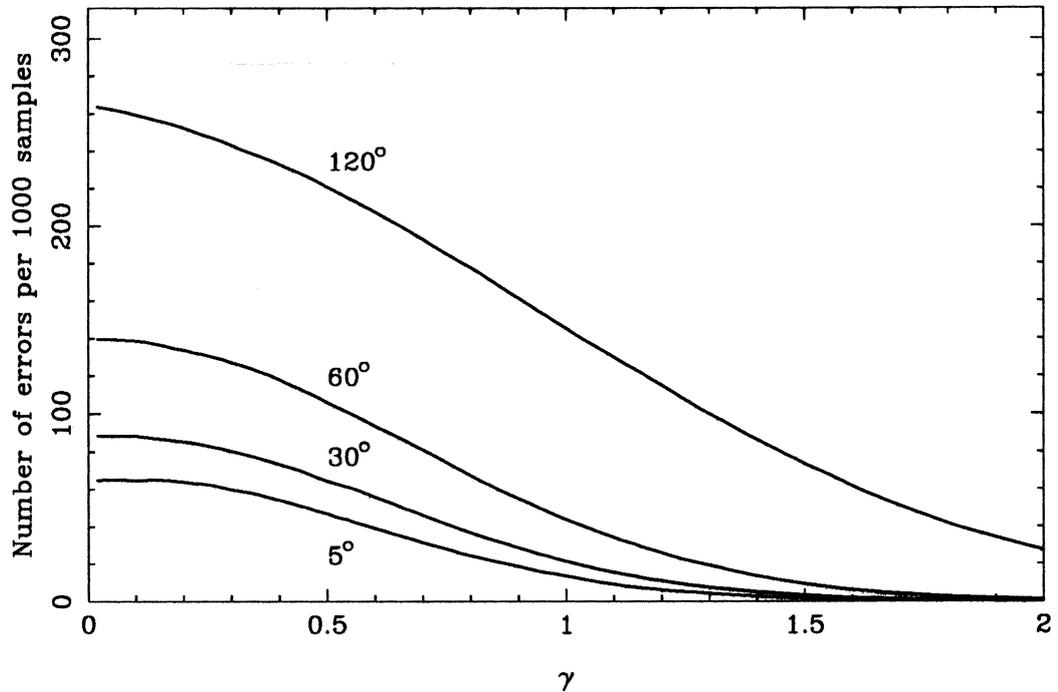


Figure 5.9: Ability of the algorithm to track a vector rotating at 5, 30, 60 and 120° per sample under different noise conditions. The vector has a S/N ratio of γ .

under various other conditions of signal-to-noise ratio and phase slope is shown in Figure 5.9. The graphs show the expected number of tracking errors of this type in 1000 tracked samples, for signal-to-noise ratios (γ) in the range 0 to 2. Examples using four phase slopes (5, 30, 60 and 120 degrees per sample) are given.

The signals used for calibration are expected to be consistently strong. Weak signal strength usually corresponds to situations where tropospheric propagation modes become important, and when the calibration may become unreliable. The simulations show therefore that under conditions in which calibration is useful (that is $\gamma \geq 2$) tracking can nearly always be maintained, even for the steepest of phase slopes, and tracking errors need never occur.

5.2.2 Consequences of 1-bit sampling

The effects of 1-bit digitisation and correlation on two bandlimited noise signals with a finite cross correlation coefficient, ρ , are well known, and are described by

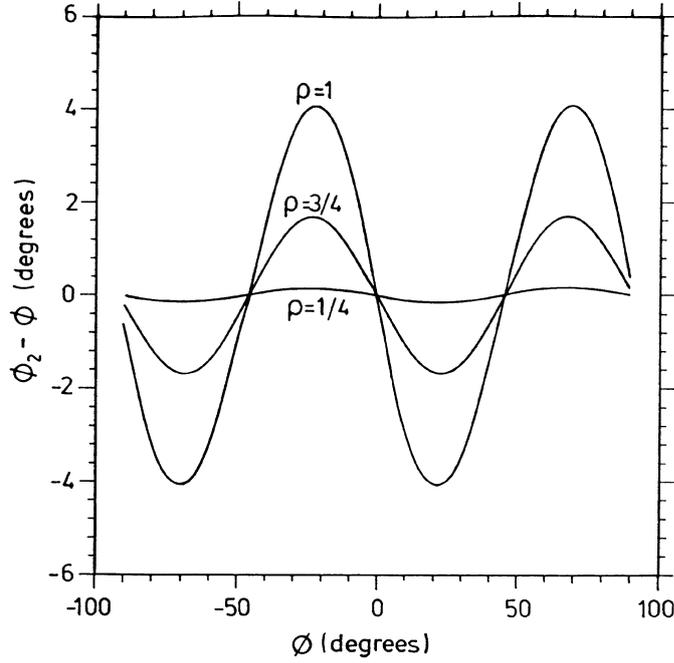


Figure 5.10: The phase error ($\phi_2 - \phi$) in the correlated signal introduced by 1-bit sampling. ρ is the correlation coefficient of the two noise signals.

the ‘Van Vleck’ relationship [99]

$$\rho_2 = \frac{2}{\pi} \arcsin \rho, \quad (5.4)$$

where ρ_2 is the measured cross-correlation coefficient. A complex correlator uses representations of both the real and imaginary components of $\boldsymbol{\rho} = \rho \exp(i\phi)$. The result of 1-bit digitisation is that the measured phase, ϕ_2 , is given by

$$\phi_2 = \arctan \left[\frac{\arcsin(\rho_0 \sin \phi)}{\arcsin(\rho_0 \cos \phi)} \right], \quad (5.5)$$

which differs from the true phase ϕ by an amount

$$\phi_2 - \phi \simeq \frac{\rho_0^2 \sin 4\phi}{8(2 - \rho_0^2 \sin 2\phi)}, \quad (5.6)$$

(Figure 5.10). In the case of well-correlated noise signals ($\rho_0 \approx 1$) the expression for ϕ_2 reduces to

$$\phi_2 = \arctan \left[\frac{2\phi}{\pi - 2\phi} \right]. \quad (5.7)$$

The measured phase can therefore deviate from the true phase by up to 4° and must be corrected. In the case of signals from an FM transmitter, however, the received

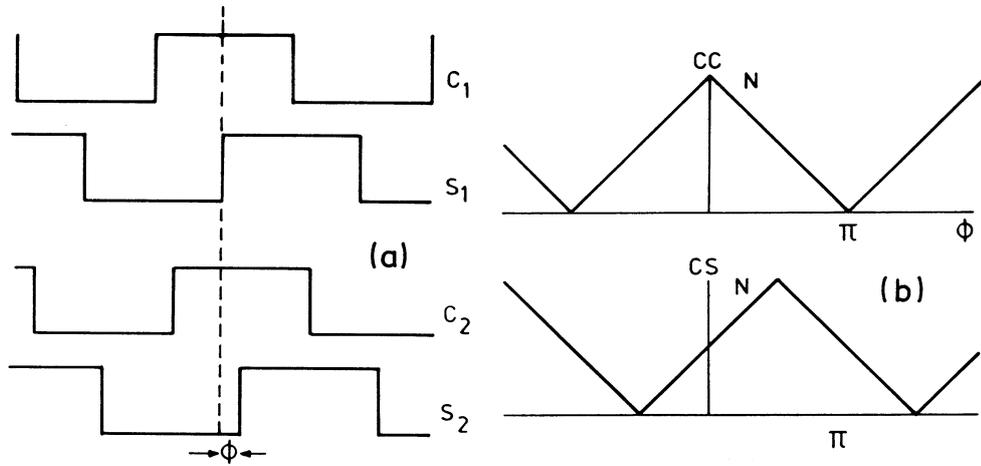


Figure 5.11: 1-bit correlation of noise-free sinusoids. (a) shows the real (C) and imaginary (S) components of the digitised signals v_{1b} and v_{2b} . The measured cross-correlation coefficients are shown in (b) between the two in-phase components (CC) and the two quadrature-phase components (CS).

power spectrum is frequently more like a δ -function than bandlimited noise, so this analysis is not appropriate. The true spectrum of the FM signal will be somewhere between the two, but to set limits an extreme case will be considered in which the two baseband signals to be correlated, $v_{1b}(t)$ and $v_{2b}(t)$, are sinusoids of frequency ω differing in phase by ϕ , i.e.

$$v_{1b} = \exp[i\omega t] = C_1 + iS_1 \quad (5.8)$$

and

$$v_{2b} = \exp[i\omega t + \phi] = C_2 + iS_2 . \quad (5.9)$$

These are then sampled using 1-bit digitisation to give real and imaginary parts with the forms of square waves (Figure 5.11). The products $\langle C_1 C_2 \rangle$ ($= \langle S_1 S_2 \rangle$) and $\langle C_1 S_2 \rangle$ ($= -\langle S_1 C_2 \rangle$) are generated in the correlator, where the angular brackets denote averages over N samples, and are given by

$$\langle CC \rangle = N(1 - 2\phi/\pi)/2 , \quad (5.10)$$

$$\langle CS \rangle = N\phi/\pi , \quad (5.11)$$

over the region $|\phi| < \pi/2$, and after subtraction of a constant offset of $N/2$. The measured phase is therefore

$$\phi_2 = \arctan \left[\frac{\langle CS \rangle}{\langle CC \rangle} \right] = \arctan \left[\frac{2\phi}{\pi - 2\phi} \right] . \quad (5.12)$$

Comparison with Equation 5.7 shows this to be the same as the result derived for a noise source from the Van Vleck relation, suggesting that it is independent of the signal spectrum. This would mean that the maximum deviation from the true phase that the 1-bit sampling process could be expected to introduce is about 4° (see Figure 5.10 for $\rho = 1$).

In the case of bandlimited noise signals, Equation 5.5 shows that as the covariance between the two signals drops, so the Van Vleck correction to the phase is reduced. It is therefore of interest to consider the effect of introducing broadband noise into the sinusoidal signals mentioned above. This noise, introduced before mixing, is uncorrelated at the two ends of the interferometer and is specifically designed to minimise the difference between the measured and true phases, while at the same time not degrading the signal-to-noise ratio appreciably. Such devices are commonly employed in signal processing to smooth out the jumps generated when a slowly varying function is quantised – the noise is usually referred to as ‘dither’ noise.

Consider therefore two real monochromatic signals that differ in phase by ϕ (the phase we wish to measure), and which contain Gaussian noise with variance σ^2 . The joint probability distribution for samples x_1 and x_2 of the two signals is simply

$$p(x_1, x_2) = g(S_1 \cos \omega t, \sigma)g(S_2 \cos(\omega t + \phi), \sigma), \quad (5.13)$$

where S_1 and S_2 are the amplitudes of the two signals, and $g(\mu, \sigma)$ is a normal distribution with mean μ and variance σ^2 . If the two signals are now 1-bit digitised and sampled N times for a time significantly greater than their period of oscillation, both samples can be expected to be greater than zero N_{11} times, where

$$\begin{aligned} N_{11} &= N \langle \int_0^\infty \int_0^\infty p(x_1, x_2) dx_1 dx_2 \rangle \\ &= \frac{N}{4} \langle (1 + \operatorname{erf} \left[\frac{S_1 \cos \omega t}{\sqrt{2}\sigma} \right]) (1 + \operatorname{erf} \left[\frac{S_2 \cos(\omega t + \phi)}{\sqrt{2}\sigma} \right]) \rangle. \end{aligned} \quad (5.14)$$

The angled brackets denote an average and erf is the error function. Similar calculations give expressions for $N_{\bar{1}\bar{1}}$, $N_{\bar{1}1}$ and $N_{1\bar{1}}$. The measured correlation coefficient can therefore be expressed as

$$\begin{aligned} \rho_2 &= \frac{(N_{11} + N_{\bar{1}\bar{1}}) - (N_{\bar{1}1} + N_{1\bar{1}})}{N} \\ &= \langle \operatorname{erf} \left[\frac{S_1 \cos \omega t}{\sqrt{2}\sigma} \right] \operatorname{erf} \left[\frac{S_2 \cos(\omega t + \phi)}{\sqrt{2}\sigma} \right] \rangle \end{aligned} \quad (5.15)$$

If the original signals are complex both the symmetric and antisymmetric correlations must be considered, but otherwise the analysis is identical. The cross-correlation coefficient between two monochromatic complex signals (signal-to-noise

ratios γ_1 and γ_2) differing in phase by ϕ is

$$\begin{aligned} \rho_2 &= \langle \text{erf}[\gamma_1 \cos \omega t] \text{erf}[\gamma_2 \cos(\omega t + \phi)] \rangle \\ &\quad + i \langle \text{erf}[\gamma_1 \cos \omega t] \text{erf}[\gamma_2 \sin(\omega t + \phi)] \rangle . \end{aligned} \quad (5.16)$$

Two limiting cases should be considered. If the signals are strong ($\gamma_1, \gamma_2 \simeq 1$), then the erf function terms become square waves, and the real and imaginary parts of the expression revert to the zero noise results stated in Equations 5.10 and 5.11. But if the signals are weak ($\gamma_1, \gamma_2 \ll 1$), the same terms become sinusoidal, and ρ_2 approaches $2\rho/\pi$. The effect on 1-bit sampling is now to leave the correlated phase unaltered, but to degrade the amplitude by $2/\pi$. So, taking the two signal-to-noise ratios to be the same for simplicity ($\gamma_1 = \gamma_2 = \gamma_0$), the measured phase, $\arg(\rho_2) = \phi_2$, deviates from the true phase ϕ by an amount depending only on γ_0 and when γ_0 is small, ϕ_2 approaches ϕ (Figure 5.12). This all looks very useful – by adding noise the systematic phase deviations generated in the quantisation process can be removed. However, it must be remembered that although the phase is now an *unbiased* estimate of the true phase, it is much more noisy, and a greater integration time required after correlation for the estimate to be *accurate*. What is needed is an optimum amount of dither noise, to be introduced before digitisation, determined by the integration time available after correlation.

In the case where $\gamma_c \sqrt{N} \gg 1$, the probability distribution of the phase noise on the correlated and integrated signal is well described by

$$p(\theta) = \gamma_c \sqrt{N/\pi} \exp(-N\gamma_c^2 \theta^2) , \quad (5.17)$$

where γ_c is the signal-to-noise ratio in one correlated sample and N is the number of samples over which the integration is made (c.f. Equation 3.12). Random fluctuations in the phase of about $\sqrt{2}/(\gamma_c \sqrt{N})$ radians can therefore be expected. This supplies the constraint on how much dither noise should be added. The total phase error is minimised with a dither that reduces the deviation from the true phase introduced by 1-bit digitisation to the level of the phase noise, i.e.

$$(\phi_2 - \phi)_{\max} = \sqrt{2}/(\gamma_c \sqrt{N}) . \quad (5.18)$$

Before the optimum noise level in the pre-correlated signals can be calculated, it is necessary to obtain a relationship between the signal-to-noise ratio of the *correlated* signal, γ_c , and that of the data streams being correlated, γ_0 . Taking the real component of these streams as an example, the i th elements in the streams ${}_a x_i$ and ${}_b x_i$, can each take the value of +1 or -1. The correlator produces the sum of N

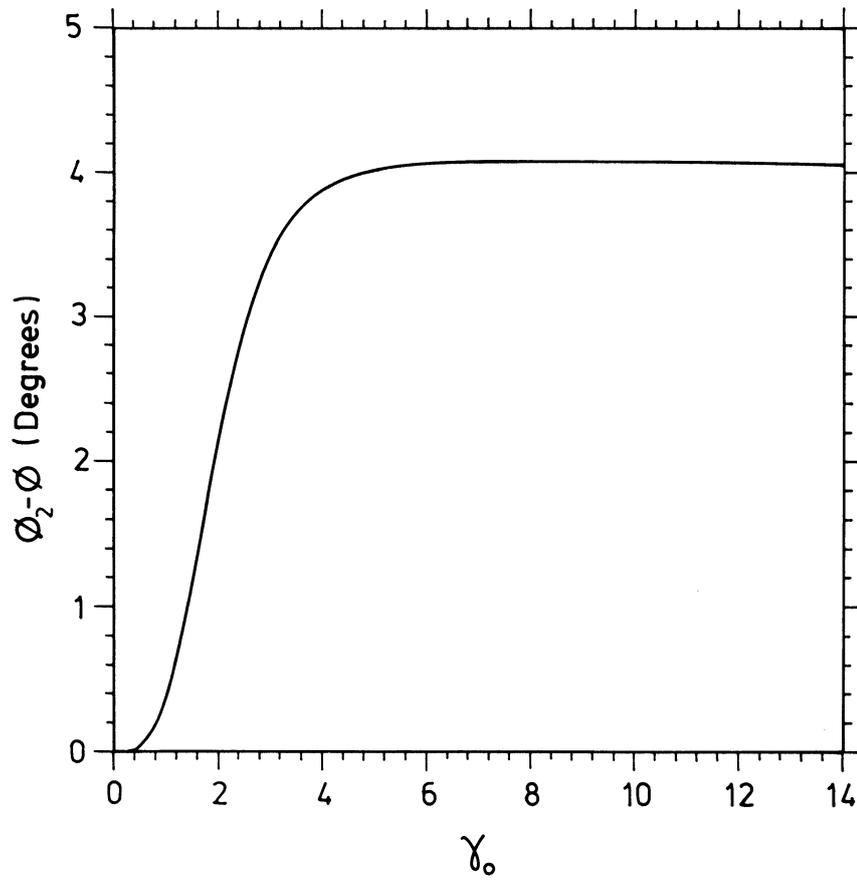


Figure 5.12: The effect of dither noise on the maximum deviation ($\phi_2 - \phi$) of the measured phase from the true phase. γ_0 is the pre-correlated signal to noise ratio.

cross-products between the two, i.e.,

$$r_{2,x} = \frac{1}{N} \sum_{i=1}^N a x_i b x_i . \quad (5.19)$$

The expectation value of $r_{2,x}$ is simply the true 1-bit correlation coefficient, $\rho_{2,x}$. The variance of $r_{2,x}$ is

$$\begin{aligned} \text{var}(r_{2,x}) &= \langle r_{2,x}^2 \rangle - \langle r_{2,x} \rangle^2 \\ &= \left\langle \frac{1}{N^2} \sum_i \sum_j a x_i b x_i a x_j b x_j \right\rangle - \rho_{2,x}^2 \\ &= \frac{1}{N} (1 - \rho_{2,x}^2) . \end{aligned} \quad (5.20)$$

Extending to the complex case, the variances of the symmetric and antisymmetric products add to give

$$\text{var}(r_2) = \frac{1}{N} (1 - \rho_2^2) , \quad (5.21)$$

where $\rho_2^2 = |\rho_2|^2$. So the complex signal-to-noise ratio in the estimate of the correlation coefficient is

$$\gamma_c = \frac{\rho_2}{\sqrt{2(1 - \rho_2^2)}} . \quad (5.22)$$

Figure 5.10 shows that the maximum phase deviation occurs when the true phase is $\pi/8$. Equating this deviation to the expected phase noise gives the optimum pre-correlated signal-to-noise ratio, γ_0 , that minimises the total phase error after integrating for N samples, i.e. the solution to

$$\arg(\rho_2) - \frac{\pi}{8} = \frac{2(1 - \rho_2^2)^{1/2}}{\sqrt{N}\rho_2} \quad \text{at } \phi = \pi/8 . \quad (5.23)$$

A plot of this function is shown in Figure 5.13. Note that $\rho_2 \simeq [\text{erf}(\gamma_0)]^2$. Care must be taken here to distinguish between γ_0 , the signal-to-noise ratio of the raw, uncorrelated data and that of the integrated and correlated data. Phase tracking is concerned with the second ratio, a number $\sim \sqrt{N}$ greater than γ_0 . For most integration times, a value for γ_0 of a little over one is optimal. This result is largely because of the rather critical dependence in that region of both γ_c and $(\phi_2 - \phi)$ on γ_0 .

To summarize, the 1-bit sampling process applied to the FM signals does affect the accuracy with which they can calibrate the interferometer. For strong signals, a systematic bias of up to 4° in the phase is present, and this can be reduced if dither noise is introduced to the data before digitisation. Although 4° is a pleasingly small effect, the evident stability of the calibration process, and intrinsic strength of the transmitter signals, make it an important factor.

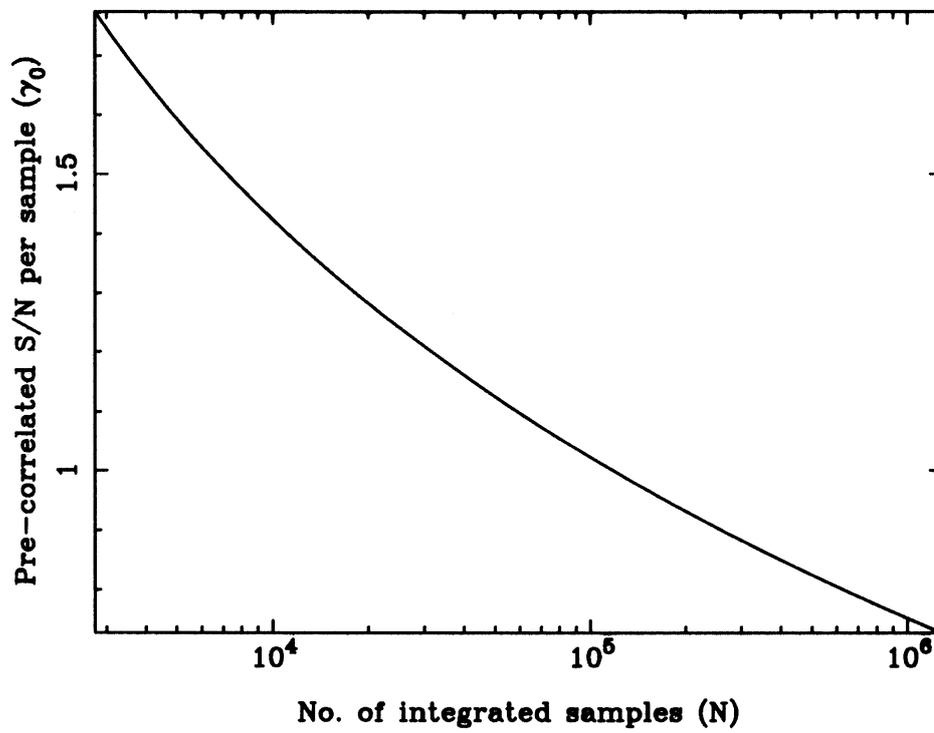


Figure 5.13: The optimum dither noise that minimises the total phase error. γ_0 is the pre-correlated signal-to-noise ratio, and N the number of samples over which the correlated signal is averaged.

5.2.3 Effects from the atmosphere

For a radio transmitter to be useful as a phase calibrator, it is important that the differential propagation path length between it and the two ends of the interferometer is constant (or calculable). Any changes in either route or refractive index along the route will show up in the phase, and hence affect the calibration. The examples given in Section 5.1.3 indicate that changes in the propagation route do not occur on time-scales greater than a few seconds (changes on shorter time scales would largely vanish in the integration). On no occasion has the tracked phase from a transmitter shown an abrupt discontinuity consistent with such an event. Small phase deviations do occur when the antennas are approached by people or passed by vehicles, but the radiation pattern can be kept more or less constant by perching the antennas on masts. This does have the drawback that any bending of the mast by the wind will be picked up in the phase so care must be taken here too, but otherwise strictly geometric effects appear unimportant.

Terrestrial transmitters with frequencies greater than a few tens of MHz suffer very little from effects caused by the ionosphere as would be the case at lower frequencies. Instead, the dominant corrupting influence on time scales greater than ~ 1 minute is the refractive index, η , of the air. The dependence of this on temperature, T , pressure, P , and the water vapour pressure, e , is described to within 1% at radio wavelengths by the relation [83]

$$(\eta - 1) \times 10^6 = (P + 4810e/T)(77.6/T), \quad (5.24)$$

where P and e are expressed in mb, and T in K. Using values of $P=1000$ mb, $T=290$ K and $e=20$ mb this gives

$$10^6 d\eta = -1.53dT + 0.26dP + 4.43de. \quad (5.25)$$

For typical changes in T , P and e the three terms are of similar magnitudes, and propagation paths can be expected to deviate from the mean by up to 2cm per km. How much this affects the *differential* path length between the transmitter and the two receivers will depend on the baseline geometry, but it would be reasonable to expect an uncertainty of up to 1° in phase per km of baseline for the geometries and frequencies used in these experiments.

On very long baselines, when the transmitter is well below the horizon with respect to one of the receivers, weak tropospheric ducting becomes the dominant propagation mode. This occurs where the refractive index of the air varies suddenly, leading to a bend in the path of the ray with a radius much shorter than a wavelength, and mostly affect the VHF band. Although the reflection coefficient may

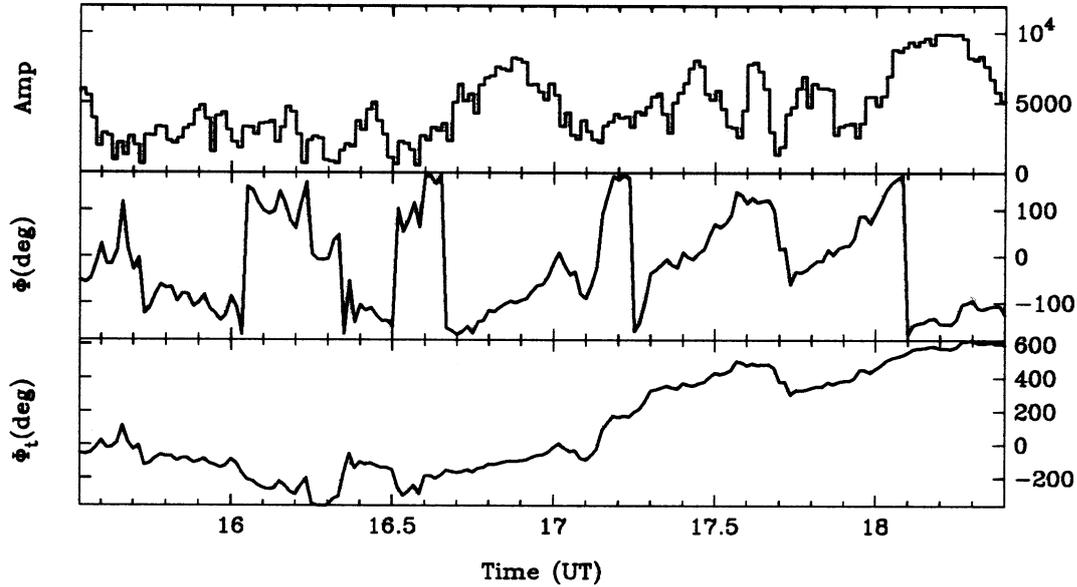


Figure 5.14: *Correlated amplitude, phase and tracked phase from Tacolneston, with the mobile end of the interferometer at Christ's Hospital. A propagation distance of 167.6 km has resulted in fading.*

be just a small fraction of a percent, the ducted signal can still be much stronger than the ground wave at long distances. Turbulence in the atmosphere can therefore create a random speckle pattern on the ground that destroys any phase information in the signal. Scintillation is thought to be strongest during the second half of the night [62], and probably contributed to a significant fraction of the phase noise present in the later portion of Figure 5.6.

These effects can be seen more clearly in Figure 5.14. This shows the amplitude, phase and tracked phase from the Tacolneston transmitter as measured on a baseline between Lord's Bridge and Christ's Hospital, near Horsham (see Figure 5.5). The distance between Christ's Hospital and Tacolneston is about 168 km, with Greater London directly between the two, so good ground wave propagation is very unlikely. The amplitude is clearly corrupted by fading, showing many dips and crinkles, and the phase also deviates markedly from the smooth curve that would be expected if it truly represented the drift of the frequency standards. This then sets a limit on the method's usefulness. Propagation paths of much more than 100 km are not sufficiently stable for the derived phases to be useful in calibration, so the maximum practical baseline, with the transmitter at the centre, is about 200 km.

5.2.4 Receiver systematics

Finally, it is important to consider the effects that the radio receivers themselves have on the phase of the correlated signal. A perfect mixer would multiply the radio signal by two sinusoids, in phase quadrature, derived from the local oscillator. The simplest, and most common systematic effect introduced by the mixer occurs if the sinusoids are no longer orthogonal, but are ‘sheared’ by a phase angle α . In general, if the receivers at the two ends of the baseline are sheared by phase angles of α_1 and α_2 respectively, then the measured correlation coefficient ρ_s is related to the true correlation coefficient ρ by

$$\rho_s = \cos(\alpha_1 - \alpha_2)\Re[\rho] + i \cos(\alpha_1 + \alpha_2)\Im[\rho], \quad (5.26)$$

where \Re and \Im denote the real and imaginary components of their arguments. As a result, any rotating correlated signal will be given an amplitude ripple of

$$\frac{\Delta\rho}{\rho} = 2 \sin \alpha_1 \sin \alpha_2 \quad (5.27)$$

and a phase ripple of $\Delta\phi$, where

$$\tan \Delta\phi = \tan \alpha_1 \tan \alpha_2. \quad (5.28)$$

The system is therefore surprisingly insensitive to small misalignments of the receivers. For example, if both ends were sheared by 10° , the resulting phase ripple would be less than 2° .

5.3 Review

These results show that a disconnected radio interferometer can be made phase-stable by receiving the broadband radio emissions from BBC VHF transmitters. The spacial coherence of this radiation is modelled as stationary, so that any measured changes can be attributed to the limited temporal coherence of the equipment. The measurements are used retrospectively to calibrate the interferometer so that it yields information on signals whose coherence is unknown.

On baselines less than about 100 km, tests show that the calibrated interferometer can be considered ‘rigid’ to just a few centimetres, and that this rigidity need be limited by neither the equipment nor the processing. On longer baselines, when the reference transmitter is well below the horizon, tropospheric scintillation dominates over ground wave propagation. The radiation no longer has a stationary spacial coherence function, and accurate calibration is lost.

Chapter 6

The structure of Cassiopeia A at low frequency – short baseline observations

6.1 Introduction

The supernova remnant Cassiopeia A (henceforth Cas A) is one of the brightest objects in the sky at radio wavelengths and has consequently received a great deal of attention, both as a test of supernova models and as a source suitable for mapping with high dynamic range and resolution. Its structure and evolution at frequencies greater than ~ 1 GHz have come under particular scrutiny in recent years, and a consistent model is emerging [13, 17]. However its behaviour at lower frequencies has been studied relatively little and there have been few structural measurements made under 0.5 GHz. Indeed, at metre wavelengths, only the CLFST has a spacing range sufficient to map the remnant. Despite attempts to site a 75 MHz array at the VLA [61], no existing low frequency instrument (other than the Cambridge long baseline interferometer) has both the short spacings and sub-arcminute resolution required to make detailed measurements. Many of the measurements that do exist are total power observations reporting anomalous low frequency flares [28, 66, 67, 102] or, in the case of Hutton *et al.* [47], a steep spectrum compact VLBI component not seen at higher frequencies. All these observations are somewhat suspect due to difficulties in data calibration, but nevertheless deserve serious consideration.

To resolve some of these anomalies, and to demonstrate the phase calibration technique described in the previous chapter, a series of LBI observations of the

remnant were made in the summer of 1987 on baselines between 0.65 km and 128 km, and with a noise figure of a few janskys. The observations fall naturally into two parts – those made on short ($\lesssim 55$ km) and long (~ 128 km) baselines. This chapter is concerned solely with the former. Chapter 7 will deal with the higher resolution observations.

Although the instrumental phase calibration was known, ionospheric corruption meant that little useful astronomical phase could be expected from the interferometer. Fourier-modulus phase-retrieval techniques do exist, but they require both high signal-to-noise ratios and significant oversampling of the transform plane for reliable operation (see [7] for a good review), neither of which were practical with a portable LBI system. Direct mapping of the remnant was therefore impossible. Instead an indirect, but potentially more informative method was employed. The VLA regularly generates maps of Cas A with arcsecond resolution as part of a continual programme to investigate the structural evolution of the remnant [14]. One such map, made at 20 cm with the array in A,B,C and D configurations, and at precisely the same time as the LBI observations, was kindly supplied by Robert Braun of NRAO. The quality and epoch of this map meant that it could be regarded as the ‘definitive’ high frequency image corresponding to the 81.5 MHz observations, so that discrepancies between the low frequency amplitude data and the modulus of the transform of this map were significant spectral features. The short baseline LBI measurements had a maximum resolution of about $14''$ and were therefore expected to reveal gross features – perhaps disentangling some of the thermal and non-thermal emission, or showing a difference in spectral index between the outer shock front and the large bow-shock features associated with the inner ring.

6.2 Current impressions of Cas A

Cas A is thought to be the remains of an optically dim type II supernova that occurred in 1680 [2]. At present the remnant has an angular diameter of about 4 arcminutes at a distance of 2.9 kpc, and a radio spectral index of 0.77 [3]. X-ray images [30] show a diffuse ‘plateau’ of emission and a more clumpy inner ring. Fabian *et al.* infer a total mass of at least $15M_{\odot}$ for the x-ray emitting gas, although this has been reassessed by Braun [13] at more like $3M_{\odot}$. Optically, the remnant shows three main features. As well as a diffuse ring of emission there are also brighter ‘knots’ with velocities that are either high (~ 5000 km/s) and directed radially outwards, or slow (~ 150 km/s) and with a more disorderly velocity distribution (see for example [97, 98]). The fast moving knots are thought to be recombining clumps of dense

ejecta, heated by their passage through the shock front of the initial blast wave, which has been decelerated by swept up interstellar matter. The slow moving knots (sometimes referred to as ‘quasi stationary flocculi’) are interpreted as material blown off by the stellar wind during an earlier, non-explosive, phase in the evolution of the progenitor star [51] and which have recently been heated as they have been overtaken by the main shock front. Both types of knot have optical lifetimes of ~ 20 years. The enhanced abundances of He and N in the slow moving knots, and the large amount of oxygen and oxygen burning products in the fast moving knots, are consistent with a very massive progenitor star, perhaps of $50 M_{\odot}$ [27].

The radio emission is thought to be generated by high velocity ejecta colliding with the interstellar material swept up by the expanding shock. It has a similar structure to the x-ray emission, with an ‘outer’ shock from the blast wave and an ‘inner’ or back shock due to the piling up of ejecta. The radio evolution does not show quite the highly ordered expansion of the fast optical knots [8, 95] but a current outer shock velocity of ~ 2770 km/s can still be inferred, together with an initial expansion velocity of ~ 11000 km/s and a mass ratio (swept-up/ejected) of 4 or 5. The physical conditions in and around the remnant have been reviewed by Braun [13] using existing data and including IRAS observations of shock-heated dust, and Contini [17] has shown that the observed X-ray, optical and radio emissions are consistent with a model of the remnant based on the above description.

The dominant process that shapes Cas A is slowly emerging from a VLA ‘animation’ programme that now includes frames over a period of eight years [14]. This clearly shows that the outer shell, powered by the initial blast wave, has been significantly decelerated and that, in general agreement with Gull [38], the majority of the radio emission comes from the Rayleigh-Taylor unstable zone between the swept-up ISM and the back shock generated by the decelerating diffuse ejecta. More slowly expanding but as yet undecelerated clumps of ejecta periodically smash through this shell. These small clumps are much denser than the shocked ISM and so pass through without significant disruption. They correspond to the fast optical knots. Larger more diffuse clumps do interact with the medium and generate significant localised shocks which compress the magnetic field and lead to particle acceleration. There are at present about eight such active sites visible around the remnant, and it is thought that these, together with the remains of earlier events, are responsible for most of the total radio emission and its blotchy appearance [14].

6.3 LBI observations

Observations of Cas A were made on thirteen ‘short’ baselines between the 4C antenna and the motorcaravan using the phase stabilisation technique described in Chapter 5. Tuffs [94] has shown that the mean flux density from the remnant at 5 GHz falls off with increasing baseline approximately as a power law. Scaling these results to 81.5 MHz ($\alpha = 0.77$) gives an approximate flux-baseline relation of

$$S(D) \simeq 28881D^{-1.477} \text{ Jy} \quad 3\text{km} < D < 60\text{km} , \quad (6.1)$$

so that the expected flux density on a 50 km baseline, corresponding to a resolution of about $15''$, is ~ 100 Jy. The equipment was sufficiently sensitive to achieve this resolution with just one or two Yagis at the mobile station. The observations were actually made with two Yagis, which were slotted into tubes bolted to the front and rear of the motorcaravan. This greatly simplified the process of establishing each new site and eliminated the need for soft ground on which to erect the antennas. Although the Yagis were near to a considerable amount of equipment and metalwork in this configuration, their performance was not noticeably affected and the method was used successfully for all the short baseline measurements.

The receiver electronics and correlator software were already configured for two-polarisation measurements, a feature not required on these short baselines. Rather than interfere with this, the Yagis were both polarised E-W but their signals processed as if they were different polarisations right up to the final stages of correlation and data reduction. The residual fringe rate between them was then removed, and the two signals added coherently. Instrumental phase calibration was taken from the Tacolneston Radio 2 transmitter, as Madingley would be too weak on the longer baselines.

The baselines used, together with their parameters, are shown in Table 6.1 and the corresponding (u, v) plane coverage for Cas A is shown in Figure 6.1. It was important that the length and orientation of the baseline vector were known to high precision. A one metre error in the measured length of an E-W baseline would have introduced a slope in the derived phase of about 15° an hour, a figure considerably greater than the expected error due to poor stabilisation. Positional errors also introduced a *differential* phase slope between different frequency channels which amounted to about half a degree per metre for the observation periods and bandwidths used. Baseline errors greater than a few tens of metres would therefore have significantly degraded the coherence of the interferometer. The technology needed for measuring long baselines to high precision does exist, notably in the form of the satellite-based ‘Global Positioning System’ (GPS) that claims accuracies exceeding

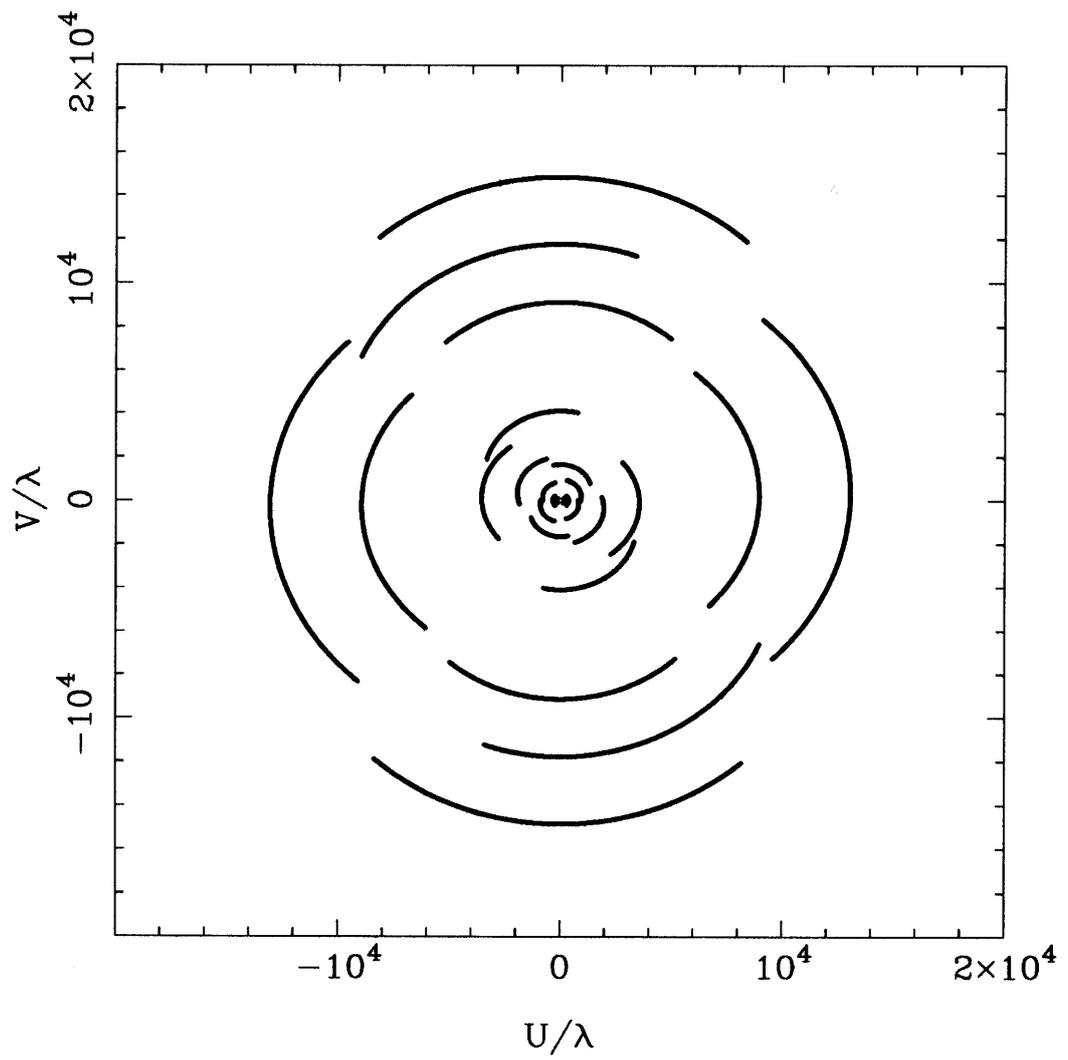


Figure 6.1: *The (u, v) plane coverage from the 'short' baseline observations (sky coordinates).*

	latitude ($^{\circ}$ ' ")			E. longitude ($^{\circ}$ ' ")			height (m)	hr ang. (rad.)	dec. (rad.)	length (m)
Mullard	52	9	50.52	0	2	27.97	20	-1.628576	0.034636	652.4
1 Mile	52	9	47.66	0	3	3.20	22	-1.545929	-0.022743	1321.3
Barton	52	10	47.24	0	4	12.72	25	-2.064604	0.352114	3194.2
Toft	52	10	55.89	0	0	-44.90	27	2.068196	0.355547	3640.1
Foxton	52	6	49.54	0	4	7.52	23	-0.527089	-0.592185	6105.5
Downing	52	11	58.11	0	7	32.31	15	-2.028207	0.327798	7567.8
Fulbourn	52	10	36.99	0	13	18.89	22	-1.662656	0.069119	13104.0
Cottenham	52	17	7.70	0	7	46.30	11	-2.584802	0.580551	15115.9
Haverhill	52	8	21.80	0	30	43.02	122	-1.509009	-0.047894	32986.5
Chatteris	52	28	2.09	0	2	38.13	7	-3.110809	0.656948	33790.4
Welney	52	31	50.57	0	14	53.27	5	-2.716093	0.611588	43425.2
Shimpling	52	7	44.28	0	43	55.91	70	-1.534478	-0.048493	48109.7
Parson D.	52	39	28.12	0	1	51.71	5	3.140172	0.655696	54983.7

Table 6.1: *Short baseline parameters for Cas A LBI, summer 1987.*

$1 : 10^6$. However, the present cost of receivers (up to £90k each) prohibited their use here. A far cheaper solution was to measure positions directly from 1:2500-scale Ordnance Survey maps. Although these maps have a resolution of about half a metre, their claimed accuracy is considerably less.

The precision of these OS maps was checked independently using two Magnavox MX 1502 satellite surveying receivers, kindly supplied by the British Antarctic Survey. These use the older TRANSIT satellite system and have an accuracy of ± 5 m for point position measurements. Such measurements, made at two sites (Lords Bridge and Cottenham), agreed with positions determined from the corresponding OS maps to within 4 or 5 metres. Phase measurements made on baselines for which Cas A was unresolved confirmed that the the position of the mobile station could be determined from the OS maps with sufficient accuracy for LBI work (see Figures 6.4 and 6.5). The transformations between eastings and northings determined from an OS map and the length and celestial co-ordinates of the corresponding baseline need careful consideration and are described in Appendix G.

6.3.1 Ionospheric effects

The main source of phase corruption in metre wavelength interferometry is the ionosphere, and much work has gone in to studying its effects. At 81.5 MHz, about

four times the plasma frequency of the medium, the excess phase length from an integrated electron column density (or ‘dispersion measure’) of DM is

$$\Delta\phi = \frac{e^2}{2m_e\epsilon_0\omega c}DM, \quad (6.2)$$

where ω is ($2\pi\times$) the observing frequency, (81.5 MHz). The interferometer measures *differences* in this excess phase at spacially separated points on the ground, and so is sensitive to the sizes and powers of the ionospheric irregularities that generate them. The scale-sizes of these disturbances range from ~ 10 cm to ~ 1000 km. The small-scale irregularities produce a fine speckle pattern on the ground that moves and evolves at the same rate as the ionosphere above and so leads to decorrelation over periods greater than a few seconds. However strong scintillation is rare [22] and at this frequency does not constitute a major problem.

The most powerful density fluctuations are not in fact caused by simple turbulent mixing of the ionosphere but are due to travelling ionospheric disturbances, or ‘TIDS’. These are large-scale quasi-sinusoidal variations in density that are thought to be passive tracers of vertical gravity waves in the neutral atmosphere beneath [33]. These gravity waves seem to be generated at high latitudes by sources such as magnetic storms, tropospheric turbulence or the local atmospheric heating caused by various auroral phenomena [46]. Although periods vary, (the largest scale TIDs, with $\lambda \geq 1000$ km, can have periods of up to 3 hours), they are characteristically similar to the natural buoyancy frequency of the neutral atmosphere – about 20 minutes, and scale-lengths between 100 to 200 km are usually present. At mid-latitudes the ionosphere has an irregular structure 90% of the time [45] due to strong anisotropies in the generation and propagation of the TIDs from higher latitudes, but the irregularities retain a definite sinusoidal nature.

The scale-lengths of TIDs are large compared to the baselines used for these observations, so their effect on the interferometric phase scales with the baseline length. If the typical change in electron density caused by a TID is taken to be 5% [101] and the mean dispersion measure of the atmosphere to be $5 \times 10^{16}\text{m}^{-2}$ (a typical nighttime figure), a sinusoidal TID of frequency ω_{tid} and wavelength λ_{tid} would be detected on a baseline d as a phase variation in the signal from a source at the zenith of about

$$\phi_{\text{tid}} \simeq 26 \frac{d}{\lambda_{\text{tid}}} \cos\alpha \sin\omega_{\text{tid}}t, \quad (6.3)$$

where α is the angle between the TID wavevector and the baseline. The ionosphere can therefore be expected to contribute $\sim 10^\circ$ of phase per km of baseline, a figure in good agreement with these observations.

The correlator determines the phases of signals at a number of radio frequencies within the band. The above ionospheric phase effects scale with the wavelength of the radiation, so the *differential* phase across the band is proportional to the *total* ionospheric component to the measured phases, if there are no other frequency dependent effects. Such multi-frequency measurements can, in principle, be used to calculate the ionospheric phase and correct for it. However, to obtain a figure to within 10° , over a fractional bandwidth of $1/100$, the differential phase must be measured to within 0.1° , requiring a signal-to-noise ratio per frequency channel ~ 600 . Correction by this method is therefore impractical on all but the shortest baselines, where the ionospheric component is small anyway. As no other reliable ionospheric phase correction method was available it was thought best to leave the phase corrupted rather than try and derive dubious correction factors.

6.4 Data reduction

Data was collected on three different frequency channels between each of the two Yagis at the remote site and the 4C antenna. To maximise the system signal-to-noise ratio it was important to combine these datasets coherently. This coherent addition could be divided into two distinct stages – alignment of the frequency channels and alignment of the data from the two antennas on the mobile station.

6.4.1 Combining the frequency channels

By far the most important contribution to the differential phase between frequency channels was due to the drifting local oscillators. Once this drift had been compensated for, using the referencing technique described in Chapter 5, the phase difference between channels changed little with time if there was good position determination for the mobile station. However, the channels did have unknown phase offsets which had to be determined prior to coherent addition. This was done by allowing the n phase offsets, α_k , to vary as free parameters and to maximise

$$I(\alpha_1, \alpha_2 \dots \alpha_n) = \int_0^\tau \left| \sum_{k=1}^n V_k(t) e^{i\alpha_k} \right|^2 dt, \quad (6.4)$$

where $V_k(t)$ is the visibility on the k th frequency channel and the integration takes place over the coherent correlation period (3 hours). This maximisation was done iteratively by aligning channels in pairs (a process for which an analytic expression exists) until the residuals in α_k were insignificant ($< 1^\circ$), and so extracted the

maximum signal possible from the data, given these free parameters. The number of degrees of freedom (2) was small compared with the number of data points (~ 500) so any overestimation of the signal due to alignment of noise was assumed to be negligible.

This procedure was carried out independently on the data from each Yagi before further processing. Both the design of equipment and the accuracy in baseline determination only allowed phases to be measured within constant unknown offset, so such offsets could be freely introduced or removed throughout the data analysis.

6.4.2 Yagi alignment

Once the frequency channels had been combined as above, the data consisted of single visibility measurements from each of the two Yagi antennas fixed to the motorcaravan. These antennas were positioned some 3.5 m apart ($\sim 1\lambda$), so the phase difference between them amounted to well over 100° during a 6 hour observation of Cas A, and the nature of this phase difference depended critically on the orientation of the van. Rather than rely on map and compass measurements, a maximisation approach similar to the one above was employed, but with the van orientation as the freely varying parameter.

If the vector connecting the two Yagis has magnitude a and makes an angle β with the local E-W line (i.e. the line perpendicular to the plane of the local meridian), the phase difference between signals reaching the antennas from a point source at hour angle H and dec δ is

$$\begin{aligned} \phi(\beta, H) = & -\frac{2\pi}{\lambda} a \cos \beta \cos \delta \sin(H + \Lambda) \\ & + \frac{2\pi}{\lambda} a \sin \beta (\cos l \sin \delta - \sin l \cos \delta \cos(H + \Lambda)) , \end{aligned} \quad (6.5)$$

where (Λ, l) are the longitude and latitude of the station. Denoting the complex visibilities measured by the front and back antennas as V_F and V_B , the orientation of the van was taken as the value of β that maximises

$$I(\beta) = \int \left| V_F(H) + V_B(H) e^{i\phi(\beta, H)} \right|^2 dH . \quad (6.6)$$

The orientation was not the only variable that must be taken into account – there was also a constant phase offset, α , between the measurements – so the final maximisation was two dimensional, in α and β .

The surface defined by $I(\alpha, \beta)$ for data gathered at Cottenham is shown in Figure 6.2 for a range of $\pm 2\pi$ in both variables. The computational burden in finding

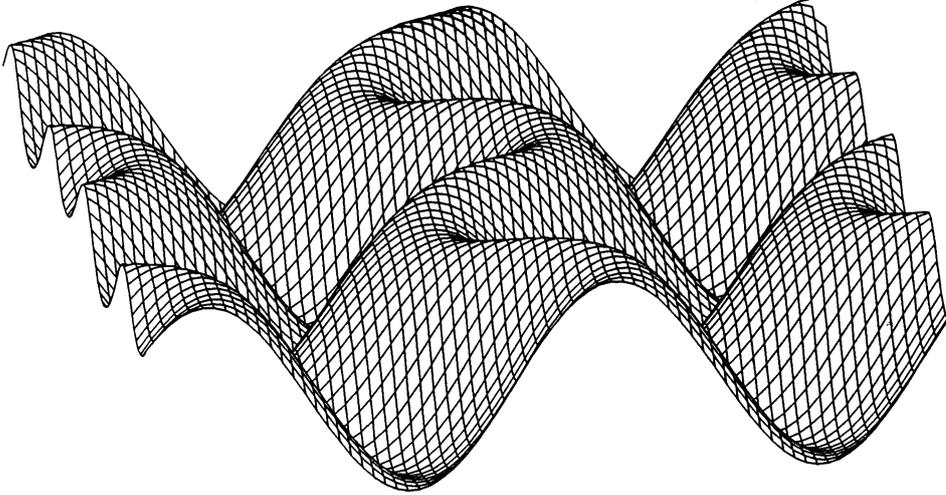


Figure 6.2: *The bearing/offset surface $I(\alpha, \beta)$ over the range $\pm 2\pi$ for the Cottenham data (Figure 6.11).*

the maximum of such a surface is quite small so a variant on the downhill simplex method was employed. As in the example shown, the maxima were sometimes quite flat, so the procedure was repeated from a number of starting positions to check for consistent results. The results could also be checked by comparing orientations derived on different days for the same baseline. These measurements agreed to within 5 or 10 degrees over all but the most distant baselines, for which map-derived orientations were used instead. A number of observations were usually made at each site, so the directions derived from each could be (vector) averaged to give an even better estimate of the van orientation.

Any error in the assumed position of either the 4C antenna or the mobile station changed the phase by an amount

$$\phi(H) = \mathcal{N}(g - b(H)) - \mathcal{E}c(H) , \quad (6.7)$$

where

$$\begin{aligned} g &= \frac{2\pi}{\lambda} \sin \delta \cos l \\ b(H) &= \frac{2\pi}{\lambda} \cos \delta \sin l \cos(H + \Lambda) \\ c(H) &= \frac{2\pi}{\lambda} \sin \delta \sin(H + \Lambda) \end{aligned} \quad (6.8)$$

and $(\mathcal{E}, \mathcal{N})$ is the error, in metres, East and North of the true position. On the short baselines Cas A is unresolved and the ionosphere is negligible, so this error could be determined directly from a χ^2 fit of Equation 6.7 to the tracked astronomical phase. Measurements of this sort on the four shortest baselines revealed a systematic error, attributed to the position of the 4C antenna, of -1.75 m East and -12.21 m North, on top of randomly distributed errors with standard deviation ~ 2 m. The small random component confirmed the accuracy of the map-based technique. The systematic error was allowed for on all baselines.

6.4.3 Amplitude calibration

One of the many advantages of a 1-bit digital correlator is that its output is simply related to the correlated power between its inputs. The true correlation coefficient, ρ , may be obtained from the correlation coefficient measured by a 1-bit system, ρ_2 , with the Van Vleck relation (see Chapter 5 and [99])

$$\rho = \sin\left(\frac{\pi}{2}\rho_2\right). \quad (6.9)$$

In turn, this coefficient is related to the antenna and system temperatures of the two stations by

$$\rho^2 = \frac{T_{A1}T_{A2}}{(T_{A1} + T_{S1})(T_{A2} + T_{S2})}, \quad (6.10)$$

where T_A is proportional to the component of the total received power that is correlated at both ends of the interferometer and T_S is proportional to that which is uncorrelated. For a particular patch of sky, the denominator of this expression is the same on all baselines, as it equals the product of the total powers received by the two stations. As both antenna temperatures are directly proportional to the (correlated) source flux density, S , ρ is proportional to S no matter how high the system or antenna temperatures. On the other hand, the signal-to-noise ratio, γ , in the product of the two antenna voltages is related to the antenna and system temperatures by

$$\gamma^2 = \frac{T_{A1}T_{A2}}{(T_{A1} + T_{S1})(T_{A2} + T_{S2}) + T_{A1}T_{A2}}. \quad (6.11)$$

Equation 6.11 only approaches 6.10 when the antenna temperatures are a lot less than the system temperatures (as is usually the case). When observing a strong source such as Cas A, the two temperatures are frequently comparable so any measurement based on signal-to-noise ratios would need careful calibration. Most systems attempt to measure the correlation coefficient directly, though with an analogue correlator the system gain must be carefully chosen both to retain sensitivity and

prevent saturation. A 1-bit digital correlator does this automatically, so that its output is directly proportional to the correlated flux density for all baselines, no matter how great the total power. The constant of proportionality depends on the sky brightness and so must be determined empirically for each observation.

As a direct result of the above, all the observations could be calibrated directly in terms of correlated flux density using a short spacing measurement and knowledge of the total flux density of the source. The response at zero baseline was inferred by extrapolating data from the shortest spacing back to the origin with an Airy function. The first zero of the function was well determined, by both the present observations and recent CLFST data, and known to be at 870λ . Using the model developed by Baars *et al.* [3], the integrated 81.5 MHz flux of Cas A in 1987 was estimated to be 16300 ± 700 Jy. Changes in the forward gain of the 4C antenna limited the accuracy of amplitude calibration to 5-10%.

The primary beam of the 4C antenna gave the data a broad envelope over hour angle and reduced the sensitivity at the extremities of the observations. Although this could be modelled by the techniques described in Appendix E, a better approach was to measure it directly on a short baseline on which the source is unresolved. This was done twice, the results averaged and a quartic fitted to the resulting amplitude. This gave an empirical beam envelope as a function of hour angle in radians (H) of

$$B(H) = \sum_{k=0}^4 a_k H^k, \quad (6.12)$$

where $a_0 = 1.117$, $a_1 = -0.200$, $a_2 = -0.709$, $a_3 = 0.321$ and $a_4 = 0.306$ (see Figure 6.3), with which all visibility measurements were normalised.

Although a number of observations were made on each baseline, daily changes in the ionosphere prevented their coherent addition. Instead the data were combined incoherently and the resulting amplitude adjusted for noise so that it represented an unbiased estimate of the true visibility modulus. By this stage the phases were of little interest, but they too were combined, by averaging their tracked values, to give a best estimate of the true phase.

6.5 Results

The processed visibility data from the thirteen inner spacings are presented in Figures 6.4 to 6.16 with the LBI data in ‘histogram’ form and fitted VLA data shown by the smooth lines (see below). The observations on each baseline are divided around transit, reflecting the 3 hour coherence time of the correlator. Amplitude calibration

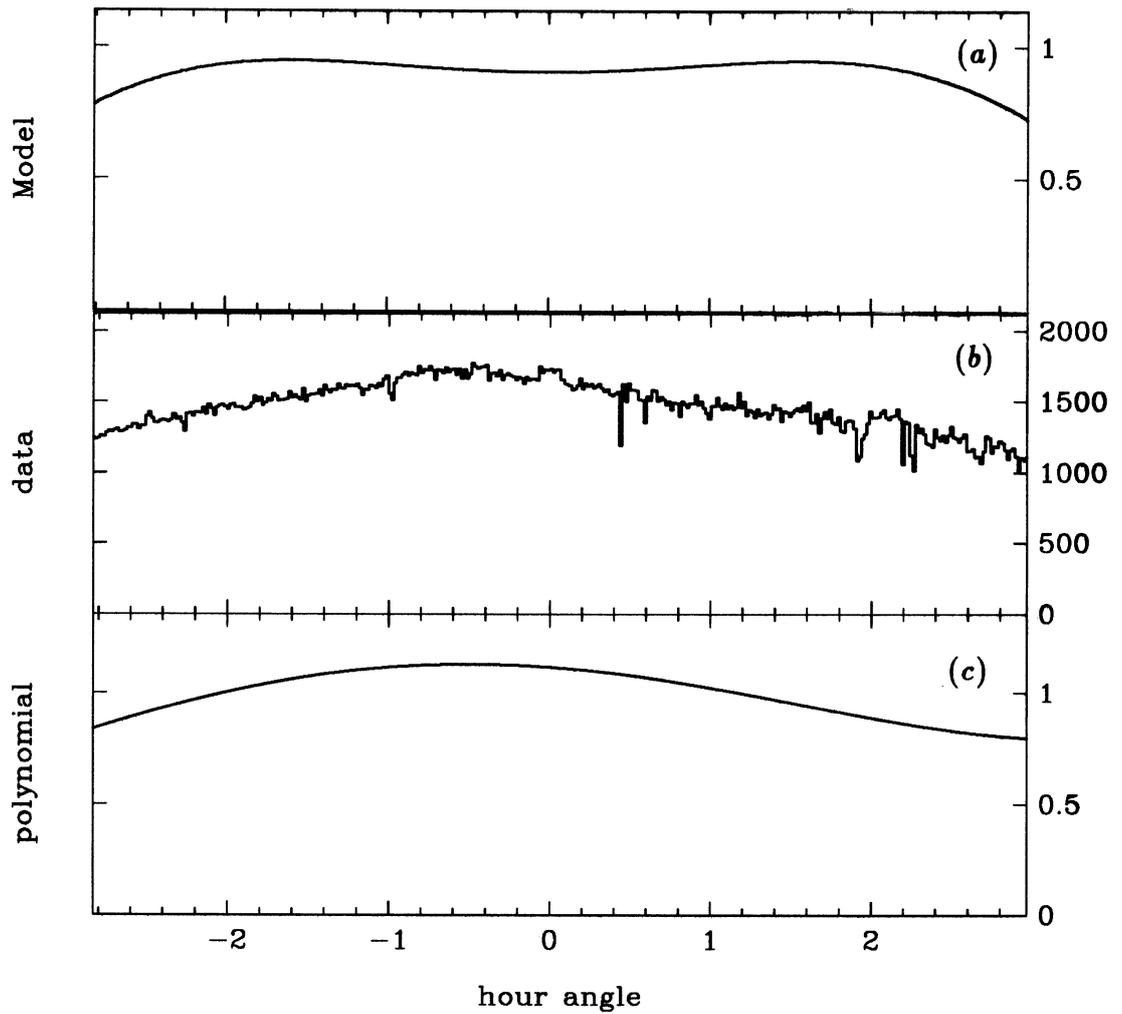


Figure 6.3: *The beam effects due to the 4C. (a) shows the theoretical response determined by similar methods to those of Appendix E. (b) shows the beam as measured on the shortest baseline and assuming Cas to be totally unresolved. (c) shows the 4th order polynomial fit to (b) used to calibrate all the amplitudes.*

is in janskys, as derived from the LBI data. The third baseline, at Barton, follows the trough of the first zero almost exactly. This can be clearly seen in Figure 6.17 which shows the inner few spacings superimposed on top of CLFST (u, v) plane amplitudes [37] determined in the summer of 1986. Both the Barton measurements and those from the other baselines are in good agreement with the 151 MHz data.

A better understanding of the LBI data can be obtained by comparing it with the 1024×1024 pixel VLA map of the remnant, made at 1381 MHz in August 1987 (Figure 6.18). The most explicit way to make this comparison is to predict, from the VLA map, what would have been observed on the baselines used had the remnant looked the same at both 81.5 MHz and 1381 MHz. The quickest method of inferring the visibility at some arbitrary point is to carry out an inverse Fast Fourier Transform on the map and de-grid in the transform plane. This process has been analysed in great detail by Tan [88] who shows the need for sophisticated de-gridding and grid-correction routines if spurious short-spacing features are to be avoided. The alternative method is to compute the ‘direct’ inverse Fourier transform of the map at the (u, v) points of interest, i.e.

$$V(u, v) \propto \sum_m \sum_n B_{mn} \exp \left[\frac{2\pi i}{\lambda} (\eta mu + \xi nv) \right], \quad (6.13)$$

where B_{mn} is the flux of the m th map pixel, (η, ξ) is the pixel size (in radians) and the sum is made over all pixels. This is computationally much more expensive, but avoids the de-gridding problem all together. The task can be reduced significantly by binning the map pixels to give an image with a resolution that is lower, but sufficient for the baselines used.

This direct transform method was used on a 128×128 map (pixel size of $3.2''$) to generate ‘mock’ data corresponding to all the LBI observations described above (maximum resolution $14''$). The long baseline measurement (Chapter 7) was modelled similarly, though its higher resolution meant that the whole 1024×1024 map had to be back-transformed. Unfortunately, the VLA map was made for epoch 1950.0 and therefore represents the remnant as rotated slightly with respect to its present orientation on the sky. However, the rotation is only a few seconds in hour angle and was therefore ignored.

Figures 6.4 to 6.16 show the VLA results overlaying the LBI observations. The VLA amplitudes were individually scaled to minimise their misfit with the LBI data, and the LBI phases were tracked and fitted to the VLA phases with an offset before being folded back to their principle range. The spectral indices inferred from the

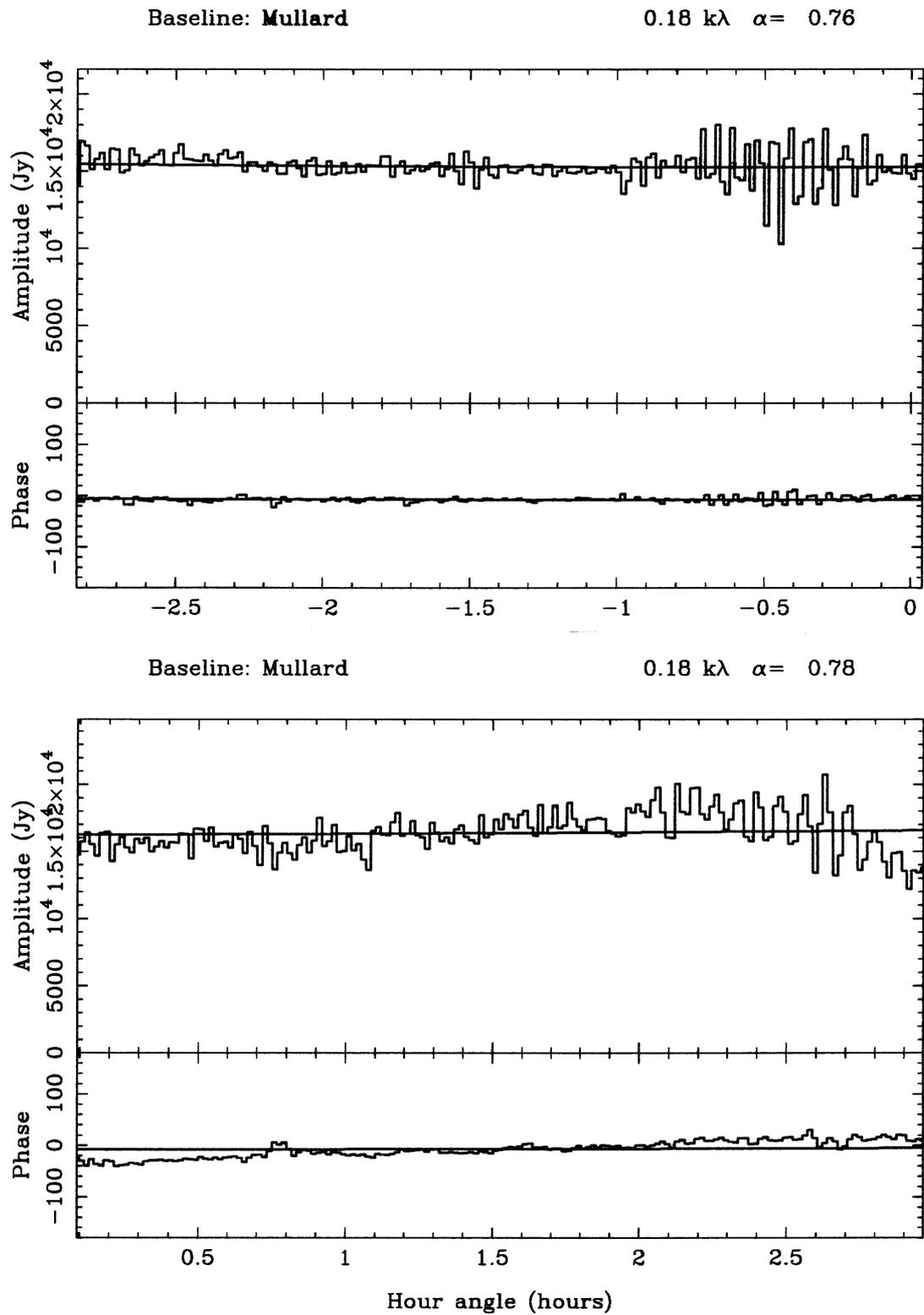


Figure 6.4: Baseline 1: Mullard. Most of the amplitude crinkles are due to interference. Note that the phase is well determined.

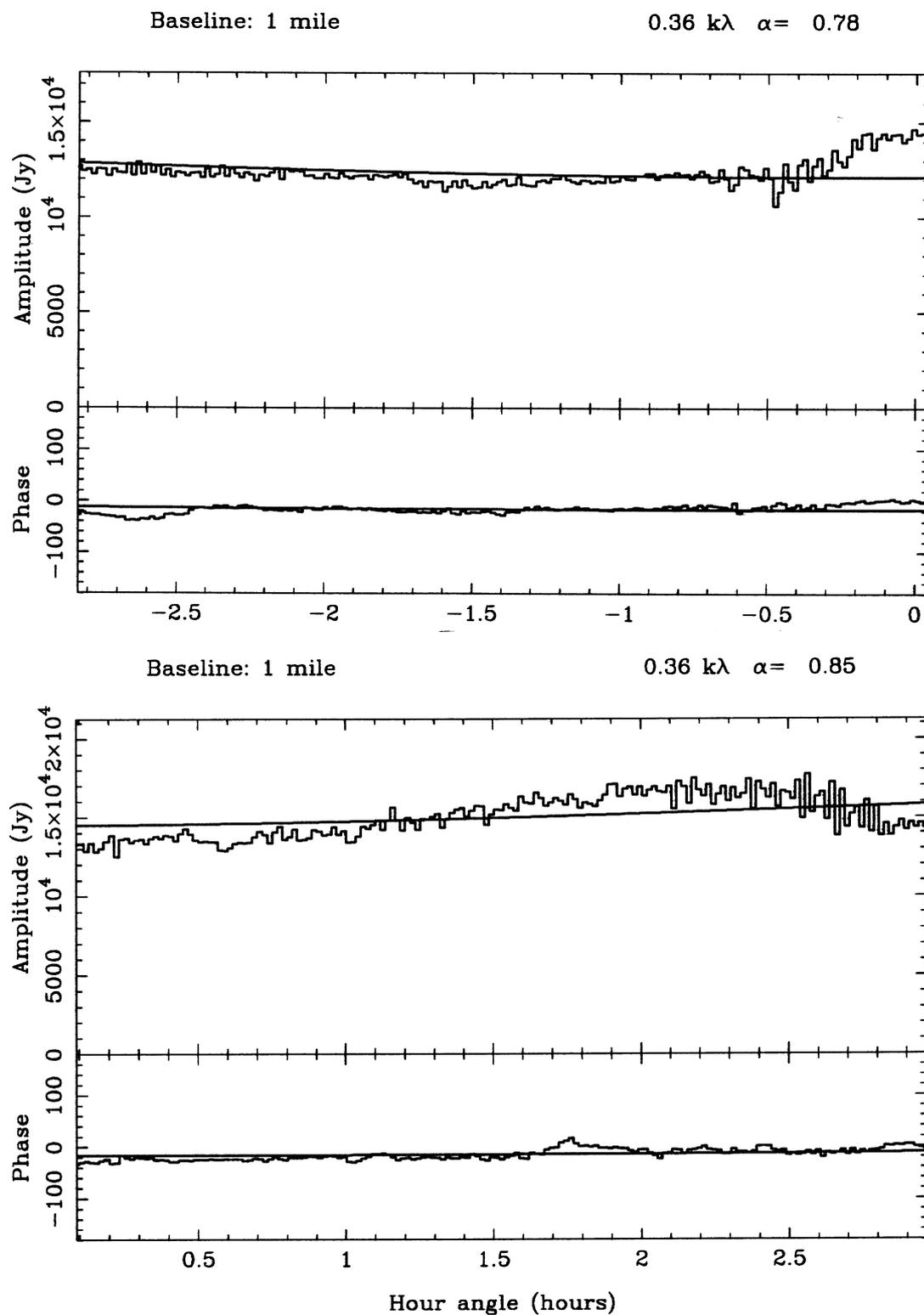


Figure 6.5: *Baseline 2: One Mile. Again, the phase is well determined.*

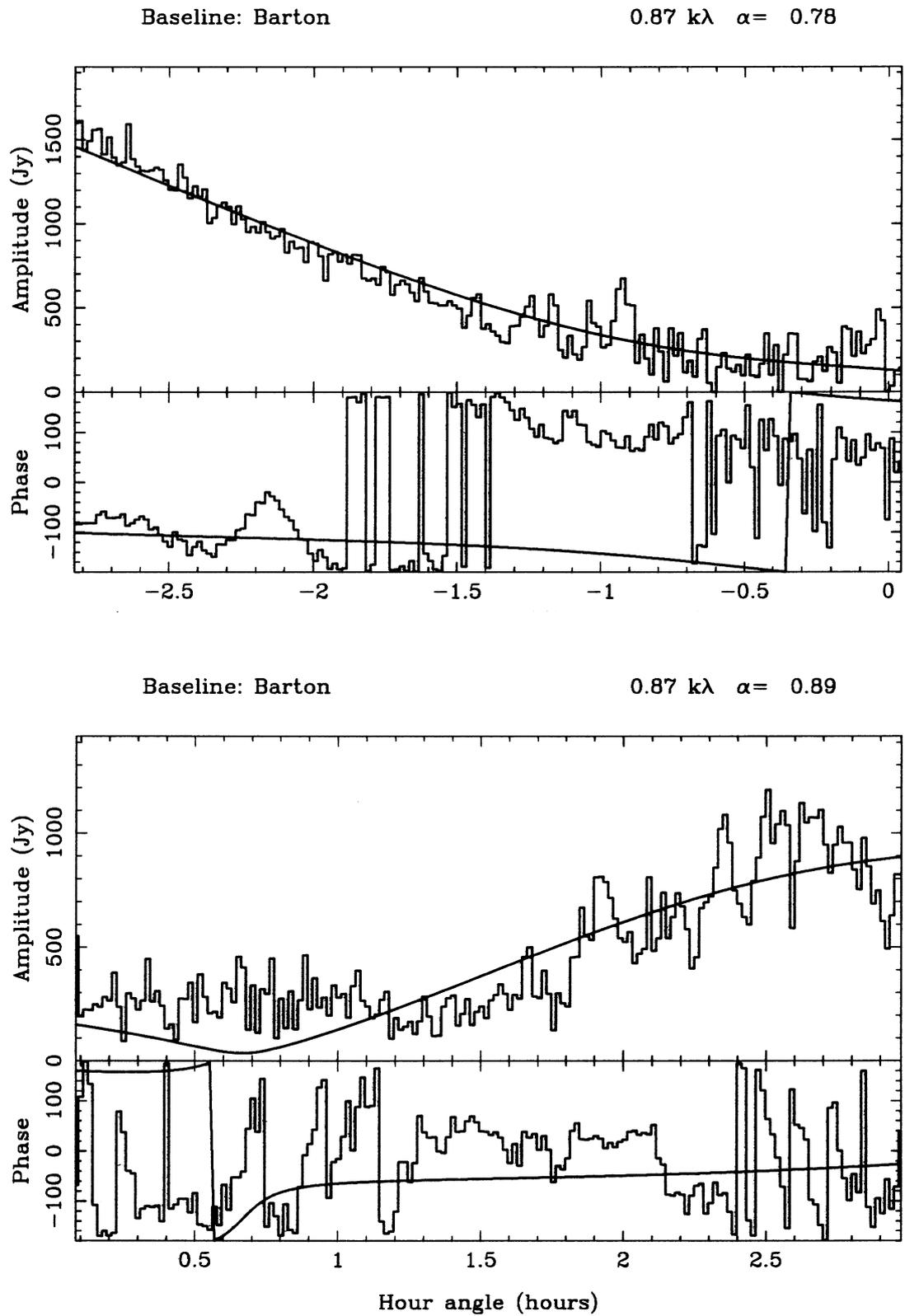


Figure 6.6: *Baseline 3: Barton.* This aperture track passes almost directly along the first zero in the transform plane.

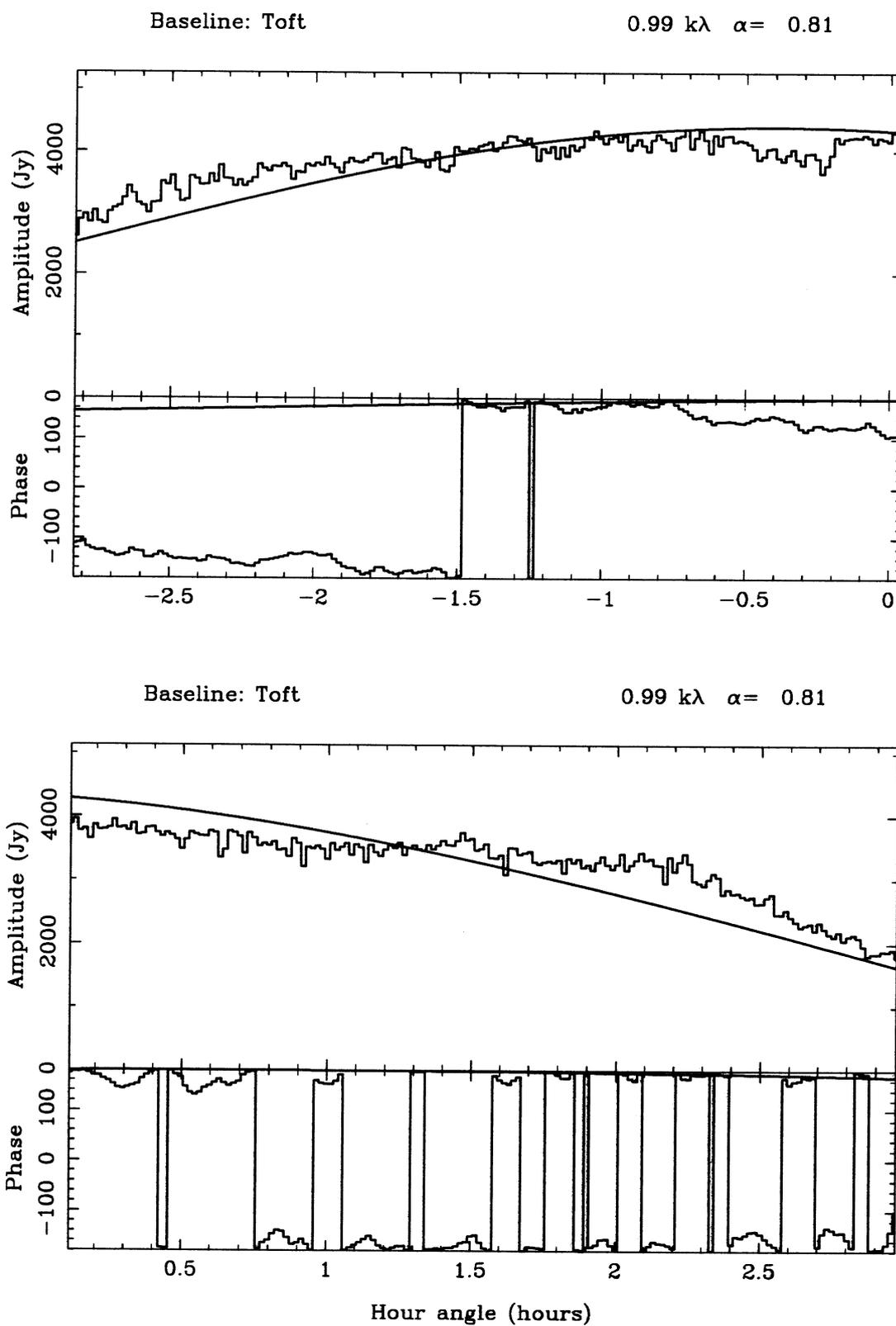


Figure 6.7: Baseline 4: Toft.

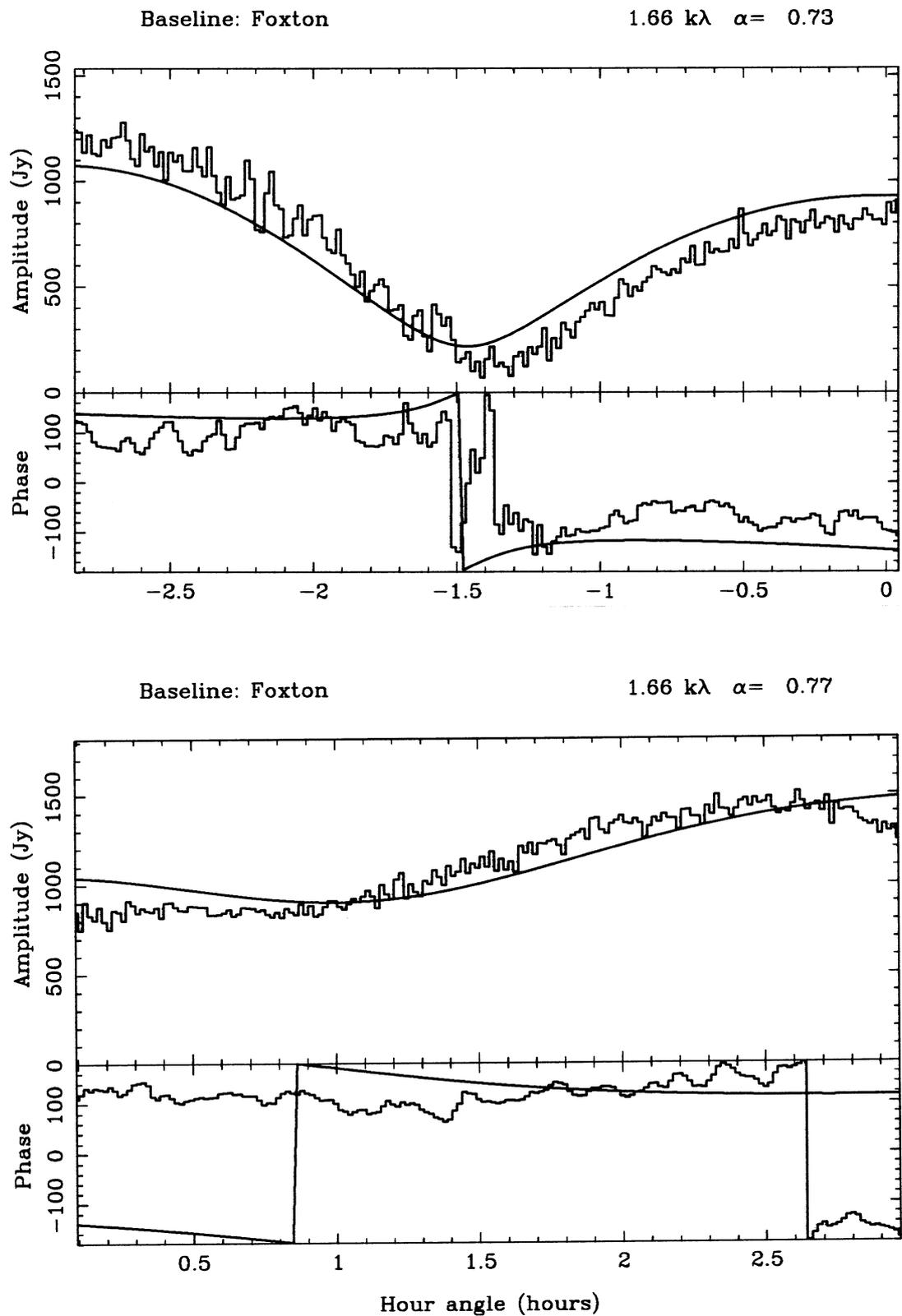


Figure 6.8: Baseline 5: Foxton. Although the LBI phase is beginning to be affected by the ionosphere, it is clearly astronomical.

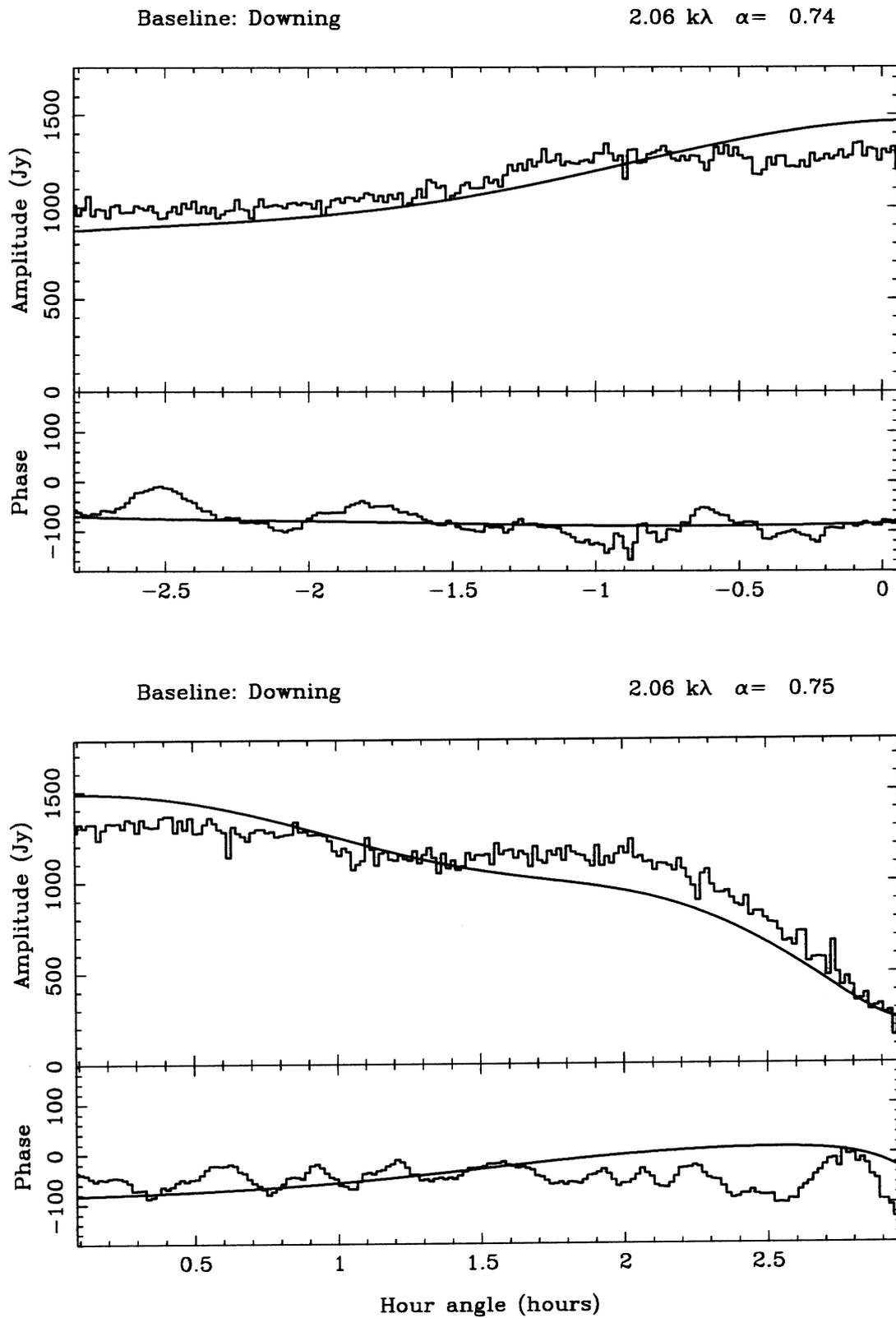


Figure 6.9: Baseline 6: Downing.

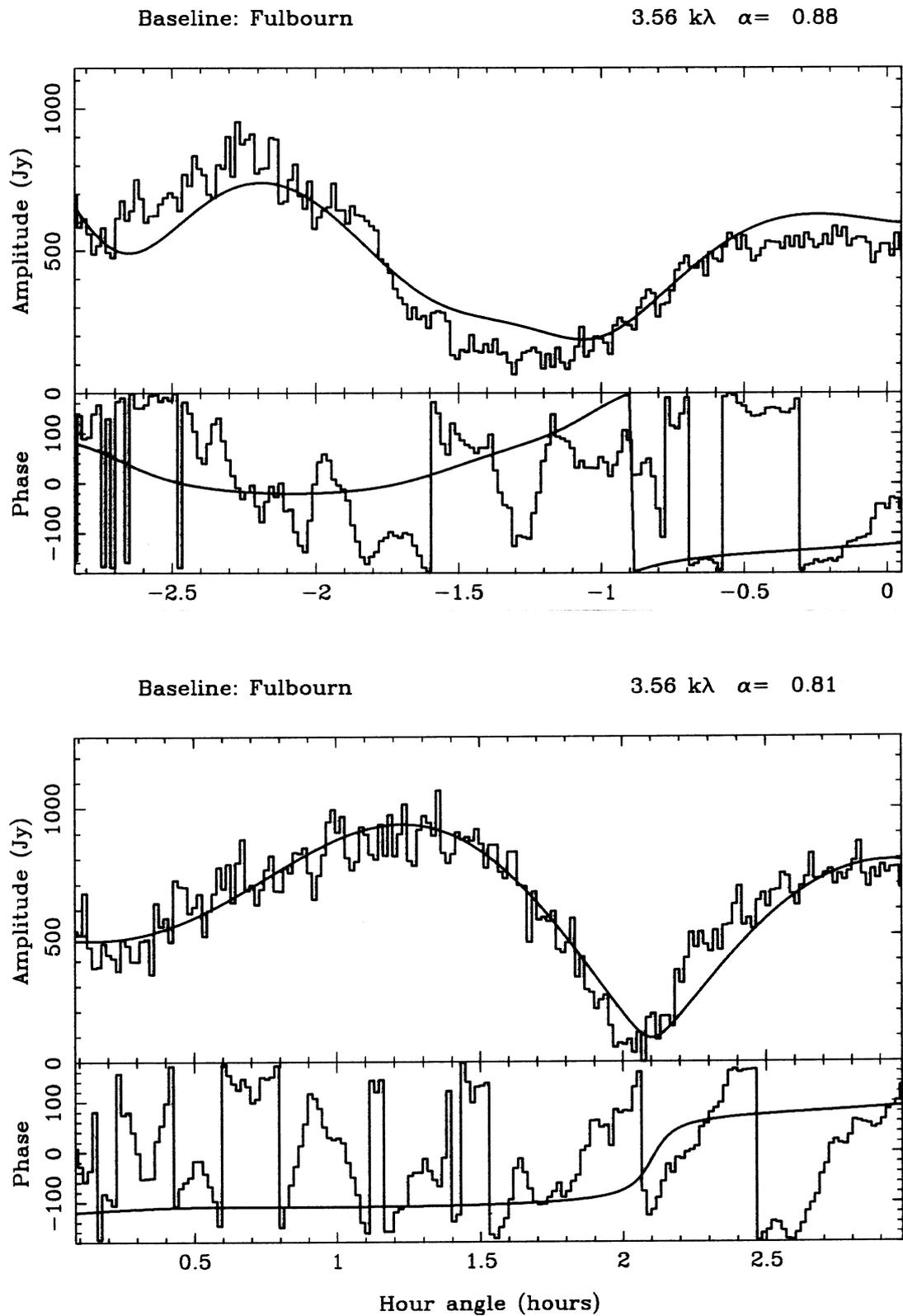


Figure 6.10: Baseline 7: Fulbourn.

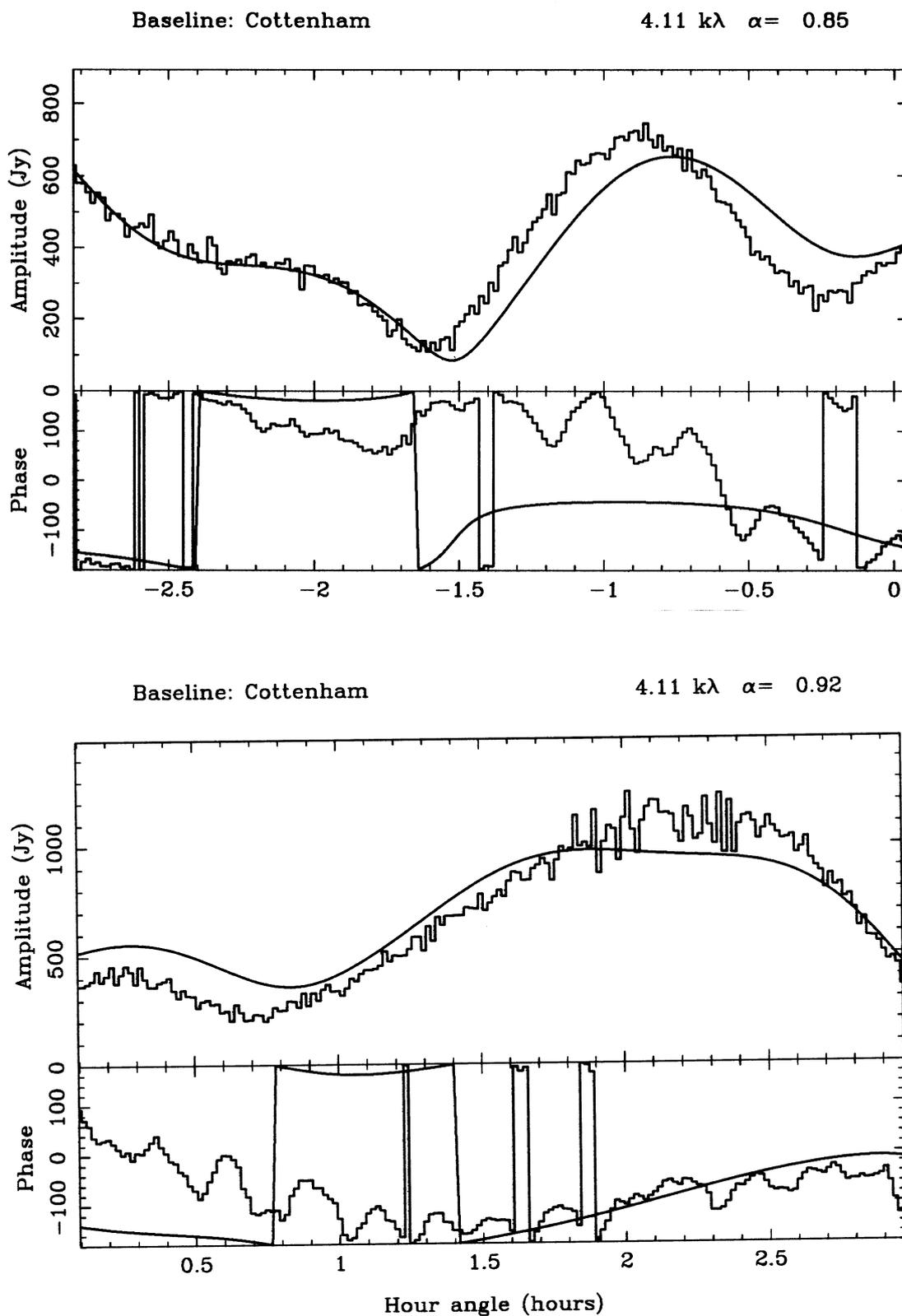
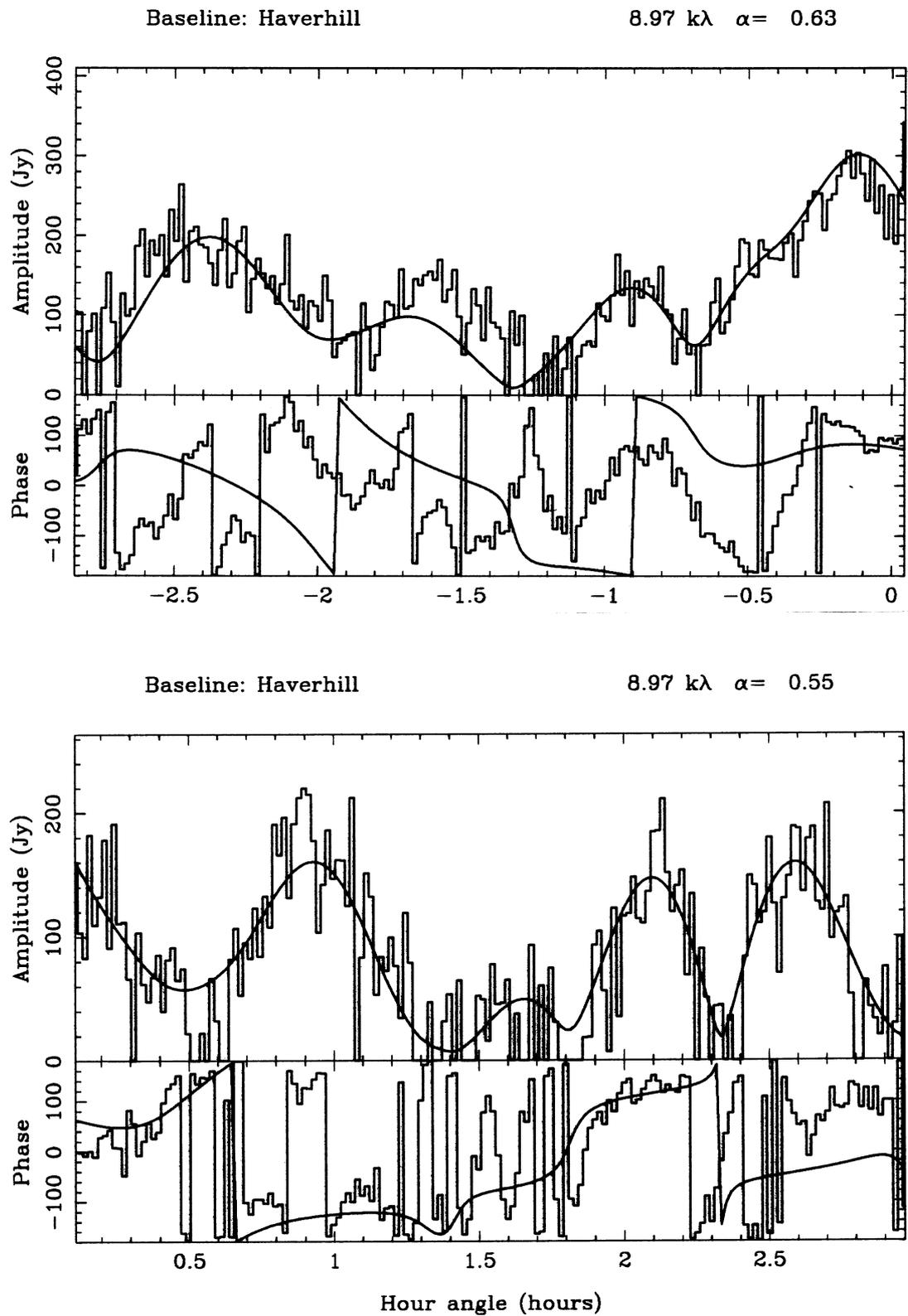


Figure 6.11: Baseline 8: Cottenham.

Figure 6.12: *Baseline 9: Haverhill.*

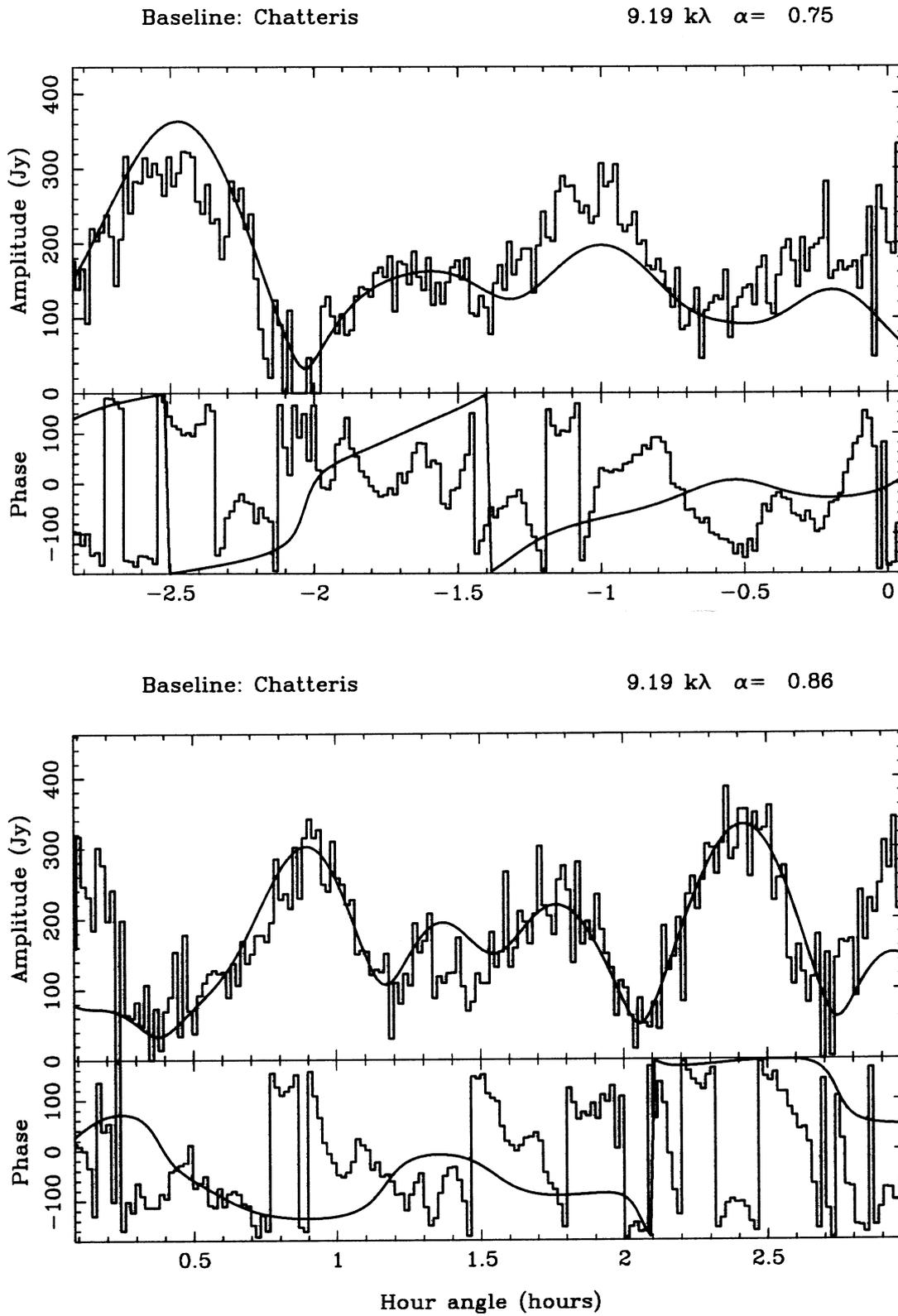


Figure 6.13: Baseline 10: Chatteris.

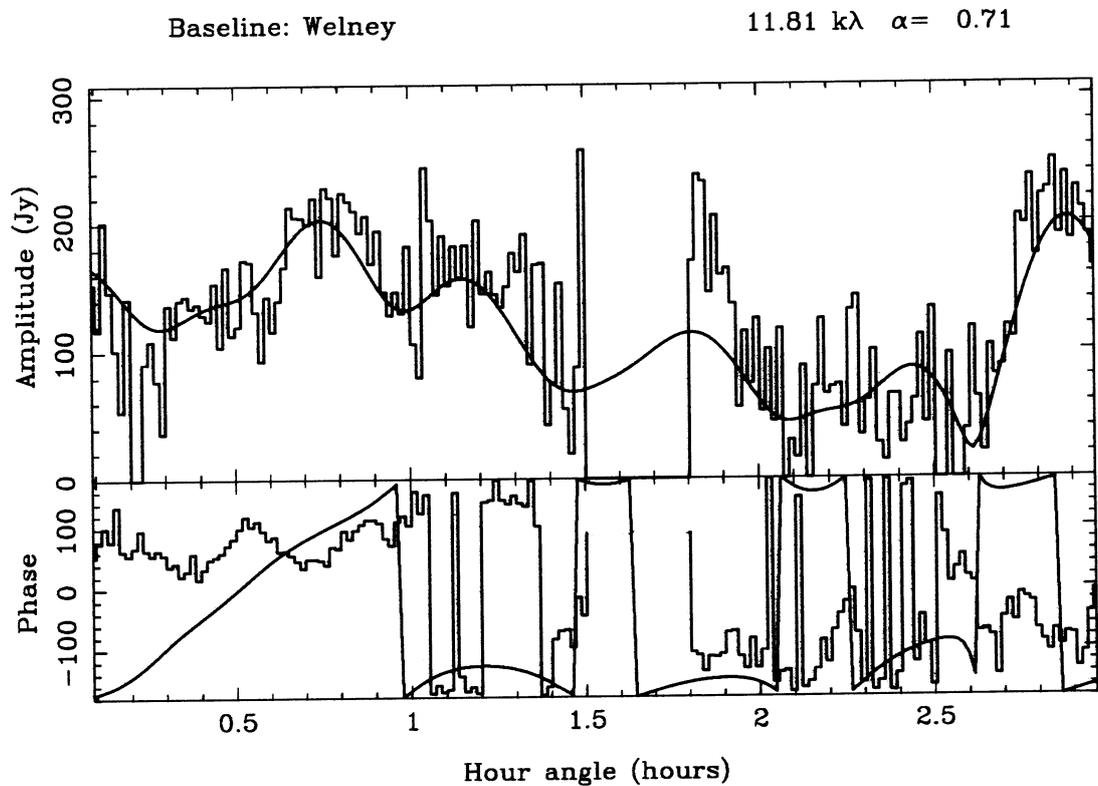
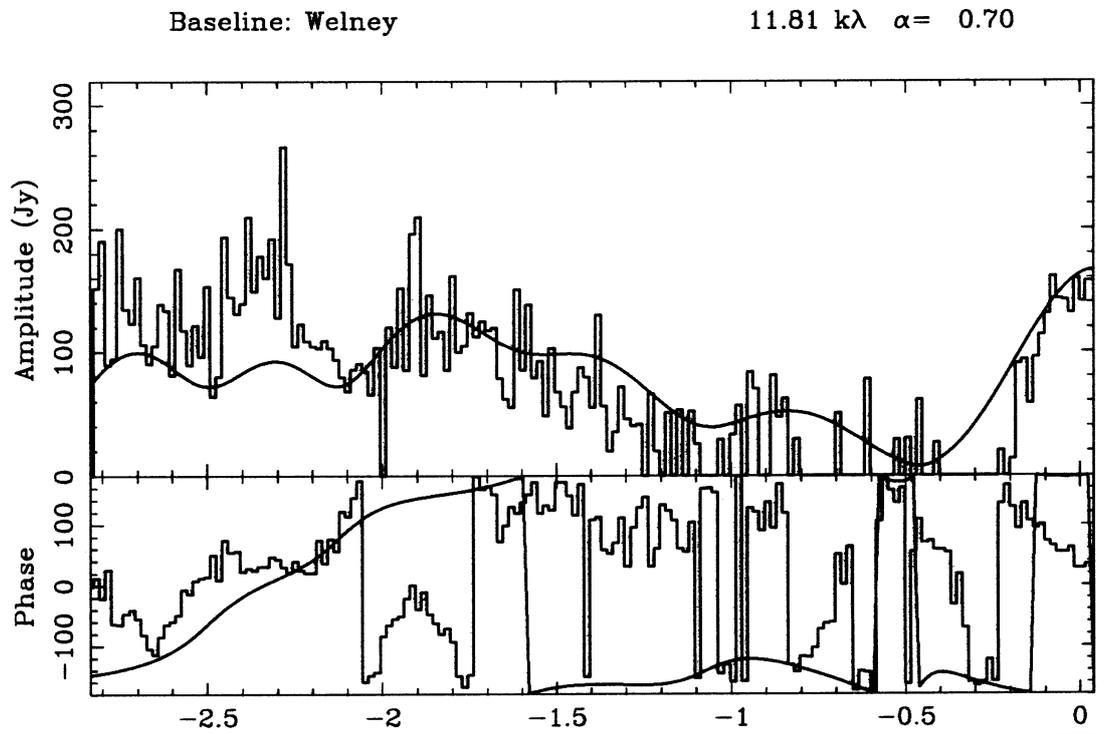


Figure 6.14: Baseline 11: Welney. The fringe rate of Cygnus and Cas were identical at about hour angle 1.65 hrs, so a section of the record was removed.

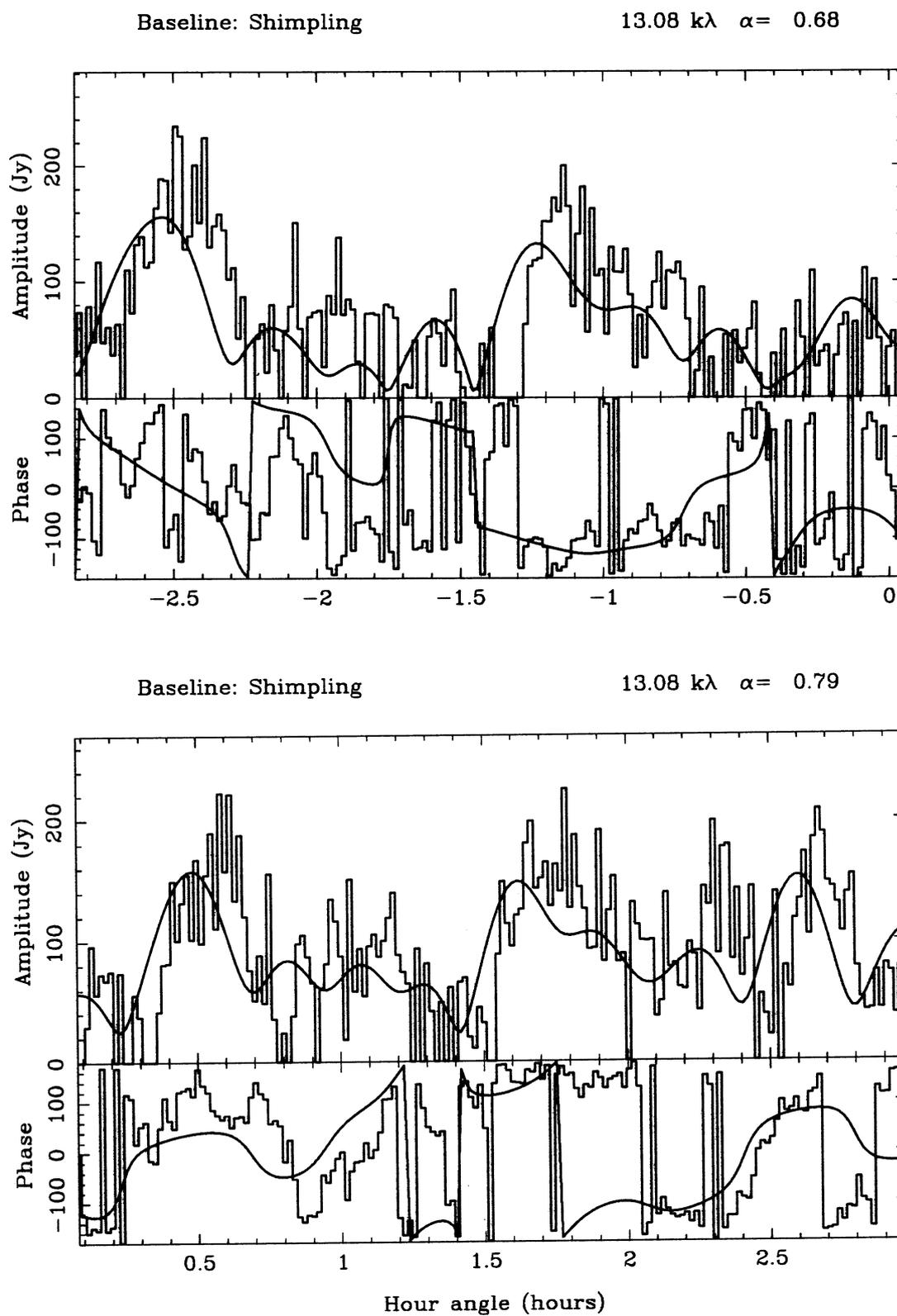


Figure 6.15: Baseline 12: Shimpling.

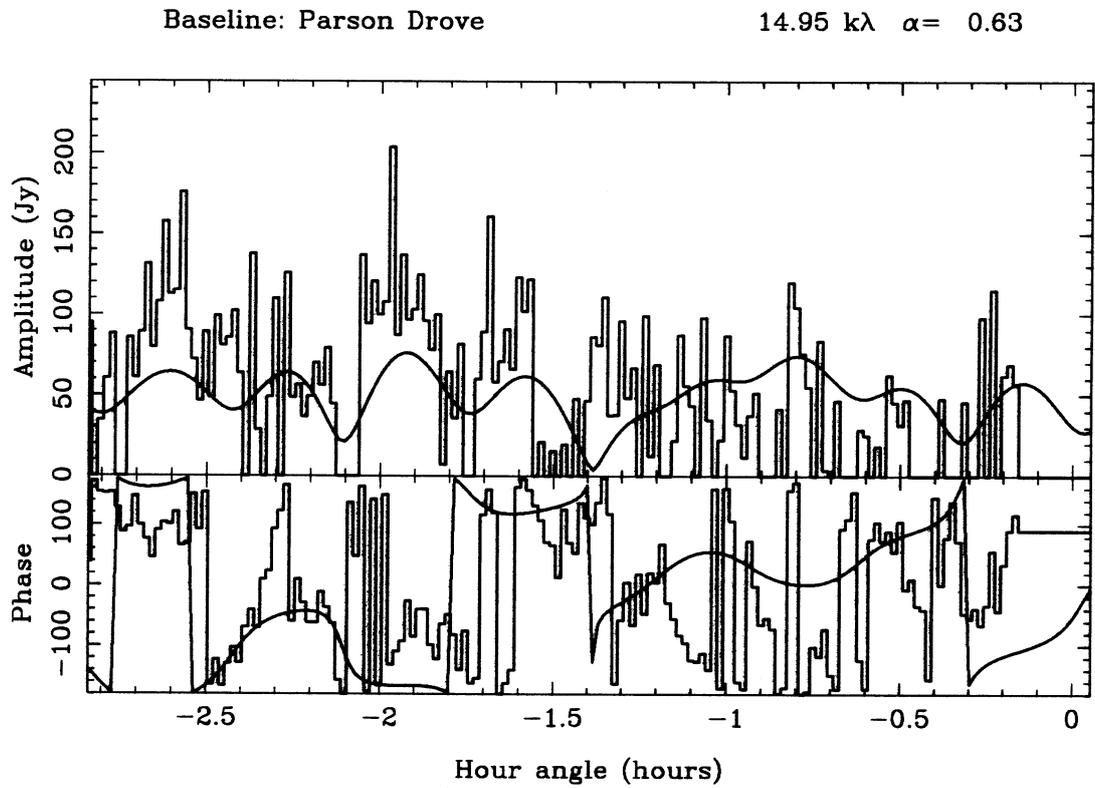


Figure 6.16: *Baseline 13: Parson Drove. The three hours after transit were lost.*

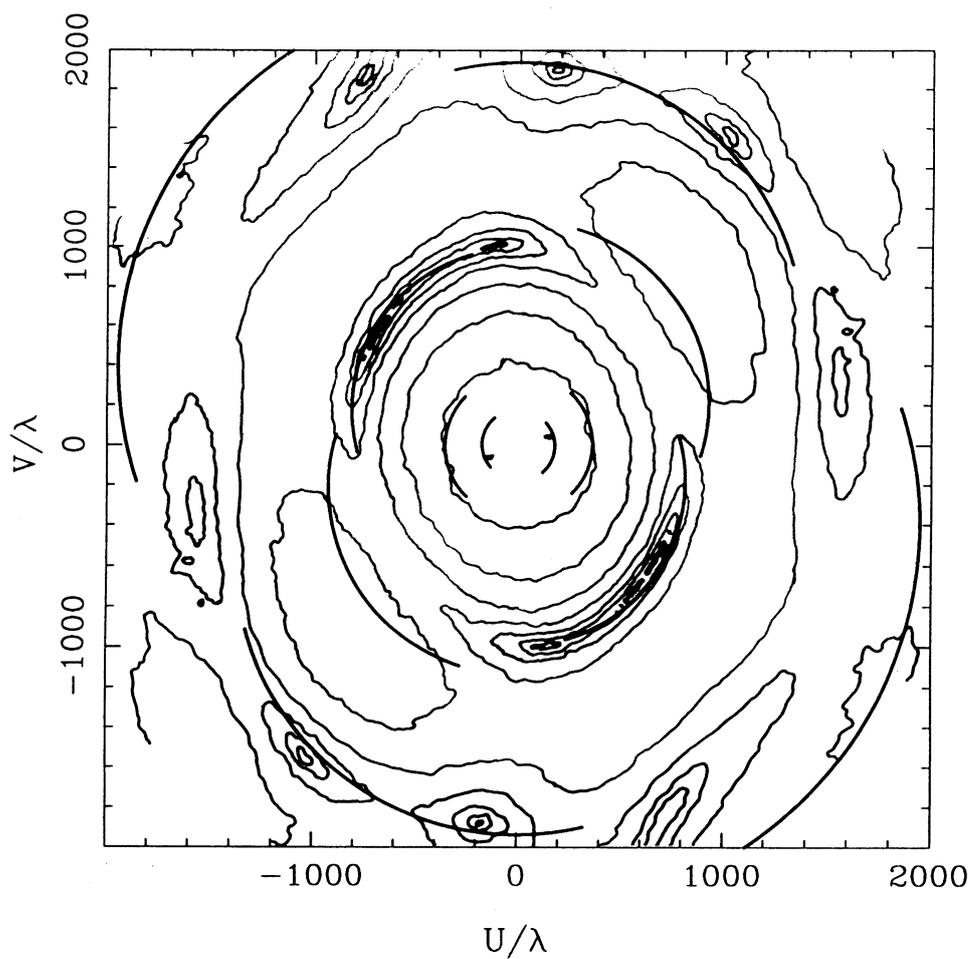


Figure 6.17: Contour map of the aerial co-ordinate aperture plane, as measured at 151 MHz [37], overlaying the inner six aperture tracks used for the new 81.5 MHz observations (solid curves). The third track (Barton) can be seen to pass very close to the first zero. Hour angle is increasing clockwise.

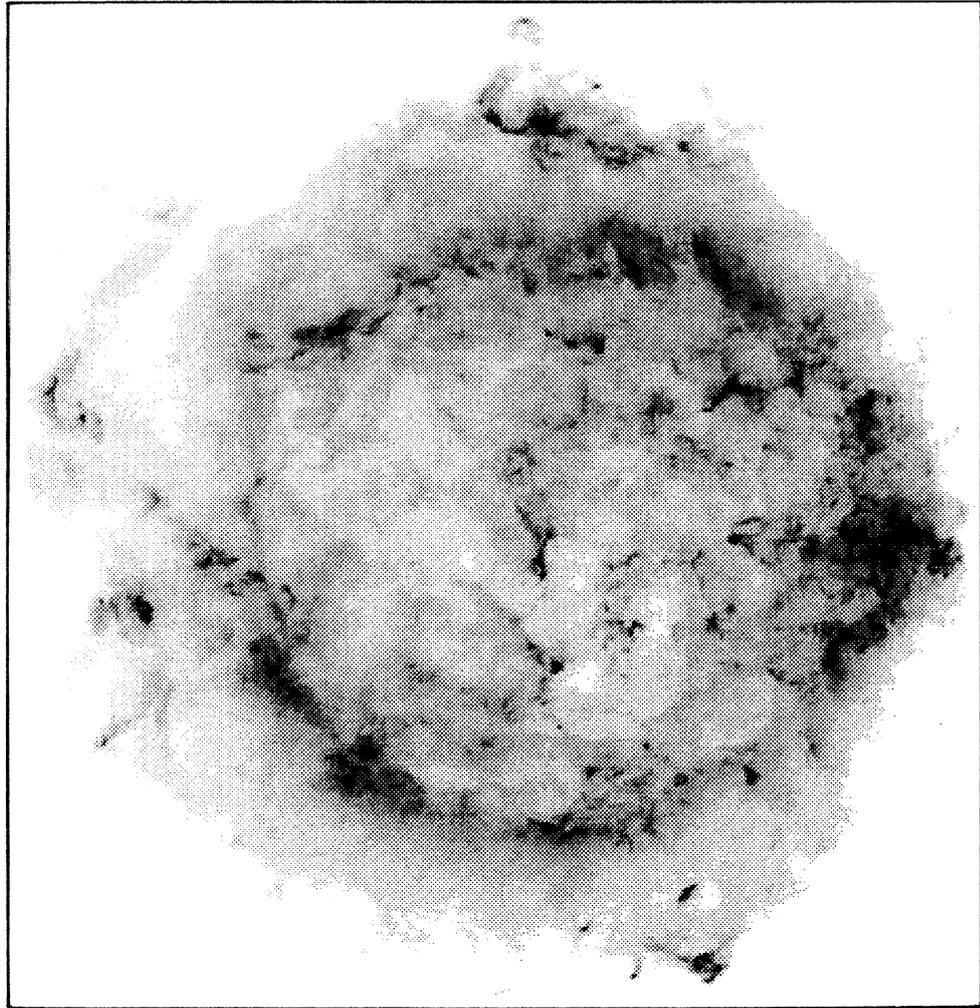


Figure 6.18: Grey-scale image of Cas A generated by the VLA at 1381 MHz. The true pixel-size is $0.4'' \times 0.4''$.

amplitude-fitting process are also shown. Gain uncertainties in the LBI data mean the values of α have an intrinsic error of ~ 0.05 .

The fit between the LBI and mock-VLA data is fairly good on most baselines. The phases on the shorter spacings (Figures 6.4 to 6.9) correlate particularly well. The phase of Figure 6.8(a) is clearly astronomical, though on this baseline, and beyond it, the ionosphere begins to dominate. On longer baselines (e.g. Figure 6.11) quasi-periodic TIDs become apparent in the phase, and the correlation between the LBI and mock-VLA data suffers accordingly. Although there is no significant trend in spectral index with baseline, there is a significant spread in the results – larger than would be expected from calibration errors. This is mainly because there are small but important differences between the mock and measured amplitudes on many of the baselines (e.g. Figures 6.8 and 6.11) which are sufficient to distort the fitted flux densities. Such differences should be interpreted with care, as they appear very similar to the effects of de-gridding errors. However the direct transform method should be free of such effects, so, assuming no errors in the geometry of the baselines, the discrepancy must indeed be between the map and the LBI data. In its ‘D’ configuration, the VLA has a minimum baseline of $\sim 180\lambda$ at 20 cm. This is very similar to the shortest LBI baseline and only about a fifth of the way out to the first minimum of the map transform, so it is unlikely that the VLA map misrepresents the extended structure of Cas. It is therefore reasonable to conclude that the differences are due to true spectral features rather than to artifacts of the processing.

With no real phase information, it is unclear how these changes in amplitude manifest themselves in the map plane. However, the differences in the remnant at high and low frequencies must be small, as the measured amplitudes are very similar. One approach is therefore to take the high frequency map as an initial model in a Maximum Entropy reconstruction constrained by the low frequency amplitude measurements. Constraints on the transform modulus are not linear functions of the image, so the entropy surface has many local maxima which would usually prevent the algorithm from selecting the ‘best’ image consistent with the data [57]. In this case though, the entropy can be assumed to be very near to its global maximum, so that a meaningful result can be achieved with the amplitude constraint alone. A second approach is to assume that the structural differences can be determined by assigning phase information derived from the VLA map to the LBI amplitudes and making a ‘hybrid’ map. To keep sidelobe ambiguities to a minimum this hybrid map must be compared with a map made from the mock-VLA data (i.e. VLA data from precisely the same aperture tracks as were used for the LBI observations) rather than the full VLA image. This second alternative, although inferior in theory, is

considerably easier to implement and was therefore preferred in this case.

Two dirty maps, one generated from the hybrid data (81.5 MHz amplitudes and 1381 MHz phases) and one from the mock-VLA data were deconvolved from the synthesised beam using the Maximum Entropy Method of Gull and Skilling [39] (Figure 6.19). It was found that the interference present on the two shortest baselines significantly degraded the LBI image, so mock-VLA data were used on these baselines for both images. The difference in the images is therefore due only to the data on baselines 3 to 13 inclusive. Because a number of the shorter baselines were oriented East-West, the algorithm has difficulty in pinning down the extended North-South flux of the remnant and this is reflected in the spurious ‘wings’ on the processed image. This map, although recognisably Cas A, contains a good deal of structure that would vanish with better aperture coverage. However, most of this structure is common to the reconstructed VLA image. Figure 6.20 shows the result of scaling the two maps for a minimum misfit and subtracting one from the other. Two contour maps are shown, one highlighting features brighter at 81.5 MHz and the other at 1381 MHz. Both maps overlay a reference image of Cas and have contours in steps of 0.164 Jy/pixel at 81.5 MHz.

6.6 Conclusions

The 81.5 MHz data have, by this stage, undergone many processes whose effects are hard to predict, however some genuine spectral properties are apparent. Most strikingly, the ‘raw’ 81.5 MHz and 1381 MHz amplitude data are very well correlated, showing that any spectral index variations across the remnant are small. This is by no means a unique property of Cas A. To take two particularly well investigated examples, Klein *et al.* [49] have shown Tycho (3C10) to have an index of $\alpha = 0.61$ that is constant to $\sim 5\%$ over the remnant between 610 MHz and 10.7 GHz, and Erickson *et al.* [29] have shown the same to be true of IC443 (3C157) which exhibits no significant departures from a spatially constant spectral index of $\alpha = 0.36$ between 30 and 1420 MHz. These results contrast sharply with the distribution of spectral indices amongst the known SNRs, which range from ~ 0.3 to ~ 0.9 . Clearly the dominant factor influencing the radiation process varies widely from remnant to remnant, but hardly at all within each.

Two explanations for this seem particularly plausible. Firstly, the energy spectrum of the radiating electrons could be determined by properties of the interstellar environment into which the supernova expands (assumed locally constant), so that the spread in spectral indices reflects the patchy nature of the ISM. Secondly, the

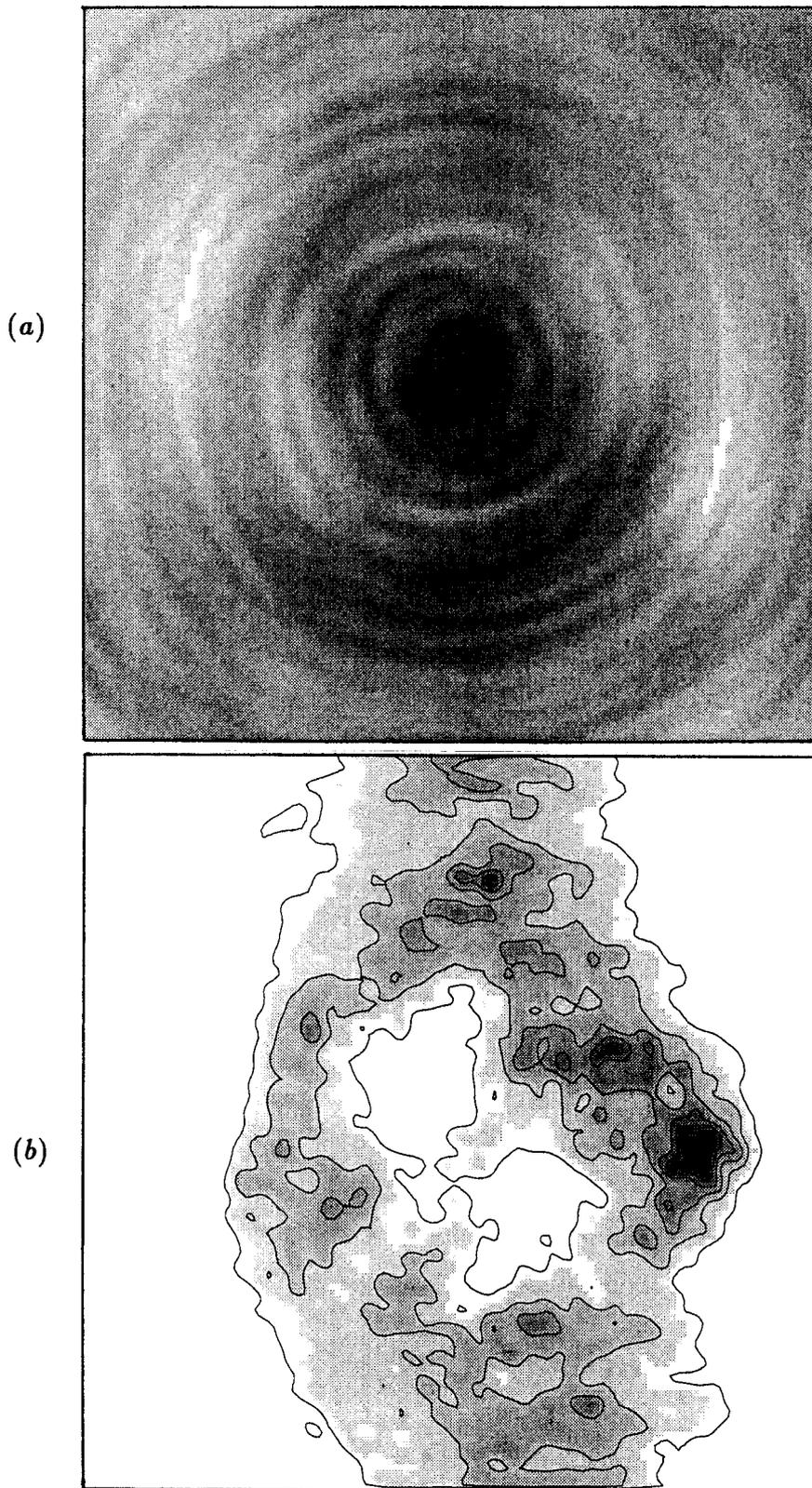


Figure 6.19: *Cas A* at 81.5 MHz. (a) shows the naturally weighted dirty beam and (b) the reconstructed image. Both maps are $410'' \times 410''$.

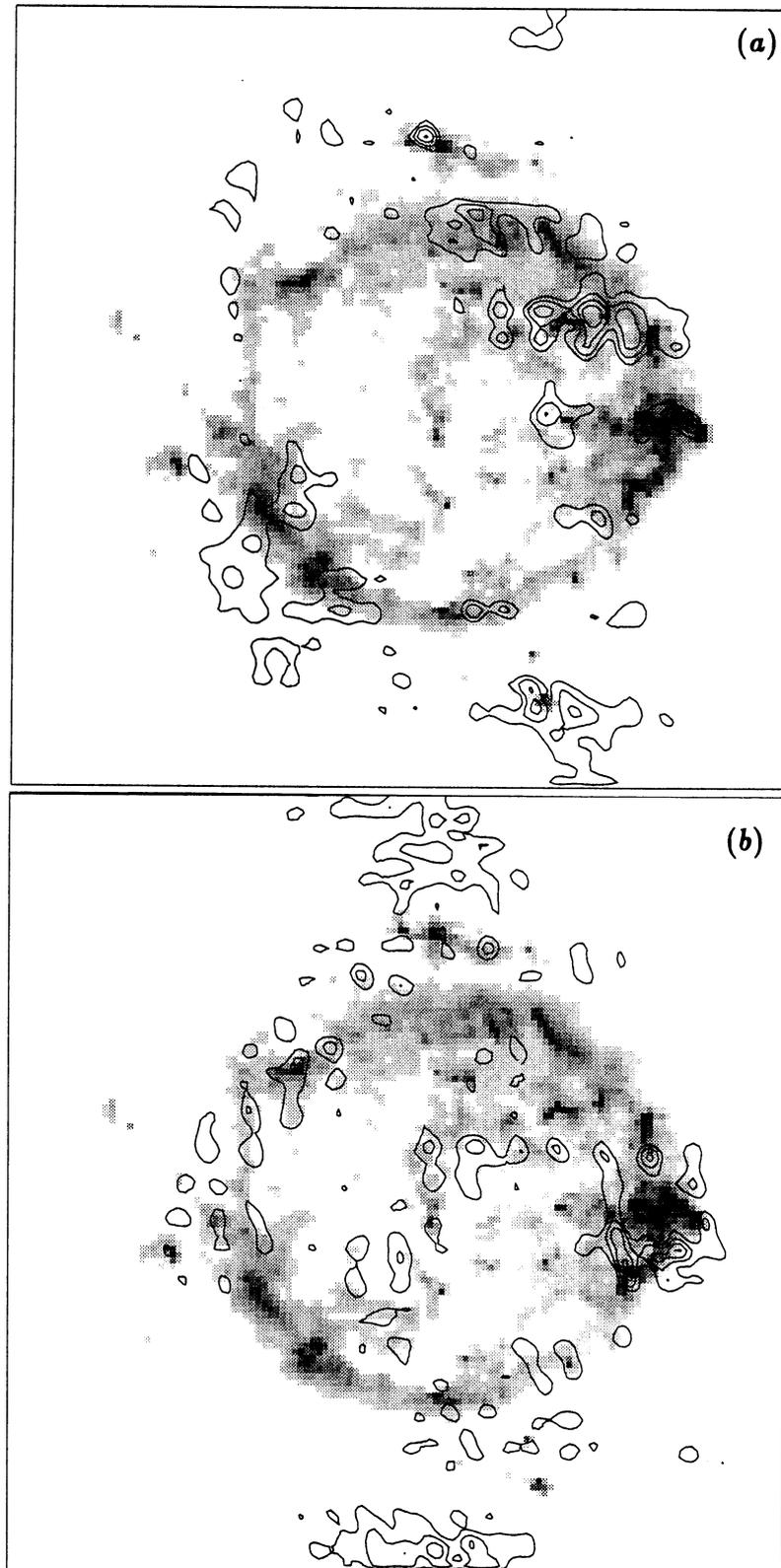


Figure 6.20: The residual differences between maps reconstructed at 81.5 MHz and 1381 MHz (contours) superimposed on full maps of the remnant (grey-scale). (a) and (b) show features brighter at low and high frequency respectively. Contours are 0.16, 0.33, 0.50 and 0.66 Jy/pixel.

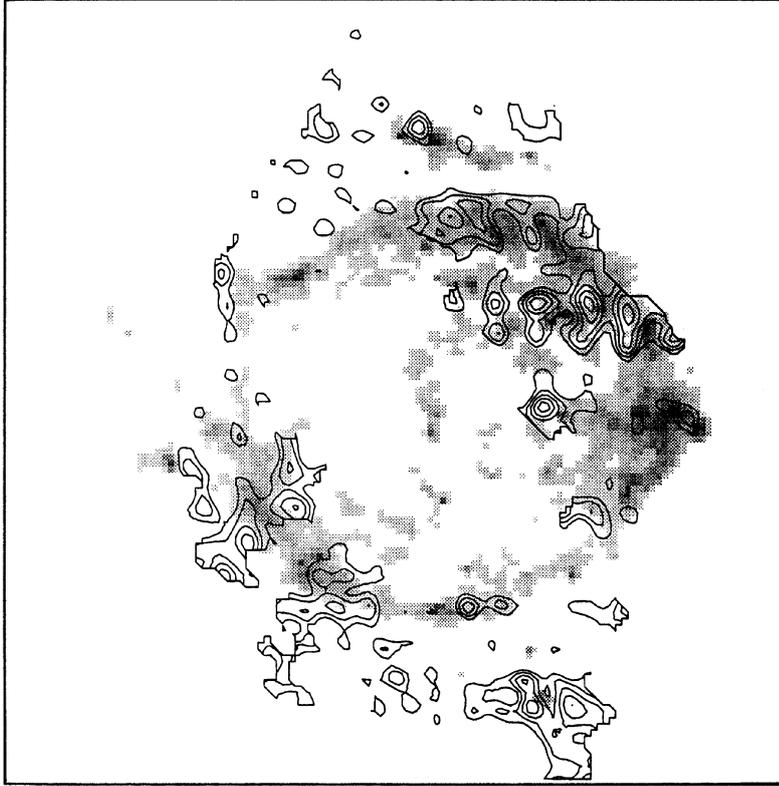


Figure 6.21: *Spectral index map of Cas A, only showing regions steeper than $\alpha = 0.77$. Contours are $\alpha = 0.79, 0.81, 0.83, 0.85, 0.87$. The grey-scale image is for reference.*

remnant's spectrum may age. The above examples (Tycho: young, steep spectrum; IC443: old, flat spectrum) suggest that spectra flatten with time. Although this idea is not supported by a good correlation between spectral index and surface brightness (assuming this to be a rough indicator of age) [36] it is the general conclusion reached by *some* numerical simulations, which predict that the energy spectrum of the radiating electrons should flatten as the shock front decelerates [21].

Figure 6.20 shows that the parts of Cas A that appear relatively brighter at low frequency tend to be associated with compact (presumably young) features, whereas the areas that seem brighter at high frequency are more randomly distributed. The effect is easier to see in the spectral index map (Figure 6.21) that only shows areas with relatively steep spectra. This is unlikely to be due to differences in the convergence of the MEM algorithm for the low and high frequency images as the effect is present over a wide range of MEM iterations. The least-square fitting procedure

used for Figure 6.20 has pulled up the mean flux of the VLA image to match the ‘spikier’ LBI image and this has made the N-S sidelobes appear brighter at the higher frequency. If the spectral index of the remnant as a whole is taken as 0.77, and assuming that residual beam effects in the data can be neglected, the low frequency brightening corresponds to a spectral index for the young sites of about 0.83. These changes in α over the remnant are in general agreement with those measured by Tuffs [94] between 2.7 and 5 GHz. Unfortunately the maps used for that comparison were of different epochs, so the results suffered from the effects of local secular brightness variations, however they did show the compact knots to vary in spectral index from ~ 0.65 to ~ 0.95 . This range in α is greater than that measured at 81.5 MHz and is probably due to the higher spacial resolution of Tuffs’ data. Perhaps more importantly, the steep spectral features measured by Tuffs correlate well with those measured at 81.5 MHz, indicating again that the spacial character of Cas A can be smoothly extrapolated to metre wavelengths. This last point is emphasized further in the following chapter.

Chapter 7

The compact structure of Cas A at metre wavelengths

7.1 Introduction

Although the previous chapter dealt with the structure of Cas A on scales $\gtrsim 15''$, some of the most puzzling observations of the remnant at low frequencies have concerned its structure on a much smaller scale ($\lesssim 4''$). These observations can conveniently be divided into four categories:

Total flux measurements A series of authors (Erickson & Perley, 1975 [28]; Read, 1977 [66, 67]; Walczowski & Smith, 1985 [102]) report seeing sudden increases in the ratio of the flux densities of Cas A to Cygnus A at frequencies at or below 38 MHz. Two ‘flares’, each with a flux density at 38 MHz of about 7000 Jy, have been detected over the last fifteen years, although no such changes have been observed at 151 MHz above the expected fading rate of $\sim 1\%$ per year. Read points out that if these flares are genuine they represent events occurring on a timescale of two to six years and that if their angular size is limited by their rate of expansion they must be smaller than $\sim 4''$.

L.B.I. Hutton *et al.* [47] made long baseline observations of Cas at 74 and 111 MHz, corresponding to baselines of 12000 and 18500 λ . By comparing their amplitude tracks with a map made at 2695 MHz some years earlier, they claimed to find a point source in their data, not present on the map, with a flux $\gtrsim 100$ Jy (scaled to 81.5 MHz) and a spectral index of ~ 2.5 . Later, but less sophisticated, observations made by Bovkoon *et al.* [12] with the URAN-1

interferometer at 20 and 25 MHz (baselines of 2840 and 3550 λ) appeared to confirm the findings of Hutton *et al.* They attributed unusually high measurements of visibility to some kind of unresolved feature with a spectral index of 2 and a flux of 2000 Jy at 20 MHz. Both sets of authors suggested that the feature was a pulsar.

Pulsar searches A number of attempts have been made to detect pulsed radiation from Cas A (e.g. Davies & Large, 1970 [24]; Seiradakis & Graham, 1980 [80]) but all have been unsuccessful. Although Cas A is a prime candidate for the site of a pulsar, none of these searches could resolve the intense shell of the remnant, so their sensitivities were low. Any pulsar within a beamwidth of the bulk emission might well have remained undetected.

IPS measurements The measured scintillation index of Cas A is significantly higher than might be expected under the assumption that its compact structure at low frequencies is the same as that mapped at higher frequencies (Tuffs and Duffett-Smith, unpublished, 1982). Although there are several possible interpretations, it suggests that either the compact structure is significantly brighter than the large scale structure at low frequencies or, that the remnant contains a pulsar. If sufficiently slow, such a pulsar would excite the scintillometer both through conventional scintillation and through its flash.

This chapter presents the results of a new series of observations which give a much clearer view of the compact structure of Cas A at 81.5 MHz. It is found that the structure is remarkably similar to what is seen at higher frequencies. Many of the measurements described above seem to be inconsistent with the present nature of the remnant and may well have been spurious.

Three main investigations are described. Firstly, sensitive long baseline observations, with a resolution of $\sim 6''$, are compared with the high resolution map of the remnant at 1381 MHz used in Chapter 6. These show there to be no steep spectrum unresolved component in the remnant brighter than a few janskys at 81.5 MHz. Secondly, the results of a deep pulsar search made at 408 MHz with a dish of the Cambridge One Mile Telescope are presented. These set a new upper limit on the power of any pulsar close to the remnant with a period, P , in the range $2 > P > 0.02$ sec. Thirdly, the IPS observations are discussed in the light of a better model of the interplanetary medium which predicts a higher scintillation index than the simple model used previously.

7.2 81.5 MHz LBI observations

Observations of Cas A were made in Autumn 1987 on a single baseline of deprojected length 128 km (resolution $6''$) in a manner very similar to that described in Chapter 6 (see Table 7.1). The simple two-Yagi mobile station employed for those observations did not have sufficient sensitivity on this baseline, so the portable phased array (Appendix B) was used instead. The array was erected in the grounds of Christ's Hospital School, near Horsham, and linearly polarised to give a total collecting area of $\sim 100 \text{ m}^2$. The differential Faraday rotation between this site and Cambridge has been measured [74] and found to exceed 45° only 10% of the time, so circular polarisation measurements were unnecessary. Successful observations (six hours each) were obtained on a total of thirteen nights and the data posted to Cambridge by boys from the school. The interferometer was phase-stabilised using signals from the Tacolneston transmitter. The propagation between this transmitter and the mobile station was often poor, but stability was usually maintained to better than a radian.

7.2.1 Data reduction

Data reduction was somewhat simpler for these observations than on the shorter spacings, as the antennas at the remote site did not need to be phased together by the reduction software. The calibrated local oscillators were sufficiently stable for coherent channel addition (see Section 6.4.1), but again the ionosphere prevented observations made on different days from being combined coherently. Instead, unbiased estimates of the amplitudes measured on the thirteen nights were averaged to give an improved estimate of the true visibility modulus. Unfortunately Cas A and Cygnus A had very similar geometrical delays about an hour after transit, giving a sharp spike in the visibility record. This also made it difficult to find the fringes due to Cas A in delay space and resulted in a drop in system sensitivity over the whole three hours following transit. However, the measurements made before transit were free from such effects, so the remaining analysis was performed on these alone.

The observations were calibrated with a short spacing observation of Cas A (0.96 km) on which the flux density was known. The flux density measured on the long baseline depended to some extent on the coherence of the frequency channels, but the calibration was certainly within 20%.

latitude	$51^{\circ} 2' 29.4''$
longitude	$0^{\circ} - 21' 57.2''$
height	56 m
hour angle	0.2772065 radians
declination	-0.6514571 radians
length	127836.4 m

Table 7.1: *Parameters of the Christ's Hospital baseline.*

7.2.2 Interpretation

As discussed in the previous chapter, the measured data can be compared with ‘mock’ data generated by a direct inverse Fourier transformation of the VLA map along an appropriate aperture track. The resolution of the VLA map ($0.4''$) is quite sufficient to accurately reproduce the region of the transform plane corresponding to the Christ’s Hospital measurement, and the direct transform ensures there are none of the degrading errors associated with faster methods.

A least-squares fit of the mock data to the calibrated LBI data is shown in Figure 7.1. The spectral index inferred from this fit is low (0.69) compared with the remnant taken as a whole, but consistent with the expected calibration error in the 81.5 MHz data. Although the spectral index is a poor measure of the difference in compact fluxes at the two frequencies (not least because the track represents a small range in hour angle) it does indicate that there are no compact features of more than, say, a few tens of janskys at 81.5 MHz that are not present at higher frequencies. What is more revealing is the overall goodness-of-fit between the two sets of data. The phase is, of course, badly corrupted, but the amplitude shows a remarkable degree of correlation. All the main features in the VLA data (except perhaps the peak at -1.4 hrs) are reproduced in the LBI data and in more or less the correct proportions. The signal-to-noise ratio in the LBI data is quite low (~ 5), but even without quantitative analysis, the data will clearly not allow a single, steep spectrum unresolved component in the remnant of more than a few janskys, as this would almost certainly shift the positions of the maxima and alter their relative heights.

Without knowledge of how many compact components are contributing to the visibility, little can be deduced about the statistics of their individual spectral indices from this single measurement. However, the data does allow strict limits to be placed

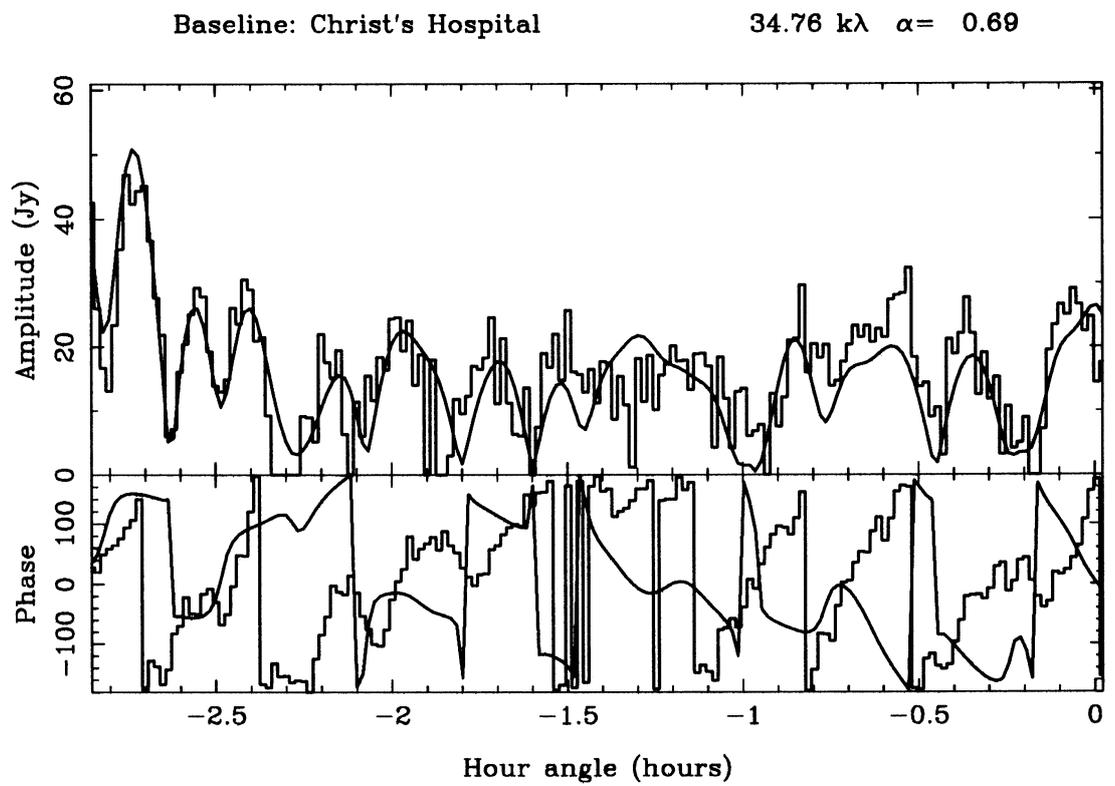


Figure 7.1: Predicted VLA data (smooth line) superimposed on the LBI observations made at Christ's Hospital.

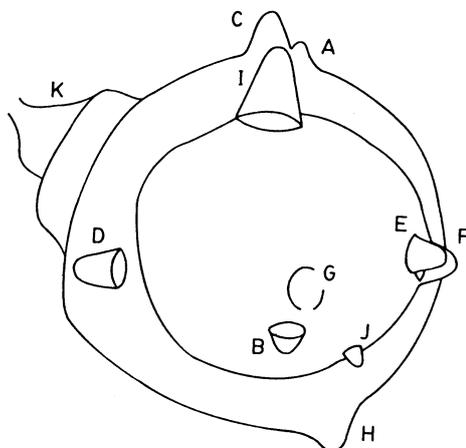


Figure 7.2: The major shell and bow-shock features identified by Braun *et al.* [14].

on the flux density of any *single* steep spectrum component, assuming that the remnant is otherwise identical at the two frequencies. Consider a set of model data, $\{m_i\}$, representing the amplitude of the visibility expected if the remnant looked exactly as the VLA map, but with an extra compact component of flux density S in pixel (j, k) . Although, strictly, the noise on the data has a Rician distribution, the signal-to-noise ratio is sufficient for Gaussian statistics to be a good approximation. It is therefore valid to associate the probability of measuring data $\{d_i\}$ given this model as

$$p(\{d_i\}|\{m_i\}) = (\sigma\sqrt{2\pi})^{-M} \exp \left[-\sum_{i=1}^M \frac{(d_i - m_i)^2}{2\sigma^2} \right], \quad (7.1)$$

where M is the number of data points and σ is the noise in the data. Assuming there is no *a priori* reason to choose one model over any other, Bayes' theorem states that this probability is proportional to the likelihood of the model given the data. The most likely model given the data is the one that maximises this function, or conversely minimises

$$\chi^2 = \sum_{i=1}^M \frac{(d_i - m_i)^2}{2\sigma^2}. \quad (7.2)$$

Figure 7.3 shows the result of applying this principle to a series of models comprising the VLA map plus an unresolved component at each of the positions of the major shell and bow shock features identified by Braun *et al.* [14] (see Figure 7.2). For each value of S the model is renormalised to minimise its misfit to the data. The flux density scale on the 81.5 MHz data is therefore considered as a free parameter over some range, though the calibration is sufficiently good to limit its freedom to within about a factor of two. This means that it is not necessary to consider point source flux densities of more than a few tens of janskys. χ^2 is shown as a function

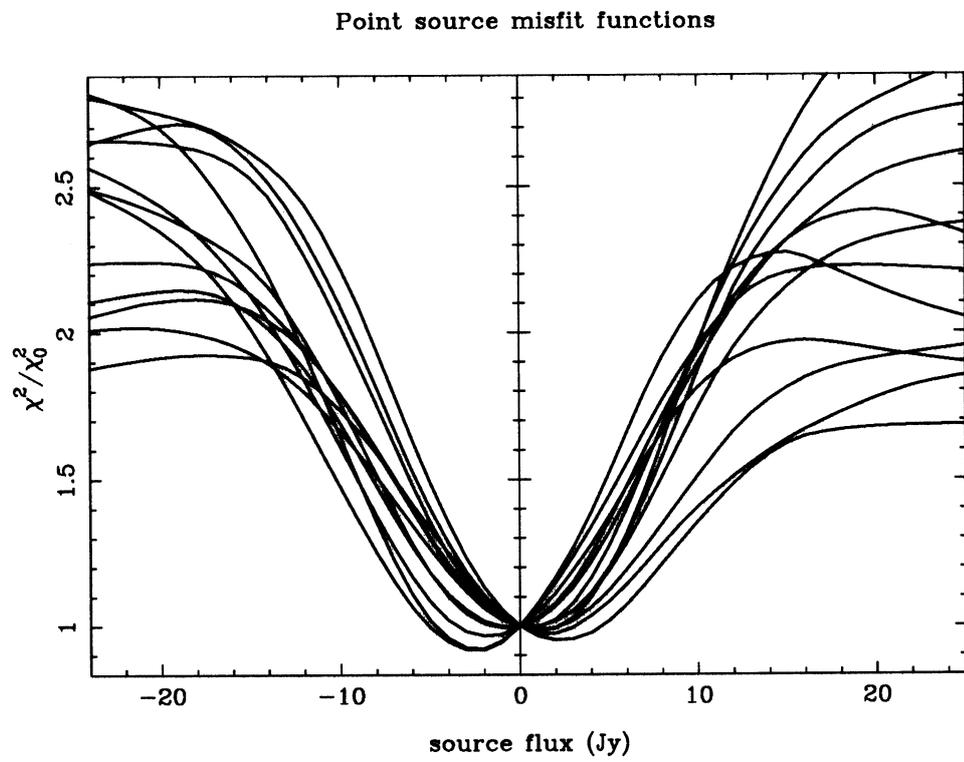


Figure 7.3: The misfit between the 81.5 MHz LBI data and various models comprising the VLA map plus a single point source at the features in Figure 7.2.

of the flux density of the compact source, S , and is normalised by χ_0^2 , its value at $S = 0$. It can be seen that at all these positions the compact flux density need only change by 4 or 5 Jy to give the most likely model parameter, S .

However, these points are only a small number of the possible positions at which a compact component may be found. In order to set a limit on the flux density from any such component it is necessary to repeat this process for all the possible positions. The baseline from which the data derives has a resolution of $\sim 6''$, so it is quite sufficient to model all possible positions in the remnant on a 128×128 grid, with a pixel size of $3.2''$. When this is done it is found that at no point on the VLA map does an additional unresolved source of more than ± 5 Jy (at 81.5 MHz) improve the fit of the model to the data. It is therefore unlikely that the remnant is the same at high and low frequency save for a point source of more than a few janskys. The next question is, ‘how unlikely?’. The confidence associated with the estimated model parameter S depends on the curvature of the χ^2 surface around its minimum. This depends on the noise in the data and is therefore hard to estimate without reasonable knowledge of this noise. In Figure 7.3 χ^2 is normalised by its value when the model comprises just the VLA map, without a point source, so the relevant curvature information has been lost. However, an approximate value for the confidence of each flux density estimate can still be determined. It is clear from Figure 7.1 that the fit between the plain VLA map and the data is extremely good. If the fit were perfect to within the noise, then its misfit $\chi_0^2 \simeq M/2$. Assuming two degrees of freedom in the model fitting procedure (the source and map flux densities), the 99.73% confidence level is reached when χ^2 has changed by 11.8, i.e. when $\chi^2/\chi_0^2 = 1 + 23.6/M$. There are 175 data points, so this confidence interval is bounded by $\chi^2/\chi_0^2 \simeq 1.135$. For the models considered in Figure 7.3 this represents an uncertainty of perhaps 4 or 5 Jy, but of course there is no reason to assume that the curvature of χ^2/χ_0^2 is the same for all source positions. However, none of the curvatures at any of the 128^2 source positions considered are sufficiently small to increase the uncertainty in the fitted flux density to more than about 10 Jy at the 99.73% level. Support for this result comes from the spread in fitted flux densities from all 128^2 possible source positions (Figure 7.4). The estimates are normally distributed with standard deviation $\sigma = 1.54$ Jy, indicating that *if* all the χ^2 curvatures are the same, the 99.73% confidence level (i.e. 3σ) is reached at 4.62 Jy.

In conclusion, the LBI data measured at 81.5 MHz on a projected baseline of $\sim 30\,000\lambda$ ($6''$ resolution) reveals two basic features of Cas A:

1. The compact ($\lesssim 6''$) structure of the remnant appears largely identical at 81.5

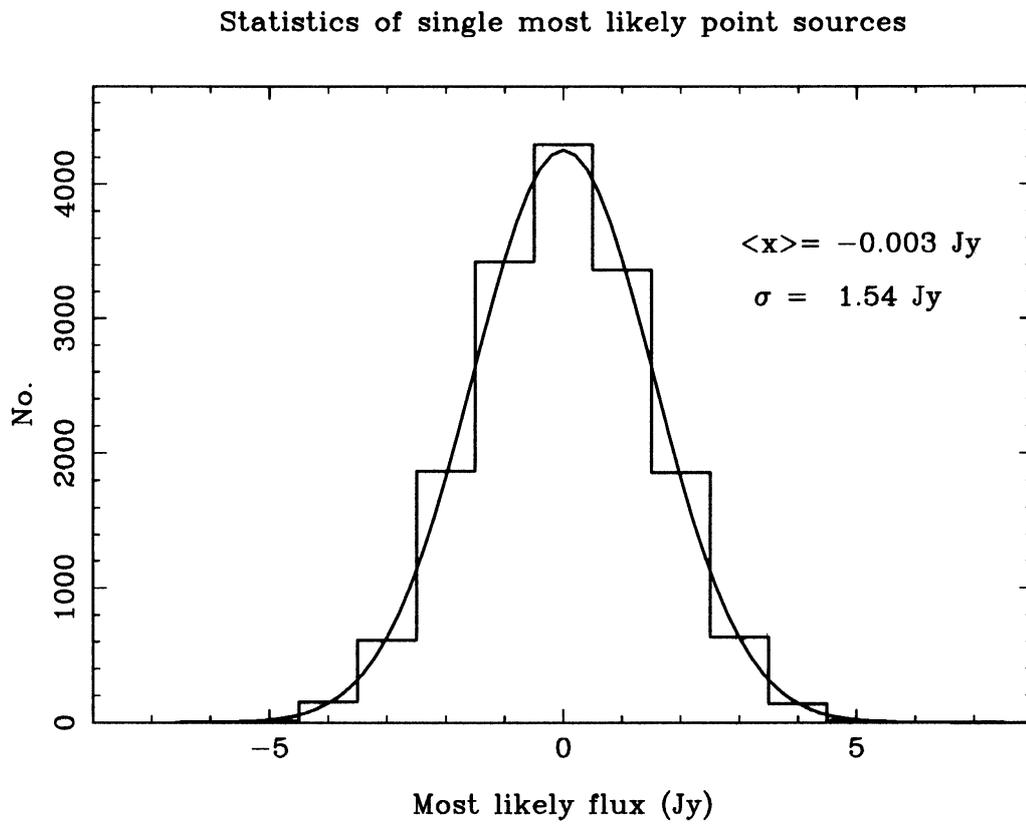


Figure 7.4: *Distribution of the most likely flux densities for a point source at each of the 16,384 possible positions in and around the remnant.*

MHz and 1381 MHz.

2. If the low frequency character of the remnant is modelled as the high frequency map plus a single point source anywhere within $205''$ of the map centre, the data restricts the flux density of this source to less than 5 or 10 Jy at 81.5 MHz.

7.2.3 Discussion

The null result obtained in the previous section contrasts sharply with earlier observations of the remnant described at the beginning of the chapter. There are a number of interpretations for this. Firstly, most of the measurements were made some years ago and, at least in the case of the 38 MHz data, indicate that the compact feature detected was transitory. It might be that whatever was seen no longer exists, and if this is the case little more can be said. However, assuming a pessimistically steep spectral index of 3 for any compact feature, the new measurements put an upper limit of about 70 Jy on its present flux density at 38 MHz, which is about 0.2% of the total flux density at this frequency. This figure must be compared with the total power measurements of Erickson *et al.* [28], Read [66, 67] and Walczowski & Smith [102] who apparently measured regular flares of 10-15%, the last of which was still very active in 1984.4. It does seem unlikely that regular, dramatic events such as these could all but vanish in the three years since measurements ceased. It seems more likely that despite the considerable amount of care that went into these observations, they contained systematic errors that were not accounted for. It is interesting to note that there is a strong anti-correlation between the measured flux density ratios of (Cas A)/(Cygnus A) and sunspot number. Without suggesting any specific mechanism, it is quite possible that this is an ionospheric effect.

The interferometric measurements made by Hutton *et al.* [47] can be compared with the present data more easily. Considering just their 74 MHz observations on a baseline of 12000λ , they claim to find (using methods similar to those described in the previous section) a major discrepancy between their data and a map made three years earlier at 2695 MHz. They model this discrepancy as a point source of $\gtrsim 100$ Jy associated with, but possibly some distance from, the remnant. In fact this baseline is closer to the largest spacings considered in Chapter 6 than the Christ's Hospital baseline, but in either case the fit of the new 81.5 MHz data to the VLA (1381 MHz) map is vastly superior to the fit between their data and map. It has been pointed out by others [66, 9] that the map data with which Hutton *et al.* made their comparison was badly undersampled in the relevant area of the (u, v) plane,

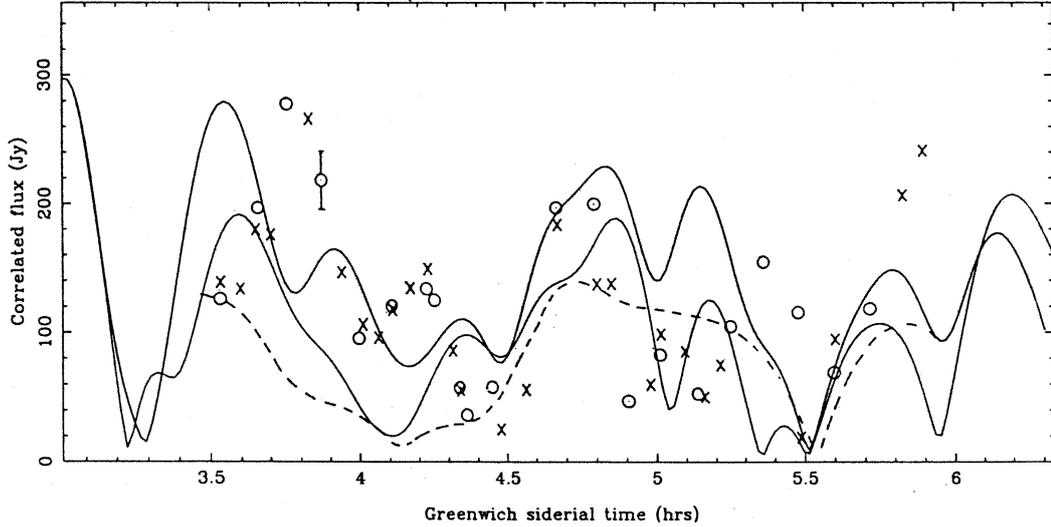


Figure 7.5: Comparison between the data due to Hutton *et al.* (circles and crosses) and the corresponding visibility predicted by their model (dashed line). The upper solid line is the prediction from the 1987 VLA data assuming no change in the remnant, and the lower line that predicted assuming a uniform expansion of 2.5%.

and that this may explain the mismatch. Confidence in the result is further shaken by observations they made at a very similar frequency (111 MHz) which are not only inconsistent with their reference map but also with the adjusted map containing an extra point source.

Although these LBI observations were made 15 years ago it is interesting to compare them, as for the 81.5 MHz data, with the corresponding (u, v) track derived from the recent VLA map. The baseline (between Sugar Grove and the recently demised 300 ft dish at Greenbank) is taken to be $l = 51287.8\text{m}$, $\delta_b = 0.15384214$ radians and $h_b = 0.31054263$ radians. The remnant has evolved considerably over the intervening period on the scale-size of the observations, so the fit between the old data and the new model is poor. Figure 7.5 shows the data points due to Hutton *et al.* together with the prediction from their model (dotted line). The two solid lines show the data predicted by the new (VLA) model assuming no change in the remnant ('upper' trace), and assuming an isotropic expansion of 2.5% over the 15 year interval ('lower' trace). The radio expansion of the remnant is known to be quite disordered, which, together with the expected changes in brightness of the more compact features, make an accurate prediction impossible for baselines of this resolution. However, it is notable that the 'quasi-sinusoidal variations', with a period

~ 20 min, that Hutton *et al.* identified as present in their data but not in their model and attributed to a steep-spectrum point source, *are* present in the predictions from the new model containing no such source. It therefore seems probable that the compact component claimed by Hutton *et al.* is due to a bad model of the remnant and is fictitious. Certainly no compact feature of its brightness exists in the remnant at present and, with this better model, there is little reason to believe one ever did.

Lastly, it is necessary to reconsider the 20 and 25 MHz interferometric observations due to Bovkoon, Braude and Megn [12]. Taking their 25 MHz observations as an example, they compare the measured visibility of Cas A over 6 hours in hour angle on a baseline of 2840λ with a very simple model of Cas A comprising a circularly symmetric ring superimposed on a disc-shaped plateau. The flux densities and extents of the two components in the model are chosen to fit observations made on (mostly) shorter baselines and higher frequencies. They associate the difference between their measured visibility (almost constant at 0.037) and that predicted from the model (0.01) with flux from a supposed steep spectrum region in the remnant. However if their simple model is replaced by the VLA map used successfully in all the above analysis, a visibility closer to 0.02 or 0.03 is predicted, so that the expected and unaccounted for fluxes are now of similar magnitude. Furthermore, according to the VLA model, the visibility along their aperture track should show strong undulations, reflecting the asymmetries in the remnant on scales of an arcminute, whilst their measurements are strikingly constant over the full 6 hours of observation. As their (u, v) track traverses 50% of the transform plane, with or without the extra steep spectrum component there should be beats in the amplitude. If the measurement is correct, Cas A looked *totally* different when these observations were made at 25 MHz from how it looks today at higher frequencies (including 81.5 MHz).

7.2.4 Conclusions

The 81.5 MHz LBI observations of Cas A described in both this and the previous chapters show its low frequency structure to be largely identical to that at 1381 MHz. In particular the remnant cannot be modelled as differing only by a single, compact component of much more than 5 Jy as suggested by previous observations. Although it is possible that its spectral structure has changed since those observations were made, it seems, for the reasons described above, more likely the feature never existed.

7.3 Pulsar Search

These recent LBI observations can only set an upper limit on the flux of a steep spectrum compact feature in the remnant – a limit too high to rule out the existence of a pulsar.

Present theories associate pulsars with spinning, highly magnetised neutron stars left behind after Type II supernovae, an idea supported by the strong correlation between the galactic distribution of pulsars, supernova remnants and their presumed progenitors, OB stars. Therefore Cas A, as the youngest known supernova in our galaxy, might well contain a young pulsar. A pulsar can be expected to last a lot longer than its associated supernova remnant and although usually acquiring a high linear velocity from the supernova event ($\gtrsim 10^2$ km/s), can further be expected to lie within the remnant while it is visible. However, of the ~ 150 known galactic SNRs, only four (Crab, MSH 15-52, Vela X, CTB 80) have a strong pulsar association, the only other association being 0540-69.3 in the LMC.¹ There are a number of possible explanations for this low figure, but they all reduce to two alternatives – either the pulsars are not there, or at least have not shown themselves yet, or they are too dim (perhaps through beaming effects) to be seen above the highly luminous remnants. It should be noted here that candidate pulsars can be identified without measuring pulsed emission, which may be highly directional. All five of the known pulsars associated with SNRs are surrounded by small, diffuse synchrotron nebulae, thought to be powered by the rotational energy lost by the neutron star. These nebulae, which should not be affected by beaming, seem most easily detected by their X-ray emission [42].

A partial solution to this dilemma may come from the observed pulsar period-luminosity relation ($L \propto P^{-1} \dot{P}^{1/3}$ [63]) which, under the assumption that magnetic dipole radiation is responsible for the loss of rotational energy from the neutron star, can best be made consistent with measured characteristic ages (P/\dot{P}) of pulsars if their ‘injection’ periods are relatively long (0.09 to 0.25 sec) [16]. It seems that either the neutron stars giving rise to pulsars are most frequently born as relatively slow rotators, or that the pulsed emission usually starts only when a neutron star with initially rapid rotation has slowed sufficiently (presumably by radiation breaking). If the second alternative were correct it might be expected that, even in the absence of particle injection from the mechanism that makes pulsars shine, the magnetic dipole radiation from the neutron star would be strongly absorbed by the remnant itself and be evident in its structure, perhaps giving the SNR more

¹There are recent, and as yet unconfirmed, reports of a sub-millisecond binary pulsar associated with SN1987A.

of a ‘filled centre’ appearance similar to the Crab. As only $\sim 10\%$ of remnants have this form, the first alternative seems more probable. Furthermore, the rate at which such slowly rotating pulsars dump their rotational energy into the nebulae is correspondingly less. Sensitive measurements of nebula X-ray luminosities are quite consistent with the expected lack of associations between pulsars and SNRs under this model. Seward and Wang [81] have gone further, using an empirically derived relationship between X-ray luminosity and rate of loss of rotational energy to infer the properties of candidate pulsars from the X-ray emission of the host remnant. In particular, by taking a bright X-ray knot near the centre of the remnant as an upper limit to the synchrotron emission from a neutron star, they set a lower limit for the period of any pulsar in Cas A at 0.33 sec, although this is under the somewhat unlikely assumption that the characteristic and true ages of a young pulsar would be the same. The existence of a ‘slow’ pulsar in Cas A is also suggested by its anomalously high degree of IPS. The evidence for this is reconsidered in Section 7.4.

This section presents a new, deep search for a pulsar in Cas A at 408 and 81.5 MHz using the Cambridge One Mile Telescope (OMT) and the 4C phased array. Although Cas A has been regularly included in previous pulsar surveys (e.g. [24, 80]), their sensitivities have been greatly reduced by the shell emission from the remnant. In fact at 408 MHz, the system temperature is so dominated by this emission that a dish of only a few metres diameter has the same sensitivity as, say, the Bonn 100 m. Under these circumstances, the sensitivity of the observations is limited by the integration period, so it seems sensible to use a small dish on which large amounts of telescope time are available. One of the dishes comprising the OMT is an almost perfect choice for this. The instrument is rarely used as an interferometer now, and at the time of the observations both the East and West aerials were idle.

Before any further description of the instrumentation used in the search, it is necessary to consider two important constraints on its design.

7.3.1 Effects of dispersion

Any pulse travelling through the interstellar medium will be both scattered and dispersed. Scattering processes result in a broadening of the pulse as received on Earth. The broadening scales as roughly ν^{-4} and introduces a sharp lower cutoff in the observed pulse spectrum. Dispersion results from the spread in arrival times of Fourier components to the pulse after propagating through an ionised medium. The group velocity of such a pulse, centred on frequency ω is

$$v_g \simeq c \left(1 - \frac{\omega_p^2}{2\omega^2} \right), \quad (7.3)$$

where ω_p is the plasma frequency and $\omega \gg \omega_p$. The time of arrival of the packet after travelling a distance L is therefore

$$\tau(\omega) \simeq \frac{L}{c} + \frac{e^2}{2\epsilon_0 mc\omega^2} \int_0^L n_e dl. \quad (7.4)$$

The line integral of the electron density in this expression is the dispersion measure (DM) of the medium. The amount of dispersion is therefore

$$\mathcal{D} = \frac{\partial \tau}{\partial \omega} = -\frac{e^2}{\epsilon_0 mc\omega^3} DM. \quad (7.5)$$

If the receiver has a square bandpass of width $\Delta\omega$, a pulse that is initially a δ -function will travel through the medium and emerge from the receiver as a ‘top hat’ shaped pulse of width $\mathcal{D}\Delta\omega$. It is important to realise that although the pulse shape has been modified, its total energy is the same as for the original pulse in this frequency band. The ‘area’ of the pulse is therefore a constant, E_0 . It is straightforward to show that such a pulse train, $f(t)$, of mark-space ratio r , ($0 < r < 1$), and period P can be described in terms of its spectral components as

$$f(t) = \frac{E_0}{P} + \frac{2E_0}{P} \sum_{n=1}^{\infty} \text{sinc}\pi nr \cos \frac{2\pi nt}{P}. \quad (7.6)$$

This result has clear implications for any pulsar search, and especially for one based on the Fourier transform. The *power* of the n th spectral harmonic is reduced by a factor $\text{sinc}^2 n\pi r$, where r is the mark-space ratio ($r = \Delta\omega\mathcal{D}/P$), so the signal-to-noise ratio in the harmonic structure is reduced. The conventional way around this problem is to ‘de-disperse’ the signal by using banks of receivers each with a small $\Delta\omega$ and summing their outputs with appropriate delays (usually assumed linear with frequency) to compress the pulse back into its original form. This minimises the mark-space ratio and therefore recovers the higher harmonics.

What constraints does dispersion place on the period range and bandwidth of a search carried out without de-dispersing equipment? It is clearly not worth trying to detect harmonics of order $n \gtrsim 1/r$. If the received signals are sampled at a frequency ν_s , this is equivalent to saying the sampling period need be no shorter than the width of the broadened pulse, i.e.

$$\nu_s \lesssim \frac{2}{\Delta\omega\mathcal{D}}. \quad (7.7)$$

To within a factor of about two, the dispersion measure to Cas A is 90 pc/cm^3 [54]. Taking a bandwidth of 4 MHz at an observing frequency of 408 MHz, a suitable Nyquist rate would therefore be 50–100 Hz. At 81.5 MHz and with a bandwidth of 1 MHz this figure drops to 3–5 Hz.

7.3.2 Sensitivity

There are two commonly used techniques for detecting periodic phenomena in a time sequence. The first, the ‘folding’ or ‘periodogram’ method, involves repeatedly folding the data over a range of different periods and filtering the results by convolution with functions approximating the expected pulse shapes. Although this method is theoretically very sensitive, and quite efficient algorithms have been devised for it (see e.g. [87]), there is no algorithm sufficiently fast for it to be applied to long time sequences. The fast Fourier transform remains the only practical method of spectral analysis for such data. Once transformed by an FFT, folding algorithms can be applied to the *power* spectrum (i.e. amplitude²) of the data which contains equally spaced harmonics whose amplitudes are determined by the pulse shape. This incoherent addition process uses relatively little computer time and can greatly increase the signal-to-noise ratio of any spectral feature. This FFT method, together with its application to multi-channel measurements of dispersed pulsars, has been reviewed in detail by Lyne [53].

The Fourier transform is basically a coherent integration process, so its sensitivity to sharp spectral features depends directly on the length of the (coherent) time series. There are three facts to kept in mind here:

1. The time series $f(t)$ is real, so its transform $F(\nu)$ is Hermitian. If the time series has a sample period τ and is M samples long where M is a power of two, the transform is specified by its $\frac{M}{2} + 1$ zero and positive frequency elements. The spectral resolution is $\Delta\nu = 1/M\tau$ and the spectral range is $0 \leq \nu \leq 1/2\tau$.
2. If the time series is noise-like, its variance equals the variance of the (complex) transform, *independent* of their lengths.
3. The strength of any true spectral feature increases as \sqrt{M} .

To illustrate these points, consider the simple example of a cosinusoidal signal of amplitude S and frequency ν_0 embedded in noise of variance σ_t^2 , i.e.

$$f(t) = S \cos \nu_0 t + \text{noise} . \quad (7.8)$$

By applying Parseval’s theorem to the signal and noise individually, it can be seen that the transform of 7.8 is

$$F(\nu) = \frac{S\sqrt{M}}{2} \delta(\nu - \nu_0) + \mathbf{n} , \quad (7.9)$$

where the variance of the complex noise \mathbf{n} equals σ_t^2 . Clearly, increasing the length of the time series, M , improves sensitivity. The signal-to-noise ratio in the complex transform, γ_F is related to that of the time series, γ_f by $\gamma_F = (\gamma_f\sqrt{M})/2$. The upper limit on M is determined by the timescale on which the signal is no longer strictly periodic, due to Doppler effects from either the pulsar or the motion of the Earth, – usually a few hours.

As much of the signal processing is done with the power spectrum, it is also important to know a relationship between the signal-to-noise ratio in F , γ_F and that in FF^* , γ_{F^2} . If the noise is Gaussian, an unbiased estimate of the signal is $(FF^* - \sigma_t^2)$. By evaluating the ratio of the expectation value and the square root of the variance of this estimator, it is straightforward to show that

$$\gamma_{F^2} = \frac{\gamma_F^2}{\sqrt{1 + 2\gamma_F^2}}. \quad (7.10)$$

Equation 7.10 shows that weak spectral components ($\gamma_F \lesssim 1$) are almost killed by the squaring process, and it takes a lot of subsequent (incoherent) averaging to recover them. The critical parameter in any pulsar search is therefore the length of the coherent time series, as this maximises γ_F . Incoherent processes, such as searching for harmonics or averaging power spectra obtained on different days, are of marginal use if the complex spectral data on each has a signal-to-noise ratio much below unity.

7.3.3 Method

The essential elements of the pulsar search are shown in Figure 7.6. The main search was carried out between April and May 1988 using the East aerial of the OMT at 408 MHz and with a bandwidth of 4 MHz. The signal was fed to a standard phase-switched receiver, with its phase switch disconnected so that it performed as a total power device, and its AGC time constant lengthened to about 30 sec so that any slow pulses would not be filtered out. The detected output was fed through a 4-pole bandpass Bessel filter (0.1–50 Hz) and then digitised. Although 1-bit digitisation reduces sensitivity by only $\sim 20\%$ [53], a full 8-bit scheme was employed so that the equipment could later be used for other purposes. Once digitised, the data was logged onto 8 inch floppy discs at a rate controlled by a rubidium standard. The details of the data acquisition equipment are presented in Appendix H. Data was collected at a rate of 100 Hz for a continuous period of 6 hours a day. All the subsequent data reduction was carried out on a microvax. For practical reasons the data was processed in two halves, giving the system a coherence time of 3

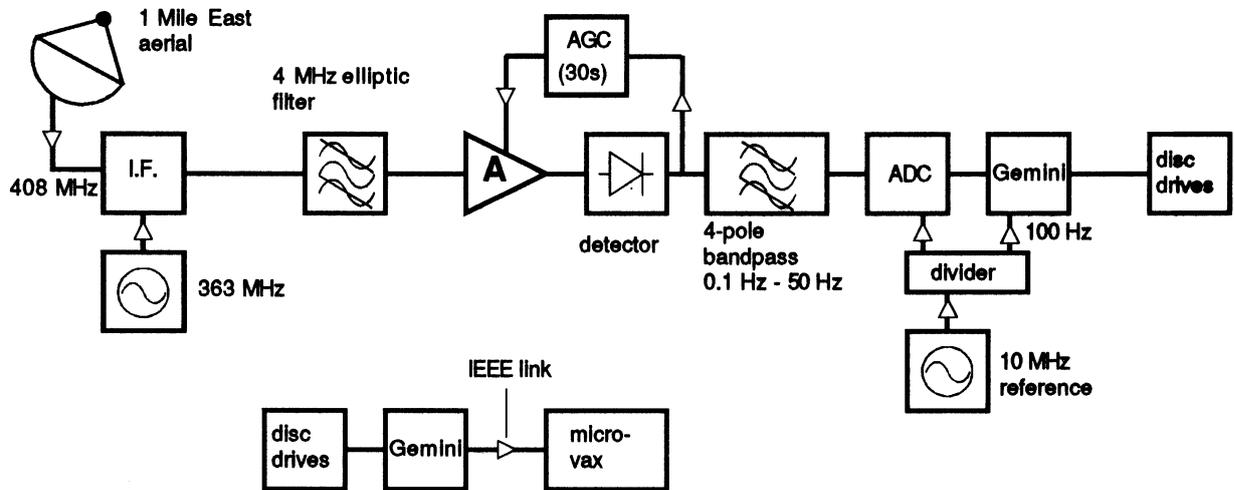


Figure 7.6: Schematic diagram of the pulsar detection system.

hours. Once a spectrum had been determined, it was re-binned to take account of the Doppler shift from the component of the Earth's velocity in the direction of the source. In this way spectra measured over an extended period could be added directly.

The system was tested by observing two well known pulsars; PSR0329+54, which has a period of ~ 0.7 s and is exceptionally bright (mean flux of 2 Jy), and the Crab (PSR0531+21, $P \simeq 0.03$ s, $S \simeq 400$ mJy). Their normalised spectra, shown in Figure 7.7, give signal-to-noise ratios that are in quite good agreement with theoretical sensitivity calculations. Furthermore, the measured barycentric frequency of the Crab (29.9853325 Hz) is well within a bin-width (9.53×10^{-5} Hz) of that predicted [84], assuming a slow-down rate of $\dot{P} = 36.46 \times 10^{-9}$ sec/day.

One major problem with the search was a considerable amount of low frequency interference, peaking at around 0.25 Hz, that was largely due to transmissions from continental weather balloons and local mobile stations. The problem was minimised by making the observations at night, but it severely reduced the sensitivity of the system to periods $\gtrsim 2$ s.

In total, Cas A was observed for 114 hours. The data were Fourier transformed in 3 hour sections and their power spectra added incoherently. A smooth baseline representing the changes in system temperature with frequency was determined from the data and used to calibrate the spectrum to give a constant noise power at all frequencies.

Any spikes in this spectrum above some critical level may be taken as the detec-

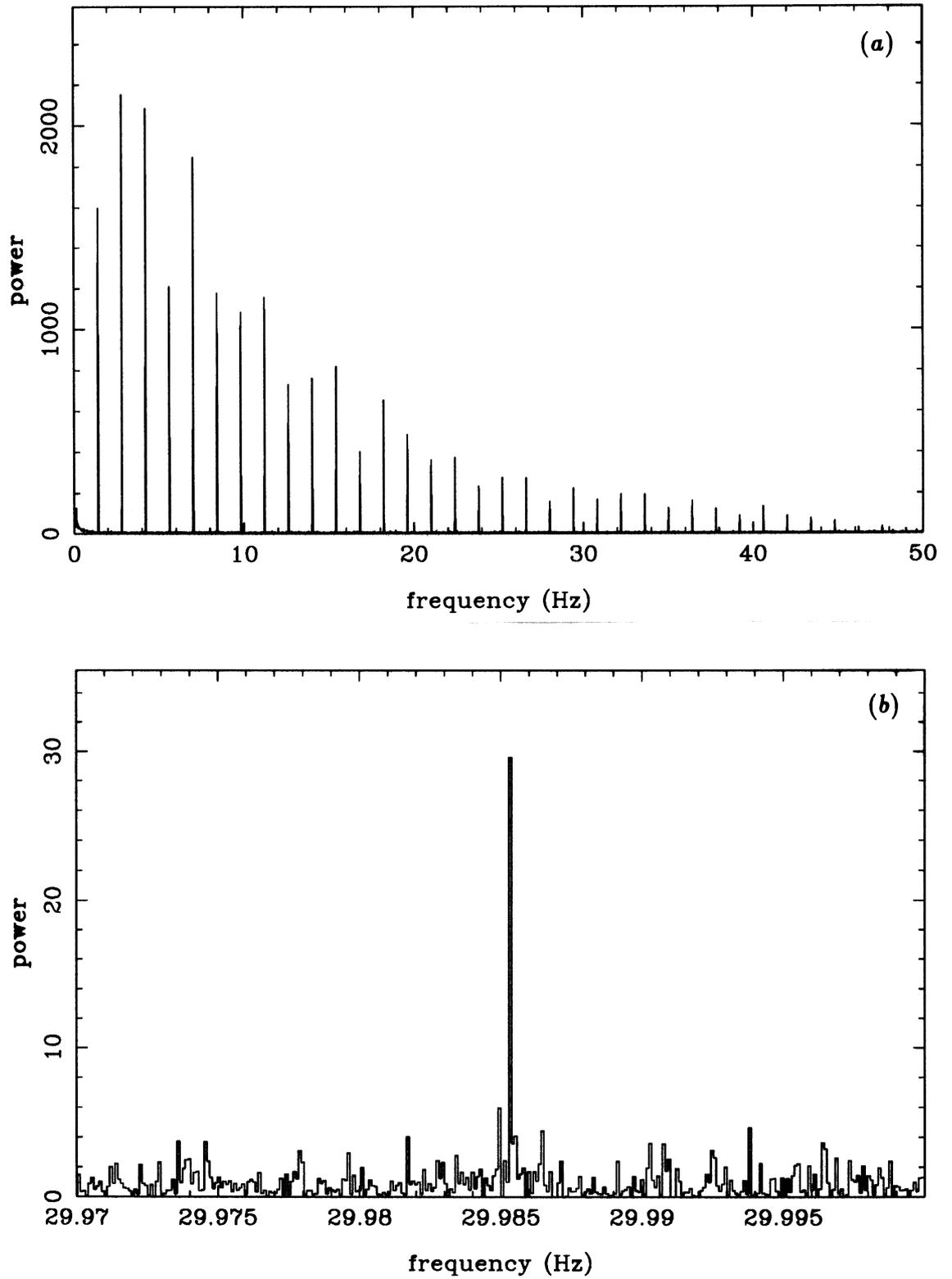


Figure 7.7: *Barycentrically corrected power spectra measurements of PSR0329+54 (a) and of the Crab (b).*

tion of a narrow-band spectral feature in the data. The determination of this level is described in Appendix I, the main results from which are as follows:

- If all the system noise is due to Cas A, the signal-to-noise ratio of a single, complex spectral measurement is

$$\gamma_F = \frac{S}{2S_{\text{Cas}}} \sqrt{\frac{\Delta\omega M\tau}{2\pi}}, \quad (7.11)$$

where S is the flux density due to the Fourier component S_{Cas} is the total flux density of the remnant and $\Delta\omega$ is the RF bandwidth.

- If the original data contains just Gaussian noise, each derived power spectra is Rayleigh distributed and the sum of N such power spectra, $y(\nu) = \sum_i F_i F_i^*(\nu)$, has a χ^2 distribution with $2N$ degrees of freedom. If y is normalised so that its mean and variance equal N , its distribution becomes

$$p(y) = \frac{e^{-y} y^{N-1}}{(N-1)!}. \quad (7.12)$$

- A detection is taken as any spike in $y(\nu)$ having only a 5% probability of occurring by chance anywhere within the spectrum. The corresponding detection level, y_d , satisfies

$$\frac{0.05}{M/2} = e^{-y_d} e_{N-1}(y_d), \quad (7.13)$$

where

$$e_{N-1}(y) = 1 + y + \frac{y^2}{2!} + \dots + \frac{y^{N-1}}{(N-1)!}. \quad (7.14)$$

- In turn, this detection level corresponds to a signal-to-noise ratio in each of the summed spectra (assumed equal) of $\gamma_{F^2} = (y_d/N) - 1$. By using Equations 7.10 and 7.11, this expression can be converted to a flux sensitivity – for this experiment a sensitivity of 80 mJy per harmonic.

Graphs for the detection levels and associated flux densities as a function of the number of power spectra taking part in the incoherent sum, N are shown in Figure 7.8. The diminishing returns obtained from an incoherent averaging process are quite evident. Tests carried out with simulated data are in good agreement with these sensitivity calculations.

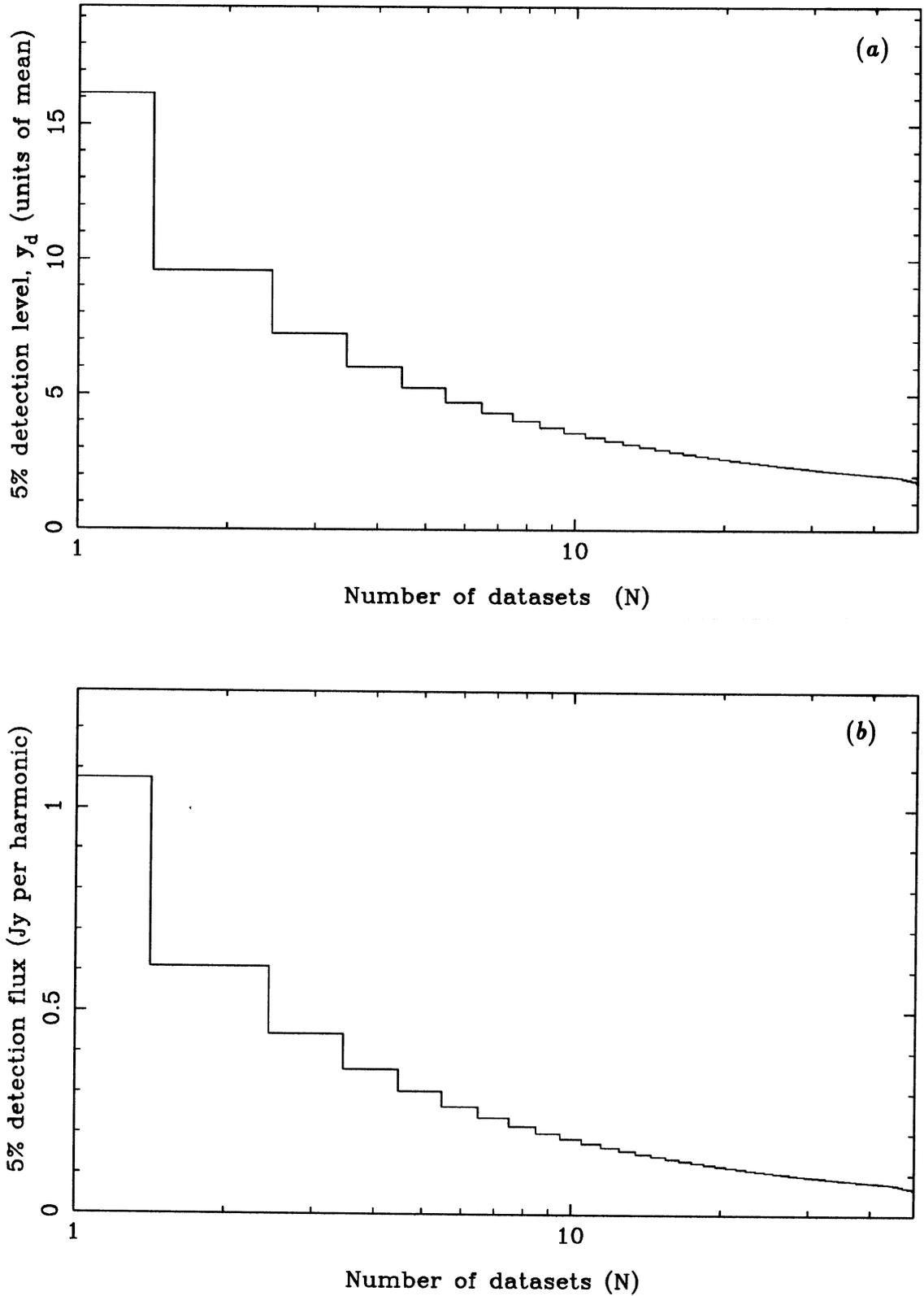


Figure 7.8: Detection limits as a function of the number of observations, N , in the incoherent sum. (a) shows the minimum signal detected at the 5% level and (b) the corresponding flux density in that harmonic.

7.3.4 Results

At 408 MHz

With the exception of the low frequency rumble mentioned at the beginning of the previous section, the summed power spectrum is remarkably free from significant spectral features. Figure 7.9 shows the spectrum in its entirety (0 to 50 Hz) normalised to units of its own mean, together with magnified views of its extremities. The 5% significance level is at a little over two units on this scale. To avoid plotting all $\sim 10^6$ channels, these traces represent only upper envelopes to the spectrum and so appear biased when many data points are present.

The only significant feature with a frequency higher than ~ 1 Hz occurs at ~ 49.96 Hz – almost the limit of the spectral range. It is also of course very near the mains frequency, and should therefore be treated with enormous caution. Having said that, the feature does have some interesting properties. Firstly, it is highly coherent. An inspection of the individual spectra show that the generally high level of activity that might be expected around 50 Hz is represented in the sum by a broad plateau, about 0.1 Hz wide. Furthermore, this feature is only present about 50% of the time, and when there is highly monochromatic over a 3 hour observation ($\Delta\nu/\nu \leq 2 \times 10^{-6}$). Secondly, the position of the peak responds quite well to barycentric correction over the 21 days of observation. Figure 7.10 shows the time evolution of the feature’s measured (‘□’) and corrected (‘+’) frequencies. Although the corrected frequency appears far from stable, it no longer contains the gross slope evident in the uncorrected measurements. Thirdly, the feature is not seen in four examples of spectra derived when Cas A was outside the primary beam, though its variable character makes this only weak evidence that the signal is celestial.

Considering its notorious frequency, the true nature of this signal will only be identified after further observation. When there, its apparent flux is a few janskys at 408 MHz. More measurements of Doppler shift (the maximum of which occurs in early February) or of interferometric fringe rate should be sufficient to expose its origin.

There are no significant features in the body of the spectrum; all other detections are at or below 1 Hz. The large spike at 1 Hz (Figure 7.9) is clearly instrumental and caused by small modulations in the power rails to the ADC from the regular flash of an adjacent LED. The regular spikes labelled ‘*’ are due to narrow pulses with a period of 10.24 seconds. This is the readout rate from the FIFO tree, so these too must be regarded as instrumental. There are only two other significant peaks in the spectrum. The first of these, at a frequency of 0.0852 Hz, is in the region of the

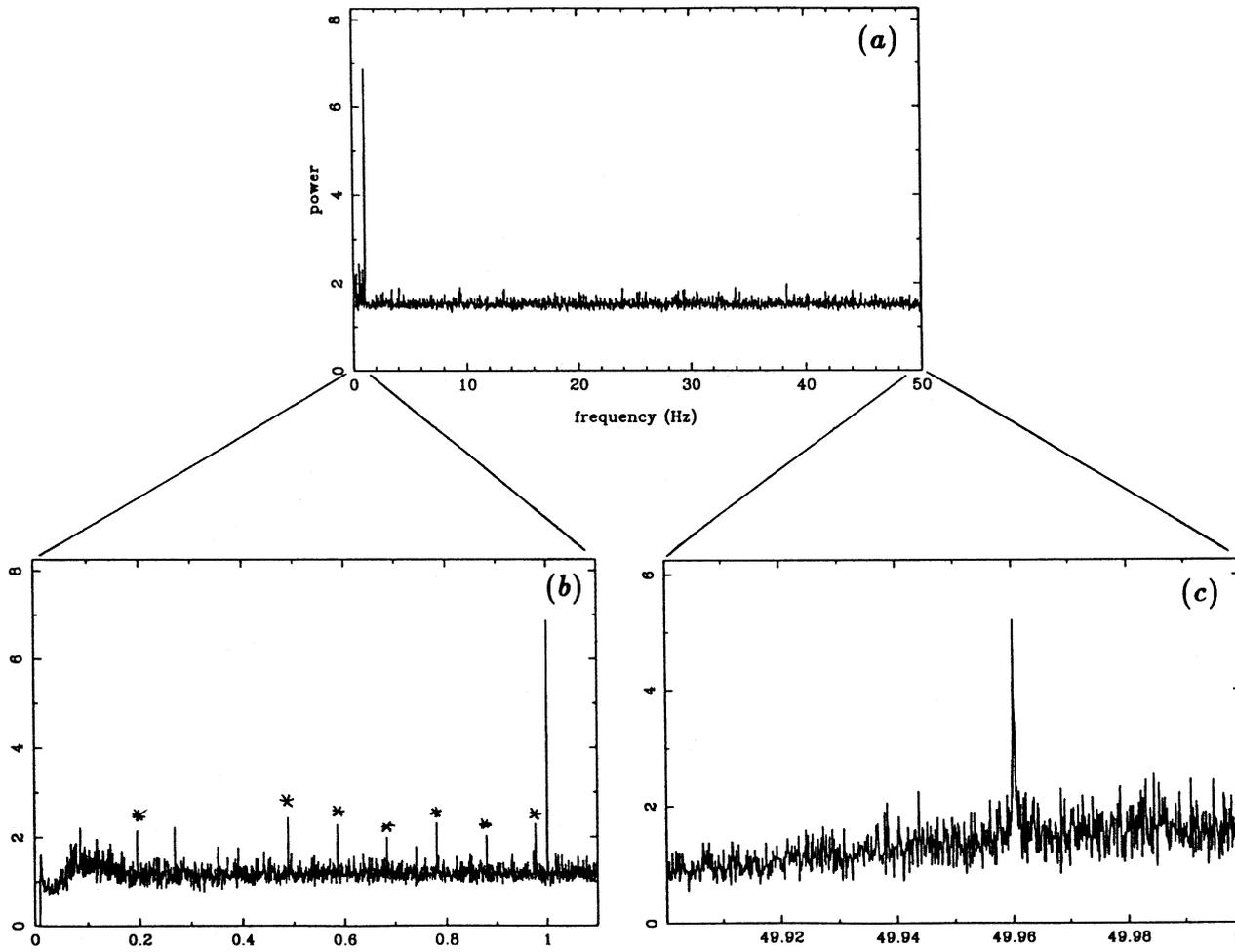


Figure 7.9: The summed power spectrum, in units of the mean, from all 38 three-hour observations of Cas A. (a) shows the full spectrum. (b) and (c) show magnified views of its extremities.

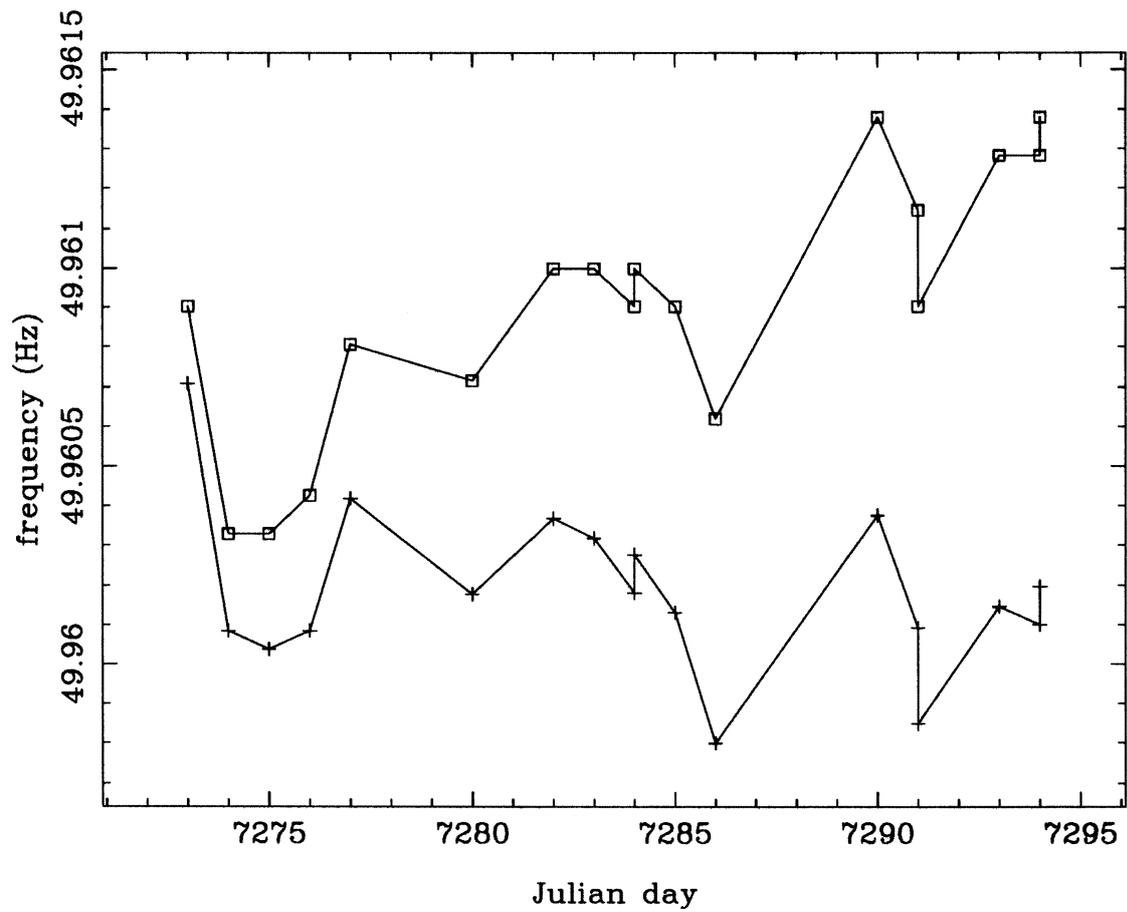


Figure 7.10: Effect of applying barycentric correction to the 49.96 Hz feature. The uncorrected data is denoted by ‘□’ and the corrected data by ‘+’.

spectrum that was badly affected by interference. Furthermore, the baseline-fitting procedure has clearly failed to remove a bump on which this peak sits. It is therefore unlikely that this peak represents a true spectral feature in the emission from Cas A. The second peak, at 0.2677 Hz, is harder to explain. Although its significance is only slightly greater than the 5% level the fit to the baseline in its proximity is good, so it is truly unlikely to be a result of the noise process. However this peak lies at the edge of the same interference band that affected the previous feature, and so it too is somewhat suspect. Unfortunately the barycentric correction is less than the spectral resolution at this frequency, so Doppler measurements are unavailable.

At 81.5 MHz

A total of 72 hours of data was collected on Cas A at 81.5 MHz using identical equipment to the above and the 4C antenna with an RF bandwidth of 1 MHz. The expected dispersion is such that all celestial signals with frequencies higher than ~ 1 Hz should be smeared out. The strong ‘50 Hz pulsar’ observed at 408 MHz is not seen at 81.5 MHz, but there is a dense forest of interference spikes around that frequency. Below 1 Hz there are a number of detections, but they can all be attributed to harmonics from pulses with a 60.1 sec period. These are almost certainly related to the update cycle of the phase rotators in the array.

7.3.5 Conclusions

This search shows there to be no strong candidate for a pulsar in Cas A with a flux density >80 mJy at 408 MHz and period >0.02 s. At 81.5 MHz there is no candidate with $P > 1$ s and $S > 100$ mJy. The 0.2677 Hz feature seen at 408 MHz is not seen at 81.5 MHz and so must be rejected. This leaves only the 49.96 Hz feature outstanding. However, if this does represent a pulsar, it would need to be quite bright. A combination of three effects (dispersion, aliasing and filter roll-off) reduce the flux sensitivity of the search at the end of its range by perhaps a factor of five. Although the apparent flux density of the pulsar is only ~ 1 Jy its true flux density could be as much as 10 Jy at 408 MHz, which would imply about 200 Jy at 81.5 MHz. This is in conflict with the LBI measurements which constrain the flux of a steep spectral feature to ~ 5 Jy at 81.5 MHz, although it is quite probable that the pulsar would have an intrinsic low frequency turnover.

7.4 IPS evidence

The intensity of the radiation from distant radio sources, when observed at low frequency, fluctuates due to diffraction by moving density irregularities in the plasma that makes up the solar wind. The phenomenon, known as interplanetary scintillation has been studied intensively since its discovery in 1964 [44], and the underlying mechanism is thought to be well understood. For sources with solar elongations greater than about 35° at 81.5 MHz, the irregularities along the line of sight fluctuate in density by only $\sim 1\%$, so weak scattering approximations can be used to predict the rms fluctuation in intensity at the Earth given a model for the scattering power and spacial spectrum of the irregularities. Sources of finite angular size scintillate with less power, as their speckle pattern on the ground is the superposition of many similar, but slightly displaced patterns from the points making up the source. The effect is only noticeable when the smearing is comparable with the speckle size. The critical source size for the interplanetary medium is about $0.5''$.

A number of IPS measurements of Cygnus A and Cas A were made in 1980 by Duffett-Smith and Tsien [93] and by Duffett-Smith & Tuffs (unpublished) showing that, although there were no discernable scintillations from Cygnus at 81.5 MHz, Cas A regularly had a scintillation index, $m = \Delta S/S$, of about 0.09%, corresponding to a scintillating flux density of about 15 Jy rms. For observations of sources with small angular diameters, the most appropriate model for the solar wind is that proposed by Readhead, Kemp and Hewish [69], which employs a Gaussian model for the spacial power spectrum of the irregularities, such that

$$F^2(u) = \pi \xi_0^2 \exp \left[-\pi^2 \xi_0^2 u^2 \right], \quad (7.15)$$

where ξ_0 is the ‘scale-size’ of the irregularities, defined at different distances, r from the Sun as

$$\begin{aligned} \xi_0 &= 175r^{0.5} \text{ km} & 0.1 < r < 0.6 \text{ AU} \\ &= 250r^{1.25} \text{ km} & 0.6 < r < 1.0 \text{ AU} \\ &= 250 \text{ km} & r > 1.0 \text{ AU} . \end{aligned} \quad (7.16)$$

This model was devised to account for the existence of an observed dominant scale-size of ~ 100 km in the fluctuations and might therefore be inappropriate when considering the scintillation of large angular diameter sources such as Cas A. Nevertheless, it predicts the scintillation index for a point source at the solar elongation of Cas A during these measurements (60°) of 34%. Assuming there is no scintillation from structure on the scale of a few arcseconds, this would mean that the rms flux from Cas in scale-sizes of less than about $1''$ is ~ 45 Jy. However, the LBI measurements described earlier in chapter show that there is only ~ 15 Jy rms of flux more

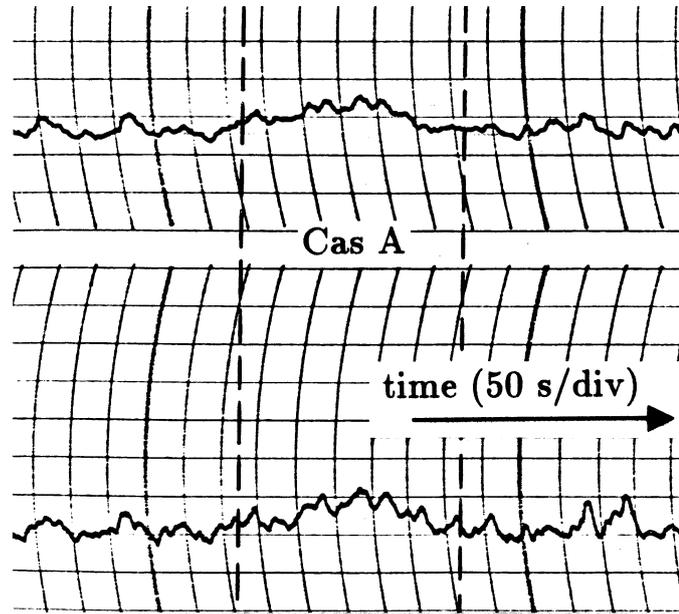


Figure 7.11: A scintillation record of Cas A. The upper trace mostly shows IPS while the lower trace shows a mixture of IPS and ionospheric scintillation.

compact than $\sim 6''$, so, by this analysis, IPS apparently overestimates the compact flux by at least a factor of three. There are three possible explanations for this:

1. The original IPS measurement was incorrectly calibrated.
2. Much of the ‘scintillating’ flux is actually pulsed flux, i.e. Cas A contains a pulsar with a period similar to the scintillation scale-time (between 0.5 and 10 sec).
3. The model of the IPM is inappropriate for sources of large angular size.

Surprisingly perhaps, the calibration (item 1) is quite straightforward. As with the pulsar search, the dominant noise term in the output of the scintillometer is the self-noise of Cas A, and is therefore independent of the system temperature or the exact forward gain of the antenna. The scintillometer used was a total power device containing a chain of filters and square-law detectors. The device had two outputs, one to register the low frequency (0.01–0.2 Hz) signals characteristic of ionospheric scintillation, and the other to measure signals in the range 0.1–1.7 Hz that are more typical of IPS. By inspecting these two records, observations that were corrupted by the ionosphere could be identified and discarded. A typical record of the scintillation from Cas A is shown in Figure 7.11. The upper trace shows the signal dominantly due to IPS, while the lower trace shows the effects

of both ionospheric and interplanetary scintillation. Both traces are proportional to the square of the scintillating flux density, and show a signal-to-noise ratio of about 1. The relationship between the signal-to-noise ratio at the output of the scintillometer, γ , and the true scintillating flux density of Cas has been determined by Duffett-Smith using methods similar to those in Appendix I, as

$$\frac{\Delta S}{S} \simeq 8.8 \times 10^{-4} \gamma. \quad (7.17)$$

All instrumental parameters such as integration times and bandwidths enter this expression as their fourth roots, so it is unlikely that the true scintillating flux density is far from the 15 Jy rms quoted at the beginning of the section. Of course these scintillations may be due to the ionosphere mimicking the effects of IPS, but this is unlikely as the scintillation is not seen on Cygnus and both Cas and Cygnus are point sources as far as the ionosphere is concerned.

The pulsar explanation (item 2) is less appealing in the light of the pulsar search described in Section 7.3. To account for the IPS observations, the pulsar would have to be slow ($P \lesssim 1$ Hz) and very bright (mean flux density of ~ 10 Jy). Even though the search sensitivity is low at these frequencies, it seems unlikely that such an object could go undetected.

The third possibility is harder to investigate. Any model of the interplanetary medium that underestimates its spectral power at short wavevectors will also underestimate the scintillating flux from an extended source. For reasonably compact sources ($\theta \lesssim 2''$) the Gaussian model is fine, as power is scattered from a relatively large range of spacial components in the plasma, and any long period anomaly in the spectrum would be relatively unimportant. However, for a source the size of Cas A ($\sim 250''$) these long period components constitute the bulk of the scattering power, so it is very important that they are modelled with accuracy. It has been stressed by Readhead *et al.* [69] that a simple power law spectrum, and specifically a Kolmogorov spectrum cannot be made consistent with the data for wavenumbers in the range $10^{-3} < k < 10^{-1} \text{ km}^{-1}$. However both spacecraft data [1] and temporal IPS observations [34] show that the spectrum at small wavenumbers is well described by a power-law of index $\sim 11/3$, and that there is considerably more extended power in the IPM than the Gaussian model would suggest. The problem is compounded by the fact that we are well within the Fresnel region when considering wavevectors of this magnitude. For this and other reasons, computer simulation is the only reliable way of predicting the amount of scintillation expected from Cas A under the various models.

7.4.1 Simulations

Provided that weak scattering approximations can be made, the response of the interplanetary medium to an extended source can be regarded as a superposition of the signals from a number of thin layers at different distances z along the line of sight between the Earth and the source. If the IPM is modelled as having radial symmetry, these layers may be thought of as slices of ‘onion skins’, centred on the Sun, with a scattering power, $\sigma(z)$, proportional to their radius⁻⁴ [69]. Each of these slices will produce a mean square scintillating power of

$$\begin{aligned} d(m^2) &= \langle g(x)^2 \rangle_x \\ &\propto \int \int \tilde{B}\tilde{B}^*(zu, zv)F^2(u, v, z)F_r^2(u, v, z)du dv, \end{aligned} \quad (7.18)$$

where $g(x)$ is the diffraction pattern on Earth, $\tilde{B}(u', v')$ is the Fourier transform of the source brightness distribution, $F^2(u, v, z)$ is the power spectrum of the electron density irregularities and $F_r^2(u, v, z)$ a ‘Fresnel filter’ term that takes account of the near-field diffraction and equals $\sin^2(\pi\lambda z(u^2 + v^2))$ (Little & Hewish, 1966 [52]). The total scintillating power is proportional to a sum of the above over all slices, weighted by their scattering powers. The constant of proportionality may be removed by evaluating the ‘scintillation visibility’, V , which is the ratio of the scintillation index of the extended source to that of a point source, m/m_0 . Hence [26]

$$V^2 = \frac{\int_{z=0}^{\infty} \sigma(z) \int \int \tilde{B}\tilde{B}^*(zu, zv)F^2(u, v, z)F_r^2(u, v, z)du dv dz}{\int_{z=0}^{\infty} \sigma(z) \int \int F^2(u, v, z)F_r^2(u, v, z)du dv dz}. \quad (7.19)$$

Both the Fresnel filter and the density power spectrum terms are independent of the orientation of the wavevector, (u, v) , so this expression can be simplified to a ratio of 1-dimensional integrals if the radial power spectral density of the source is used instead of $\tilde{B}\tilde{B}^*(u, v)$.

A series of simulations were carried out based on Equation 7.19 but using different irregularity spectra. The integration in z was carried out by summing the wavenumber integrals (determined by standard NAG quadrature routines) in steps of 0.01 AU out to 2.5 AU from the Earth. The power spectrum of Cas A was evaluated from the high resolution VLA map introduced in the last chapter, as the LBI observations show this to be a good approximation to its form at low frequency. The high-contrast view of the map transform in Figure 7.12(a) clearly shows discrete visibility tracks on the longer baselines, so the map underestimates the high spatial frequency power in the remnant. This bias was largely removed from the

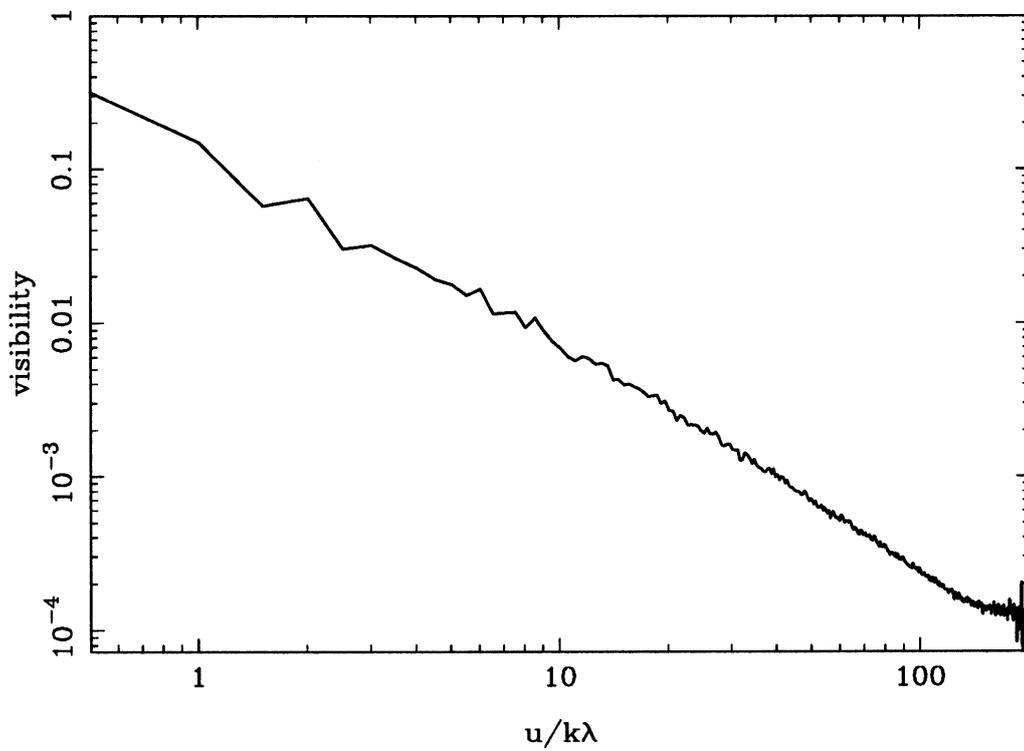
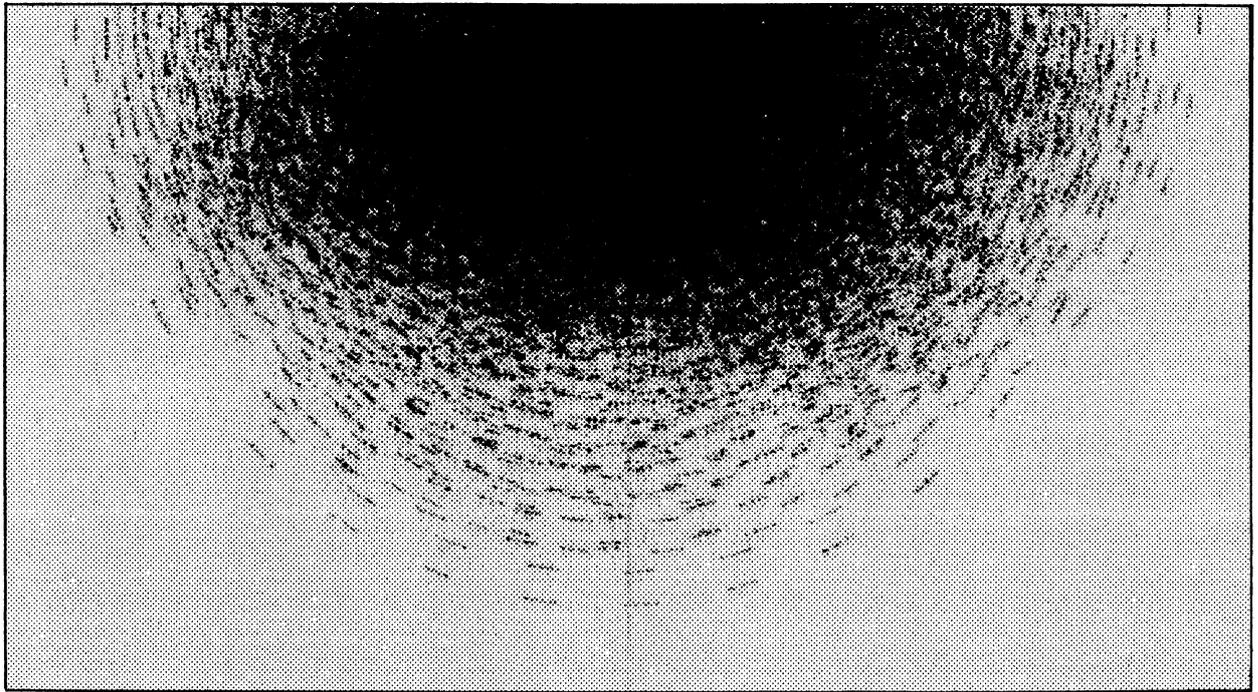


Figure 7.12: (a) The transform plane corresponding to the map shown in Figure 6.18. (b) The unbiased radial visibility function derived from the transform.

model	$V/10^{-3}$
Gaussian	0.06
$\alpha=10/3$	1.5
$\alpha=11/3$	2.6
$\alpha=12/3$	5.3
(observed)	2.7

Table 7.2: *The predicted scintillation visibility of Cas A under various models of the IPM.*

radial profile shown in Figure 7.12(b) by only including (u, v) pixels greater than the noise.

Two types of spectra were tried, the Gaussian spectrum of Readhead, Kemp and Hewish [69] and various power laws ($F^2 \propto u^{-\alpha}$) including $\alpha = 11/3$ (Kolmogorov). Table 7.2 contains the results of the simulations for a solar elongation of 60° .

A number of conclusions can be drawn from these results. Firstly, the model containing a Gaussian spectrum grossly underestimate the scintillation index of Cas A. This is perhaps not surprising as it was designed to fit observations over quite a limited range in wavenumber. Nevertheless it is worth pointing out that the model gets the wrong answers when applied to sources of this size. Secondly, the power law models do a lot better, and the Kolmogorov spectrum fits the observation remarkably well. Furthermore, the amount of scintillation depends critically on the power law index. For calculations with $\alpha \leq 4$ the power law can be extended to a wavenumber of zero without the integrals diverging as Fresnel effects win over the spectral power. However, for steeper spectra an outer scale-size would need to be defined. Thirdly, the true spectrum must be some hybrid between a power law at low wavenumber and a Gaussian at high wavenumber where a true scale-size appears to exist. Almost any hybrid would reduce the scintillating visibility of Cas A from that derived using the purely power law model, as the extra small-scale structure would produce proportionately less scintillation from Cas A than from a point source.

In conclusion, the observed scintillation of Cas A is consistent with a power-law model for the interplanetary medium with an index of $11/3$ or greater. Under these circumstances most of the scintillation actually comes from its extended structure. The lack of scintillation from Cygnus A, a source of comparable size and power, probably tells us that Cygnus is considerably smoother than Cas A and does not contain sufficient power at the spacial frequencies that are important.

Appendix A

The fringe rotator transfer function

To compute $G(\omega)$, the correlator's complex transfer function, it is necessary to know the exact form of $\mathcal{F}(t)$ – the digital fringe function. The function used is divided into 12 steps, and unwinds the phase with a fringe rate that is the nearest integral multiple of 0.1 Hz to the true fringe rate. If this fringe rate is ν Hz, The software introduces a new step every τ seconds, where

$$\tau = \frac{256^2}{500 \times 157 \times 10 \times \nu} . \quad (\text{A.1})$$

Steps are added for exactly one second (t_0), after which the phase of \mathcal{F} is reset, and the process repeated.

In Chapter 2, $G(\omega)$ is identified as

$$G(\omega, t) = \frac{e^{i\omega t}}{t_0} \int_0^{t_0} \mathcal{F}(t') e^{i\omega t'} dt' , \quad (\text{A.2})$$

the integral term of which can be rewritten as

$$I(\omega) = \int_0^{t_0} [\mathcal{C}(t') + i\mathcal{S}(t')] e^{i\omega t'} dt' . \quad (\text{A.3})$$

The functions \mathcal{C} and \mathcal{S} are realised digitally and can only take the values -1, 0 and 1 as a function of the step number. Their values on the n^{th} step will be denoted by \mathcal{C}_n and \mathcal{S}_n . If the fringe rotator completes N whole steps in one integration time, then real and imaginary parts of Equation A.3 can be written as

$$\Re[I(\omega)] = \sum_{n=1}^N [\mathcal{C}_n \cos \alpha_n - \mathcal{S}_n \sin \alpha_n] \sin \frac{\omega\tau}{2} +$$

$$[\mathcal{C}_{N+1} \cos \beta - \mathcal{S}_{N+1} \sin \beta] \sin \frac{\omega(1 - N\tau)}{2} \quad (\text{A.4})$$

$$\Im[I(\omega)] = \sum_{n=1}^N [\mathcal{S}_n \cos \alpha_n + \mathcal{C}_n \sin \alpha_n] \sin \frac{\omega\tau}{2} +$$

$$[\mathcal{S}_{N+1} \cos \beta + \mathcal{C}_{N+1} \sin \beta] \sin \frac{\omega(1 - N\tau)}{2}, \quad (\text{A.5})$$

where

$$\alpha_n = \omega n\tau - \omega\tau/2,$$

$$\beta = \omega(1 + N\tau)/2.$$

These expressions include the contribution from the partially completed last step. Once evaluated from the above relations, $I(\omega)$ can be tabulated as a function of fringe rate, and the appropriate amplitude and phase correction factors applied to the data with little further computation.

Appendix B

Portable phased array

The portable phased array consists of sixteen Yagi antennas, pointing at the zenith, connected by cables of equal length to a central ‘phase box’. This box contains sixteen RF amplifiers and phase rotators controlled by a microcomputer in the van. The computer is able to introduce phase into signals from each of the antennas in steps of 22.5° before they are combined and fed to the receivers. The data for this is sent via a parallel link from the host computer.

The antennas are erected in two lines of eight (Figure B.1) with the North line polarised N-S and the South line polarised E-W. If the most Easterly antenna of each line is denoted as antenna 1, the line of antennas can be phased up to a point on the sky of altitude a and azimuth A by adding a phase

$$\phi_n = \frac{2\pi}{\lambda} \left(n - \frac{9}{2}\right) D \cos a \sin A \quad (\text{B.1})$$

to the n th antenna, where D is their E-W separation. If circular polarisation is required, the two rows are then added together with a net phase difference of $\pi/2$ (right-hand circular) or $-\pi/2$ (left-hand circular). This means that for right-hand circular polarisation, once geometrical effects have been taken into account, all the antennas in the N-S polarised line need an additional phase of

$$\Delta\phi = \frac{\pi}{2} - \frac{2\pi}{\lambda} w \cos a \cos A, \quad (\text{B.2})$$

where w is the distance between the two rows.

A linearly polarised array can be generated by aligning all the Yagis to the same polarisation and omitting the $\pi/2$ offset in Equation B.2.

Appendix C

Automatic tape switching

Although the correlator is only able to process two tapes at a time, it can be made to behave like a three-station device if three VCRs are used, and are switched together pair-wise. In this way the three baselines can be automatically correlated in turn rather than all together. The net result is the same as for a multi-station correlator, although the processing is of course three times slower.

It is a relatively simple matter to switch the three sets of data and control connections from the VCRs (frame sync, on/off, clock, data and audio) to either of the two correlator inputs under the control of the correlator computer. A UART on the control board (Figure C.1) accepts a single serial byte from the host computer at the beginning of each correlator run. The bottom three bits of this define which tape to connect to the first correlator input and bits 5,6 and 7 define which to connect to the second. Conflicting requests are rejected. The digital switching (data and clock) is done via conventional logic gates, whilst the analogue switching (frame sync, on/off and audio) is achieved with bi-lateral switches.

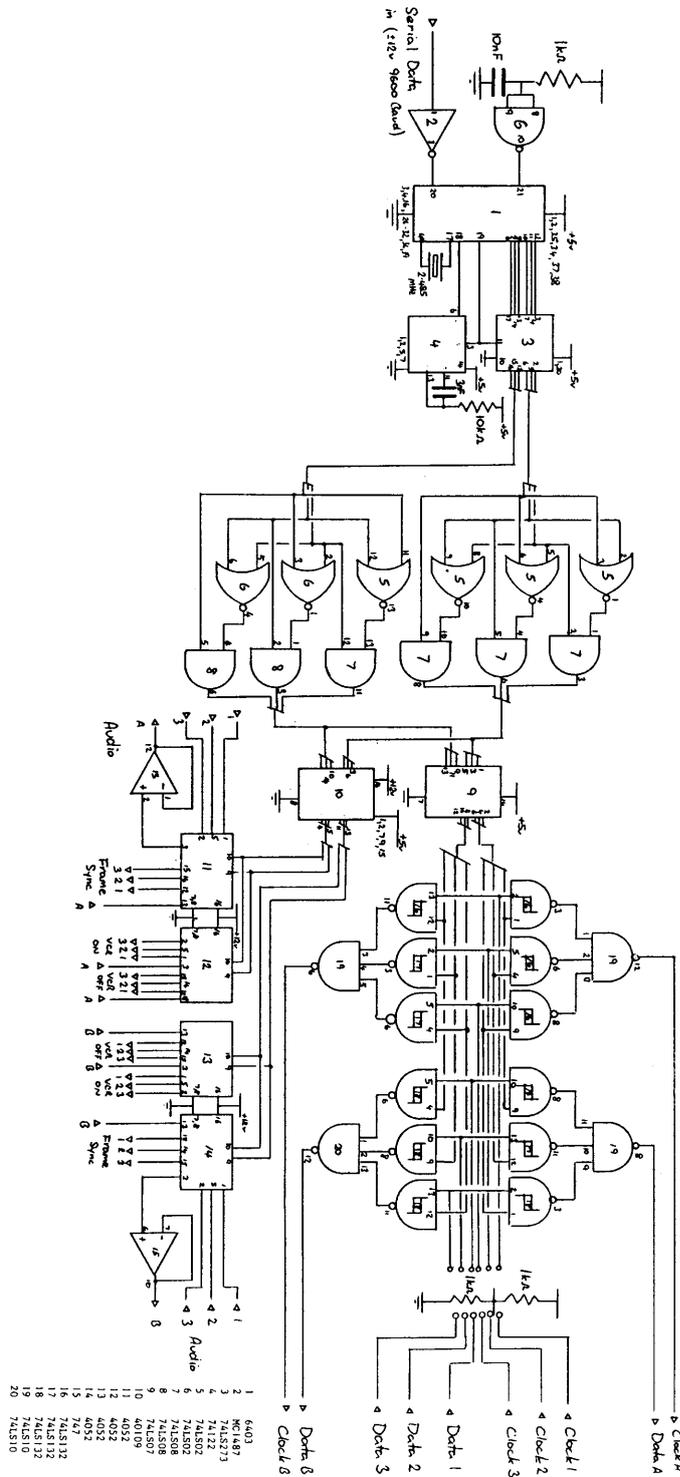


Figure C.1: VCR switcher, enabling the two-station correlator to process three-station data automatically.

Appendix D

Gaussian phase noise

Consider a complex measurement R of a real signal S made in the presence of complex Gaussian noise n (Figure D.1). If the real and imaginary components of the noise are uncorrelated, the probability of measuring a value of R with components x and y is

$$p(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{(x-S)^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right]. \quad (\text{D.1})$$

If the noise processes are considered to be independent of the signal and to have a uniform phase power distribution, then this expression becomes, in polar coordinates,

$$p(R, \theta) = \frac{R}{2\pi\sigma^2} \exp\left[-\frac{R^2 - 2RS \cos \theta + S^2}{2\sigma^2}\right]. \quad (\text{D.2})$$

The probability of measuring a phase θ is simply $p(\theta) = \int p(R, \theta) dR$. By making the substitution $z = (R - S \cos \theta)/(\sigma\sqrt{2})$, this integral can be rearranged to give

$$p(\theta) = \sigma^2 \exp\left[-\frac{S^2 \cos^2 \theta}{2\sigma^2}\right] + \sqrt{2}\sigma S \cos \theta \int_{-\frac{S \cos \theta}{\sqrt{2}\sigma}}^{\infty} e^{-z^2} dz. \quad (\text{D.3})$$

The integral term on the right hand side is proportional to the complementary error function. However it is usual to express the result in terms of the error function itself, defined as

$$\text{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-z^2} dz. \quad (\text{D.4})$$

Equation D.3 can be further simplified by defining the signal-to-noise ratio in R as $\gamma = S/(\sigma\sqrt{2})$. The factor of $\sqrt{2}$ results from the combined contributions of both the x and y components of the noise, each with variance σ^2 . In this notation the probability density function for the phase fluctuations due to the noise becomes

$$p(\theta, \gamma) = \frac{1}{2\pi} e^{-\gamma^2} [1 + \gamma\sqrt{\pi} \cos \theta e^{\gamma^2 \cos^2 \theta} (1 + \text{erf}(\gamma \cos \theta))]. \quad (\text{D.5})$$

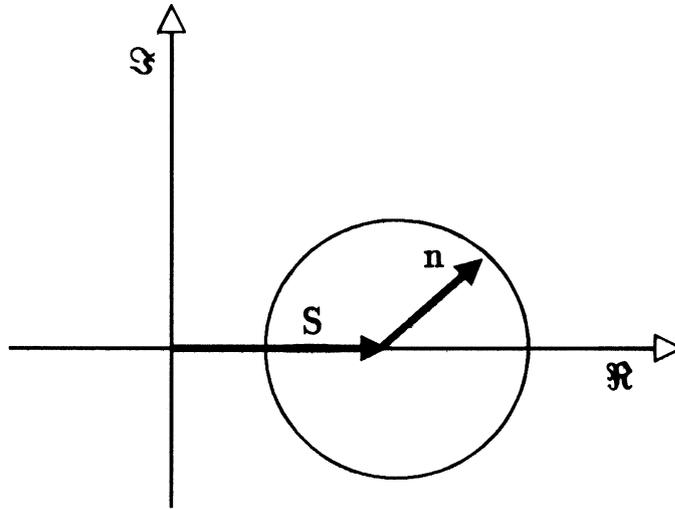


Figure D.1: *Complex noise, n , superimposed on a real signal, S .*

This standard result is displayed graphically in Figure D.2. When the signal-to-noise ratio is high the corresponding phase deviations are small, so the distribution approximates to a Gaussian with variance $1/(2\gamma^2)$. When the noise dominates, $p(\theta)$ flattens to an almost uniform distribution, with a residual bump proportional to $\gamma \cos \theta$.

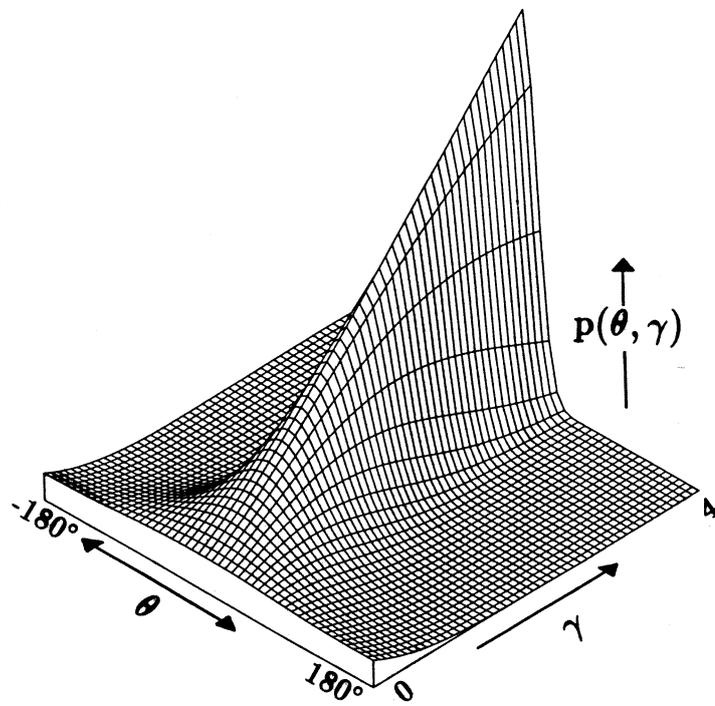


Figure D.2: The phase distribution for a measurement of a real signal within complex Gaussian noise. Note the change in character at $\gamma \simeq 1$.

Appendix E

Theoretical beam patterns

In Chapter 4 it is necessary to simulate the beam shapes of both the 4C and 3.6 ha arrays to carry out confusion simulations. A simple approach is adopted in which both arrays are modelled as perfect rectangular apertures, and the beams computed from their far-field diffraction patterns.

The response of a general two dimensional aperture to a point source in the direction \mathbf{s} is taken as

$$R(\mathbf{s}) \propto \int F(\mathbf{r}) e^{ik\mathbf{r}\cdot\mathbf{s}} d^2r, \quad (\text{E.1})$$

where $F(\mathbf{r})$ is the aperture function and k the wave-number of the radiation (Figure E.1). If the aperture is rectangular in the (x, y) plane, ($|x| \leq L/2, |y| \leq W/2$), the response to a source at an angle α to the x -axis and β to the y -axis will be

$$R(\alpha, \beta) \propto \text{sinc}\left(\frac{\pi L}{\lambda} \cos \alpha\right) \text{sinc}\left(\frac{\pi W}{\lambda} \cos \beta\right). \quad (\text{E.2})$$

If the x -axis is taken as W-E and the y -axis as S-N, the cosine angles can be related to azimuth(A) and altitude(a) of the source by

$$\begin{aligned} \cos \alpha &= \cos a \sin A \\ \cos \beta &= \cos a \cos A. \end{aligned} \quad (\text{E.3})$$

The equation is of more use if expressed in terms of the hour angle (H) and declination (δ) of the source. Using the standard first order transformation between the local horizon and equatorial coordinates systems [105], the array response becomes

$$R(H, \delta) = \text{sinc}\left(\frac{\pi L}{\lambda} \cos \delta \sin H\right) \text{sinc}\left(\frac{\pi W}{\lambda} (\sin \delta \cos \phi - \sin \phi \cos \delta \cos H)\right). \quad (\text{E.4})$$

This is the response of an E-W rectangular aperture at a geographical latitude ϕ that is pointing at the zenith. If the array were tilted away from the zenith along

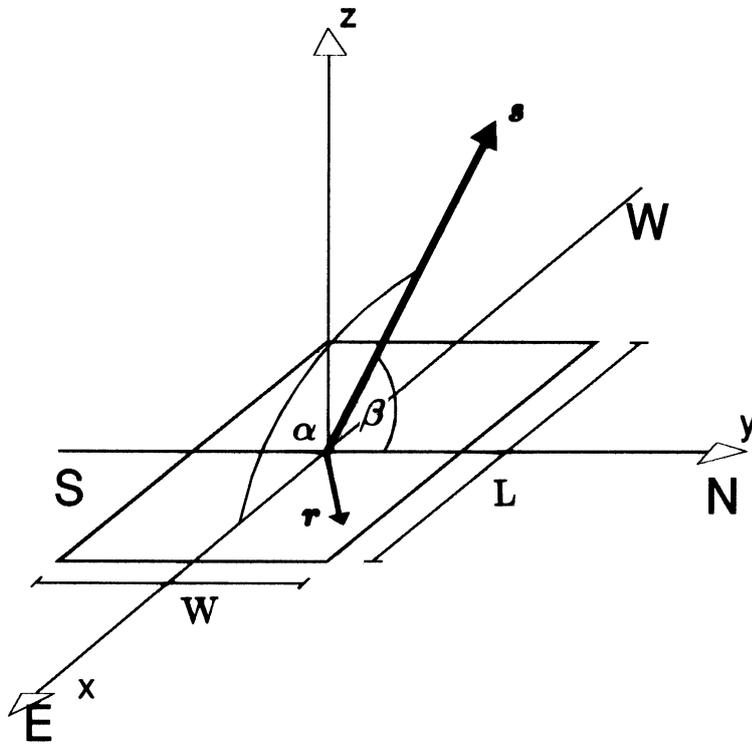


Figure E.1: The rectangular aperture described in the text together with the source position angles, α and β .

the meridian, ϕ would be its *effective* latitude, that is the latitude at which it would be pointing at the zenith if moved there without rotation. This is equivalent to identifying ϕ as the declination of the array.

In deriving Equation E.4, the polarisation properties of the radiation and the array have been ignored. In fact, both the 4C and 3.6-ha arrays are only sensitive to E-W linear polarisation, so there is an ‘obliquity’ factor of $1 - \cos^2 \delta \sin^2 H$ to be included in the equation to take account of its overall dipole anisotropy. In practice however, this is a small effect, and as long as the beam is compact it can be absorbed in the normalisation.

The 4C is a phased array. This means that a phase gradient may be imposed along its length to maximise its response to any particular point (H_0, δ_0) on the sky whilst keeping its physical orientation fixed. Equation E.4 can be modified to take account of this by adding a phase offset to the first sinc term, i.e.

$$R(H, \delta, H_0, \delta_0) = \text{sinc} \left(\frac{\pi L}{\lambda} (\cos \delta \sin H - \cos \delta_0 \sin H_0) \right) \times \text{sinc} \left(\frac{\pi W}{\lambda} (\sin \delta \cos \phi - \sin \phi \cos \delta \cos H) \right). \quad (\text{E.5})$$

Equation E.5, together with the following parameters, represents the model of the 4C used in the simulations:

$\lambda(\text{m})$	$L(\text{m})$	$W(\text{m})$	$\phi(\text{deg})$
3.68	235	19.8	35.5

(Note that only half the 470m long array was used for the observations.)

The model of the 3.6-ha array needs to be a little more sophisticated. Firstly, the aperture is divided into North and South portions, each of which can be modelled as 16-line diffraction gratings with line spacing d ($=2.39\text{m}$) and length L ($=470\text{m}$). These lines can be electrically phased to give an overall declination response that can be modified in steps of $\pi/16$. In fact the amplitude response of each section of the array is

$$R_i(H, \delta) = \text{sinc} \left(\frac{\pi L}{\lambda} \cos \delta \sin H \right) \frac{\sin 16\xi}{\sin \xi}, \quad (\text{E.6})$$

where

$$\xi = \frac{\pi d}{\lambda} (\sin \delta \cos \phi - \sin \phi \cos \delta \cos H) + (B - 10) \frac{\pi}{16} \quad (\text{E.7})$$

and B is the beam number [89]. B is chosen to obtain a maximum signal from the source at transit. In the case of 3C48, $B=7$. Secondly, the two halves of the array are phased together so that the sum of the two signals is a maximum for

3C48 at transit. This corresponds to a phase angle of $kl \sin \delta_0 \cos \phi$ between the two signals, where δ_0 is the declination of 3C48, ϕ the latitude of the instrument and l the distance between its North and South sections. Therefore the overall response of the phased-up array to a source at (H, δ) is

$$\begin{aligned}
 R(H, \delta) &= R_1(H, \delta) && \text{(E.8)} \\
 &+ R_2(H, \delta) \exp ikl[\cos \phi(\sin \delta - \sin \delta_0) \\
 &- \sin \phi(\cos \delta \cos H + \cos \delta_0 \cos H_0)] .
 \end{aligned}$$

Another factor may be added to take account of the blaze angle of the reflecting screens that sit behind each row in the array. These screens should give the whole array something like a sinusoidal dependence on declination, peaking at the ecliptic. However, it is thought that this is not the case [64]. Whatever the true overall envelope of the array, it is not terribly important in these calculations as they are only concerned with the comparatively narrow response of a single beam.

Appendix F

Simulation source list

The following tables contain the source parameters used in the confusion simulations.

No	4C	S/Jy	RA/hms	Dec/dms	No	4C	S/Jy	RA/hms	Dec/dms
1	1805	6.2	1 0 0	18 55 31	2	2503	9.7	1 0 11	25 36 23
3	5103	5.5	1 0 17	51 10 46	4	1404	28.0	1 0 17	14 36 34
5	4805	5.2	1 2 14	48 34 5	6	4204	9.1	1 3 52	42 14 8
7	3205	35.4	1 4 41	32 7 36	8	5301	6.4	1 7 28	53 44 21
9	3102	30.1	1 7 33	31 31 23	10	5503	9.7	1 8 12	55 32 52
11	1709	11.0	1 9 11	17 38 7	12	2504	6.2	1 8 41	25 49 38
13	2702	5.9	1 8 55	27 6 58	14	4702	4.6	1 9 3	47 8 5
15	4904	25.1	1 9 5	49 12 46	16	4101	9.6	1 9 26	41 39 21
17	3103	4.6	1 10 1	31 52 17	18	2902	6.2	1 10 39	29 42 48
19	5402	8.2	1 11 12	54 19 42	20	4806	13.2	1 11 25	48 8 2
21	4304	4.6	1 12 3	43 16 37	22	1504	6.8	1 13 24	15 24 8
23	4005	5.2	1 13 32	40 1 33	24	4503	21.1	1 15 2	45 20 39
25	3502	5.5	1 15 2	35 43 19	26	4603	6.8	1 15 24	46 56 20
27	3104	5.5	1 16 46	31 55 6	28	3704	6.2	1 19 5	37 43 51
29	4905	9.2	1 19 43	49 51 9	30	4205	4.8	1 20 17	43 5 38
31	4006	7.1	1 20 31	40 31 26	32	2304	8.4	1 21 7	23 56 15
33	3206	20.4	1 23 55	32 57 33	34	2505	9.5	1 23 56	25 43 29
35	1806	13.9	1 24 11	18 57 46	36	4906	8.7	1 24 16	49 30 48
37	2305	7.3	1 24 25	23 56 15	38	3404	5.5	1 24 25	34 40 8
39	3207	26.0	1 25 12	32 28 52	40	2804	35.0	1 25 42	28 47 41
41	2107	4.6	1 25 44	21 4 18	42	1905	9.6	1 26 19	19 3 8
43	2306	21.3	1 27 25	23 23 11	44	2006	5.2	1 27 25	20 41 3
45	2506	8.7	1 27 25	25 53 3	46	2507	13.1	1 28 39	25 5 16
47	2703	4.6	1 29 3	28 5 43	48	1906	5.0	1 30 25	19 57 48
49	2402	11.9	1 30 46	24 12 42	50	2704	11.4	1 30 40	27 31 42

Table F.1: Confusion simulation sources, $1^{\text{h}}0^{\text{m}}0^{\text{s}} \leq \text{RA} \leq 1^{\text{h}}30^{\text{m}}40^{\text{s}}$. The position epoch is 1950.0.

No	4C	S/Jy	RA/hms			Dec/dms			No	4C	S/Jy	RA/hms			Dec/dms		
51	3805	5.5	1	30	47	38	7	50	52	4604	8.7	1	30	50	46	42	5
53	3705	23.5	1	32	33	37	38	57	54	5203	5.2	1	32	35	52	25	56
55	2007	64.4	1	33	39	20	42	17	56	2903	19.8	1	34	4	29	58	57
57	3208	83.7	1	34	49	32	54	21	58	3806	5.9	1	34	53	38	36	21
59	3105	5.5	1	35	16	31	17	15	60	2202	14.8	1	35	40	22	45	29
61	1603	6.4	1	36	20	16	44	24	62	3904	29.5	1	36	32	39	42	1
63	3405	5.2	1	37	58	34	35	2	64	3209	5.5	1	39	18	32	53	46
65	3807	11.9	1	40	34	38	47	4	66	3303	15.0	1	41	22	33	56	54
67	2604	5.0	1	42	5	27	29	48	68	2705	9.8	1	43	36	27	50	49
69	4404	10.6	1	43	53	44	40	51	70	4807	5.7	1	44	28	48	34	36
71	2108	8.1	1	44	40	21	13	9	72	4305	5.9	1	44	54	43	4	54
73	5302	22.8	1	45	13	53	17	34	74	2605	10.4	1	47	14	26	49	54
75	2008	12.9	1	48	1	20	2	40	76	4703	7.9	1	48	34	47	32	40
77	1604	9.3	1	50	23	16	26	33	78	4007	4.6	1	50	49	40	40	34
79	4504	4.8	1	51	46	45	45	9	80	2606	4.6	1	52	3	26	8	30
81	2403	5.9	1	52	13	24	11	33	82	2203	4.6	1	52	13	22	14	15
83	4704	4.8	1	52	18	47	7	15	84	4306	18.7	1	52	25	43	31	16
85	2904	7.1	1	53	6	29	40	58	86	4102	11.6	1	53	18	41	47	49
87	2009	8.4	1	53	47	20	48	50	88	2805	45.3	1	54	18	28	37	9
89	3106	9.3	1	54	23	31	39	13	90	5303	11.6	1	55	54	53	14	54
91	4008	23.1	1	57	4	40	35	27	92	5304	11.6	1	57	13	53	5	37
93	4405	8.3	1	57	31	44	12	38	94	3905	5.9	1	57	46	40	33	1
95	3304	6.8	1	57	49	33	24	6	96	3107	5.5	1	58	1	31	32	16
97	2706	5.2	1	58	37	27	19	47	98	2905	16.9	1	58	43	29	19	18
99	1807	8.0	1	58	54	18	23	18	100	4808	16.3	1	59	49	48	4	11
101	1907	15.5	2	0	41	19	37	30	102	5840	17550	23	21	10	58	32	30
103	4040	16118	19	54	43	40	35	50									

Table F.2: *Confusion simulation sources, $1^{\text{h}}30^{\text{m}}47^{\text{s}} \leq \text{RA} \leq 2^{\text{h}}0^{\text{m}}41^{\text{s}}$. The position epoch is 1950.0.*

Appendix G

Co-ordinate transformations

Ordnance Survey maps present positional information, in Transverse Mercator projection, as metres North and East of a primary reference point off the South West coast of Britain. Heights are in metres above the local geoid (loosely ‘sea level’). The eastings and northings may be transformed into latitude (l) and longitude (Λ , East positive) on a reference ellipsoid (in this case the Airy ellipsoid) by a numerical routine presented in ‘*The History of the Retriangulation of Great Britain, 1935-62*’, published by the Ordnance Survey. A point $P(l, \Lambda, h)$ referred to this ellipsoid has Cartesian co-ordinates (x, y, z) specified by

$$\begin{aligned}x &= (\nu + h) \cos l \cos \Lambda \\y &= (\nu + h) \cos l \sin \Lambda \\z &= (\nu(1 - \epsilon^2) + h) \sin \phi ,\end{aligned}\tag{G.1}$$

where ν is the local radius of curvature of the ellipsoid $a(1 - \epsilon^2 \sin^2 l)^{-1/2}$, ϵ^2 its eccentricity (0.00667054) and a its semi-major axis (6377563.3964 m). Thankfully the planes of the geoid and the Airy ellipsoid coincide to within about 1 m in East Anglia, so a height above the ellipsoid (h) can be assumed equal to height above the geoid.

Once in Cartesian coordinates, the transformations become straightforward and have been well charted by others (see for example Thompson Moran and Swenson [91] Chapter 4). A baseline (D_x, D_y, D_z) has a declination, hour angle and length in the equatorial system defined by

$$\delta_B = \arctan \left[\frac{D_z}{\sqrt{D_x^2 + D_y^2}} \right]$$

$$\begin{aligned}
h_B &= \arctan \left[\frac{D_y}{D_x} \right] \\
D &= (D_x^2 + D_y^2 + D_z^2)^{1/2}.
\end{aligned} \tag{G.2}$$

These are related to its (u, v, w) ('sky') co-ordinates in and above the plane perpendicular to the source direction (δ, H) by

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = D \begin{bmatrix} \cos \delta_B \sin(H - H_B) \\ \sin \delta_B \cos \delta - \cos \delta_B \sin \delta \cos(H - H_B) \\ \sin \delta_B \sin \delta + \cos \delta_B \cos \delta \cos(H - H_B) \end{bmatrix}. \tag{G.3}$$

The source hour angle H is taken as being measured with respect to the Greenwich meridian rather than the local meridian.

Lastly, a small change $(\mathcal{E}, \mathcal{N})$ in the eastings and northings of a point corresponds to a change in its Cartesian co-ordinates of (dx, dy, dz) where

$$\begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} \simeq \begin{bmatrix} -\sin \Lambda & -\sin l \cos \Lambda \\ \cos \Lambda & -\sin l \sin \Lambda \\ 0 & \cos l \end{bmatrix} \begin{bmatrix} \mathcal{E} \\ \mathcal{N} \end{bmatrix}. \tag{G.4}$$

This is useful for determining the effect of small baseline errors on the interferometer's phase.

Appendix H

Pulsar data acquisition system

The data acquisition system used for the pulsar search described in Chapter 7 is shown schematically in Figure H.1. Bandlimited analogue signals are fed to an 8-bit ADC whose clock and sample frequencies are derived from a divider chain driven by a 10 MHz rubidium standard, and the digitised data are clocked at a constant rate into a FIFO daisy chain. The FIFO is read in 1024 byte blocks by a Gemini microcomputer, the software for which was written by Dr. P.J. Duffett-Smith, and the data stored on floppy disc. This buffer system is necessary maintain a constant data acquisition rate, even when the host computer is otherwise engaged.

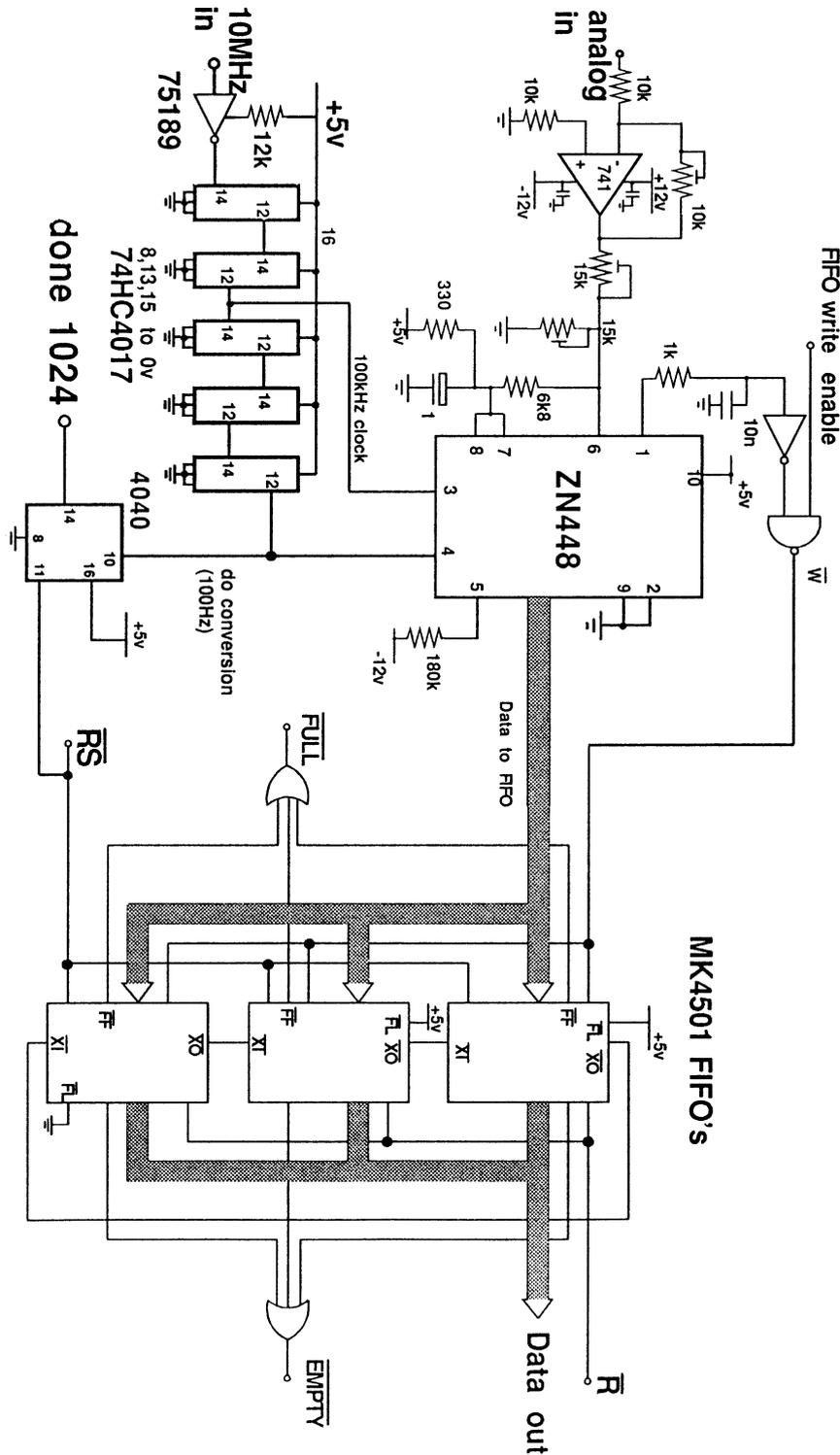


Figure H.1: Data acquisition system. The FIFO chain accepts data at a constant rate even when being read by the host computer.

Appendix I

Pulsar search sensitivity

An analysis of the sensitivity achieved by the pulsar search (Chapter 7) can be conveniently divided into two sections – the relationship between the pulsed flux and the signal-to-noise ratio in the time series at the output of the telescope, and the sensitivity of the subsequent processing (i.e. what minimum value of signal-to-noise ratio can be detected with a reasonable confidence).

I.1 Flux calibration

Consider a voltage signal from the antenna of the form

$$v(t) = s(t) + n(t) , \quad (\text{I.1})$$

where $s(t)$ represents a noise-like voltage due to the pulsar and $n(t)$ represents a noise-like system voltage, which in this case is mostly due to the extended emission from the remnant. The voltage v is squared and low-pass filtered giving a post-detection signal of

$$r(t) = \langle v^2(t) \rangle . \quad (\text{I.2})$$

The noise power spectrum of $r(t)$ defining the signal-to-noise ratio of the pulsar in r , is best approached via its autocorrelation, $\rho_r(\tau)$. It is straightforward to show that

$$\rho_r(\tau) = 2\langle (s(t) + n(t))(s(t - \tau) + n(t - \tau)) \rangle^2 + \rho_v^2(0) . \quad (\text{I.3})$$

The last term in this expression is a constant over the integration time scale, and represents the square of the total power received by the telescope. The slowly varying component to this is the power due to the pulsar, the remainder, the unchanging power from from the remnant, is removed by a high-pass filter with a cutoff at 0.1 Hz.

The ‘signal’ is therefore the power received from the pulsar alone. If the frequency response of the receiver is represented by $H(\nu)$, this signal can be expressed in terms of an antenna temperature as

$$\text{pulsar signal} = kT_A \int |H(\nu)^2| d\nu . \quad (\text{I.4})$$

The first term in Equation I.3 represents the noise on the measurement of the pulsar signal. Assuming that the pulsar power is a lot less than the noise power, the term reduces to the square of the autocorrelation of the original noise, $\rho_n^2(\tau)$. The Fourier transform of this is the noise power spectrum after detection:

$$\rho_n^2(\tau) \Leftrightarrow k^2 T_S^2 |H(\nu)|^2 \otimes |H(\nu)|^2 , \quad (\text{I.5})$$

where T_S is the system temperature due to this noise, \Leftrightarrow represents a Fourier transform relationship and \otimes denotes convolution. This noise spectrum is now low-pass filtered with a frequency response $H_L(\nu)$. Assuming that the width of this filter is a lot less than the characteristic width of the noise power spectrum, the total noise power surviving the filter may be approximated by

$$\text{noise power} = 2k^2 T_S^2 \cdot 2\Delta\nu_{LF} \int |H(\nu)|^4 d\nu , \quad (\text{I.6})$$

where $\Delta\nu_{LF}$ is the noise equivalent bandwidth of the low-pass filter, i.e.

$$\Delta\nu_{LF} = \frac{1}{|H_L(0)|^2} \int_0^\infty |H_L(\nu)|^2 d\nu . \quad (\text{I.7})$$

The signal-to-noise ratio in the detected, filtered data is therefore

$$\gamma_f = \frac{kT_A \int |H(\nu)|^2 d\nu}{\sqrt{4k^2 T_S^2 \Delta\nu_{LF} \int |H(\nu)|^4 d\nu}} . \quad (\text{I.8})$$

If the receiver is assumed to have a rectangular RF passband, this expression reduces to

$$\gamma_f = \frac{T_A}{T_S \sqrt{2}} \sqrt{\frac{\Delta\nu_{RF}}{\Delta\nu_{LF}}} . \quad (\text{I.9})$$

Measurements of the noise equivalent bandwidth of the low pass filter show it to equal almost exactly 50 Hz. The signal-to-noise ratio can therefore be freely written in terms of the sampling period, $\tau=100$ Hz, to give

$$\gamma_f = \frac{T_A}{T_S} \sqrt{\Delta\nu_{RF} \tau} . \quad (\text{I.10})$$

If it is assumed that T_S is entirely due to the bulk emission from the remnant, $T_A/T_S = S_{\text{pulsar}}/S_{\text{remnant}}$. In fact the system temperature due to the electronics of the OMT is ~ 185 K, and that due to Cas is ~ 474 K. Here then is the first result – a relationship between the signal-to-noise ratio in the time series, the system temperature and the flux from the pulsar.

I.2 Minimum detectable signal-to-noise ratio

The detected and filtered data considered in the previous section are recorded to make a time series of length M , which is then Fourier transformed. Assuming the complex frequency data contains only noise of variance σ^2 in each of its real and imaginary components, the probability distribution characterising the *power* spectrum, $y(\nu)$ (rather than the amplitude spectrum, which is Rayleigh distributed) is

$$p(y) = \frac{1}{2\sigma^2} \exp\left[-\frac{y}{2\sigma^2}\right], \quad y \geq 0. \quad (\text{I.11})$$

Note that y represents any *single* frequency channel in the Fourier transform. The probability distribution for the sum of N such power measurements can be evaluated from the characteristic function of $p(y)$, $\tilde{p}(k) = (1 - 2\sigma^2 ik)^{-1}$, as

$$p_N(y) = \frac{1}{(2\pi)^{N/2}} \int_{-\infty}^{\infty} \tilde{p}^{-N}(k) e^{-iky} dk, \quad (\text{I.12})$$

a contour integration of which gives

$$p_N(y) = \frac{(2\sigma^2)^{-N}}{(N-1)!} y^{N-1} \exp\left[-\frac{y}{2\sigma^2}\right]. \quad (\text{I.13})$$

This is a χ^2 distribution with $2N$ degrees of freedom.

It is now necessary to evaluate the cumulative probability, $P_N(y_0)$, that this sum of N numbers is less than some total, y_0 . A straightforward integration gives

$$P_N(y_0) = p_N(y < y_0) = 1 - e_{N-1} \left(\frac{y_0}{2\sigma^2}\right)^{N-1} \exp\left[-\frac{y_0}{2\sigma^2}\right], \quad (\text{I.14})$$

where

$$e_{N-1}(x) = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^{N-1}}{(N-1)!}.$$

Until now, the statistics of only one frequency ‘bin’ in the Fourier transform have been considered. If a power level is to be set that corresponds to a detection of a periodic signal, it must be sufficiently high for there to be little chance that *any* bin reaches it by random means. There are about half a million bins in the transform, so this level will clearly be a lot higher than two or three standard deviations for any one bin. The *joint* cumulative probability that none of the $M/2$ bins contain a signal greater than y_0 is simply $[P_N(y_0)]^{M/2}$, which also equals the probability that such a feature is not an artifact of the noise process. If M is large, this confidence, c , approximates to

$$c = 1 - \frac{M}{2} e_{N-1} \left(\frac{y_0}{2\sigma^2}\right)^{N-1} \exp\left[-\frac{y_0}{2\sigma^2}\right]. \quad (\text{I.15})$$

So here is the second result. Working in units of the variance of the original time series ($2\sigma^2$), the 5% detection level ($c = 0.95$) over $M/2$ bins and after summing N spectra is the value of y_d that satisfies

$$\frac{0.05}{M/2} = e^{-y_d} \sum_{n=0}^{N-1} \frac{y_d^n}{n!}. \quad (\text{I.16})$$

This detection level can be converted directly into a signal-to-noise ratio which in turn can be related to the signal-to-noise ratio in each of the constituent power spectra, γ_{F^2} , (assuming they are equal) to give

$$\sqrt{N}\gamma_{F^2} = \frac{y_d - N}{\sqrt{N}}. \quad (\text{I.17})$$

Assuming the noise on the original time series is Gaussian, the signal-to-noise ratio in the power spectrum, γ_{F^2} , is related to that of the complex spectrum, γ_F , by

$$\gamma_{F^2} = \frac{\gamma_F^2}{\sqrt{1 + 2\gamma_F^2}}. \quad (\text{I.18})$$

Furthermore, taking a single spectral component of the pulse, the Fourier transform of a time series with M samples and signal-to-noise ratio γ_f has itself a signal-to-noise ratio of

$$\gamma_F = \frac{\gamma_f}{2} \sqrt{M}. \quad (\text{I.19})$$

These equations, together with Equation I.10, give the calibrated search sensitivity in terms of janskys per pulse harmonic. Of course the sensitivity of the next stage in the search, in which these harmonics are summed incoherently, depends on the pulse profile.

References

- [1] Armstrong, J.W., Woo, R., Estabrook, F.B.: 1979, *Astrophys. J.*, **230**, 570.
- [2] Ashworth, W.B.: 1980, *J. Hist. Astr.*, **9**, 1.
- [3] Baars, J.W.M., Genzel, R., Pauliny-Toth, I.I.K., Witzel, A: 1977, *Astron. Astrophys.*, **61**, 99.
- [4] Baldwin, J.E., Warner, P.J.: 1979, *Image Formation from Coherence Functions in Astronomy*, C. van Schoonveld, Ed. D. Reidel, Dordrecht, Holland.
- [5] Bare, C., Clarke, B.G., Kellerman, K.I., Cohen, M.H., Jauncey, D.L.: 1967, *Science*, **157**, 189-191.
- [6] Barthel, P.D.: 1987, *Superluminal Radio Sources*, p. 148, Eds. Zensus, J.A., Pearson, T.J., C.U.P.
- [7] Bates, R.H.T., Mnyama, D.: 1986, *Adv. in Electronics and Electron Phys.*, **67**, 1.
- [8] Bell, A.R., Gull, S.F., Kenderdine, S.: 1975, *Nature*, **257**, 463.
- [9] Bell, A.R.: 1977, *Ph.D Thesis, University of Cambridge*, .
- [10] Boroson, T.A., Oke, J.B.: 1982, *Nature*, **296**, 397.
- [11] Bowers, F.K., Klinger, R.J.: 1974, *Astron. Astrophys. Suppl.*, **15**, 373.
- [12] Bovkoon, V.P., Braude, S.Ya, Megn, A.V.: 1982, *Astr. Sp. Sci.*, **81**, 221.
- [13] Braun, R.: 1987, *Astron. Astrophys.*, **171**, 233.
- [14] Braun, R., Gull, S.F., Perely, R.A.: 1987, *Nature*, **327**, 395.
- [15] Bridle, A.H.: 1967, *Mon. Not. R. astr. Soc.*, **136**, 219.
- [16] Chevalier, R.A., Emmering, R.T.: 1986, *Ap. J.*, **304**, 140.

- [17] Contini, M.: 1987, *Astron. Astrophys.*, **183**, 53.
- [18] Cordey, R.A.: 1986, *Mon. Not. R. astr. Soc.*, **219**, 575.
- [19] Cornwell, T.J.: 1987, *Astron. Astrophys.*, **180**, 269.
- [20] Cornwell, T.J., Wilkinson, P.N.: 1981, *Mon. Not. R. astr. Soc.*, **196**, 1067.
- [21] Cowsik, R., Sarkav, S.: 1983, *Supernova remnants and their X-ray emission*, IAU symp. 101, p. 187, Eds. J. Danziger, P.Gorenstein.
- [22] Crane, R.K.: 1977, *Proc IEEE*, **65**, 180.
- [23] D'Addario, L.R.: 1984, *VLBA Memo. No. 332*, National Radio Astronomy Observatory, Charlottesville, Virginia.
- [24] Davies, J.G., Large, M.I.: 1970, *Mon. Not. R. astr. Soc.*, **149**, 301.
- [25] DuCastle, F., Faynot, J.M.: 1964, *Nature*, **204**, 984.
- [26] Duffett-Smith, P.J.: 1980, *Mon. Not. R. astr. Soc.*, **192**, 33.
- [27] El Eid, M.F., Langer, N.: 1986, *Prog. in Part. and Nucl. Phys.*, **17**, 295.
- [28] Erickson, W.C., Perely, R.A.: 1975, *Astrophys. J.*, **200**, L83.
- [29] Erickson, W.C., Mahoney, M.J.: 1985, *Astrophys. J.*, **290**, 596.
- [30] Fabian, A.C., Willingale, R., Pye, J.P., Murray, S.S., Fabbiano, G.: 1980, *Mon. Not. R. astr. Soc.*, **193**, 175.
- [31] Fitt, A.J., Alexander, P., Cox, M.J.: 1988, *Mon. Not. R. astr. Soc.*, **233**, 907.
- [32] Fort, D.N., Yee, H.K.C.: 1976, *Astron. Astrophys.*, **50**, 19.
- [33] Francis, S.H.: 1975, *J. Atmos. Terr. Phys.*, **37**, 1011.
- [34] Gapper, G.R., Hewish, A.: 1981, *Mon. Not. R. astr. Soc.*, **197**, 209.
- [35] Gehren, T., Fried, J., Wehinger, P.A., Wyckoffs, S.: 1984, *Ap. J.*, **278**, 11.
- [36] Green, D.A.: 1984, *Ph.D. Thesis, University of Cambridge*.
- [37] Green, D.A.: 1988, IAU coll. 101, Eds. Roger and Landecker, C.U.P.
- [38] Gull, S.F.: 1973, *Mon. Not. R. astr. Soc.*, **161**, 47.
- [39] Gull, S.F., Skilling, J: 1984, *Indirect Imaging*, Ed. J.A. Roberts, C.U.P.

- [40] Hartas, J.S.: 1983, *Ph.D. Thesis, University of Cambridge*.
- [41] Hartas, J.S., Rees, W.G., Scott, P.F., Duffett-Smith, P.J.: 1983, *Mon. Not. R. astr. Soc.*, **205**, 625.
- [42] Helfund, D.J.: 1983, *Supernova remnants and their X-ray emission*, IAU symp. 101, p. 471, Eds. J. Danziger, P. Gorenstein.
- [43] Hewish, A.: 1952, *Proc. R. Soc. Lond. A*, **214**, 494-514.
- [44] Hewish, A., Scott, P.F., Wills, D.: 1964, *Nature*, **203**, 1214.
- [45] Hunsucker, R.D.: 1971, *Radio Sci.*, **6**, 535.
- [46] Hunsucker, R.D.: 1982, *Rev. Geophys. and Sp. Phys.*, **20**, 293.
- [47] Hutton, L.K., Clark, T.A., Erickson, W.C., Resch, G.M., Vandenberg, N.R., Knowles, S.H., Youmans, A.B.: 1974, *Astr. J.*, **79**, 1248.
- [48] Jennison, R.C.: 1958, *Mon. Not. R. astr. Soc.*, **118**, 276.
- [49] Klein, U., Emerson, D.T., Haslam, C.G.T.: 1979, *Astron. Astrophys.*, **76**, 120.
- [50] Kühr, H., Nauber, U., Pauliny-Toth, I.I.K., Witzel, A.: *Max-Planck-Institut für Radioastronomie*, preprint number 55.
- [51] Lamb, S.A.: 1978, *Astrophys. J.*, **220**, 186.
- [52] Little, L.T., Hewish, A.: 1966, *Mon. Not. R. astr. Soc.*, **134**, 221.
- [53] Lyne, A.G.: 1988, *NATO Adv. Workshop on Gravitational Wave Data Analysis*, Eds. Krolak and Schutz, Reidel, Dordrecht.
- [54] Manchester, R.N., Taylor, J.H.: 1977, *Pulsars*, San Fransisco, Freeman.
- [55] Morgan, M.G., Calderon, C.H.J., Ballard, K.A.: 1978, *Radio Sci.*, **13**, 729-741.
- [56] Napier, P.J., Crane, P.C.: 1982, *Synthesis Mapping*, Proc. NRAO-VLA Workshop Ch. 5, Eds. Thompson, A.R., D'Addario, L.R., NRAO Green Bank.
- [57] Narayan, R.: 1987, *Interferometric Imaging in Astronomy*, ESO, NOAO Workshop.
- [58] Neugerbauer, G., Sofier, B.T., Miley, G.K.: 1985, *Ap. J.*, **295**, L27.
- [59] Papoulis, A.: 1965, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York.

- [60] Pearson, T.J., Readhead, A.C.S.: 1984, *Ann. Rev. Astron. Astrophys.*, **22**, 97-130.
- [61] Perely, R.A., Erickson, W.C.: 1984, *VLA Scientific Memo*. No 146, National Radio Astronomy Observatory, Charlottesville, Virginia.
- [62] Picquenard, A: 1974, *Radio wave propagation*, Macmillan.
- [63] Prószyński, M., Przybycień, D.: 1984, *Proc. Green Bank Workshop on Millisecond Pulsars*, p. 151, Eds. S.P. Reynolds, D.R. Stinebring.
- [64] Purvis, A.: 1981, *Ph.D. Thesis, University of Cambridge*.
- [65] Purvis, A., Tappin, S.J., Rees, W.G., Hewish, A., Duffett-Smith, P.J.: 1987, *Mon. Not. R. astr. Soc.*, **229**, 589.
- [66] Read, P.L.: 1977, *Mon. Not. R. astr. Soc.*, **178**, 259.
- [67] Read, P.L.: 1977, *Mon. Not. R. astr. Soc.*, **181**, 63p.
- [68] Readhead, A.C.S., Hewish, A.: 1974, *Mem. R. astr. Soc.*, **78**, 1.
- [69] Readhead, A.C.S., Kemp, M.C., Hewish, A.: 1978, *Mon. Not. R. astr. Soc.*, **185**, 207.
- [70] Readhead, A.C.S., Wilkinson, P.N.: 1978, *Ap. J.*, **223**, 25.
- [71] Readhead, A.C.S., Walker, R.C., Pearson, T.J., Cohen, M.H.: 1980, *Nature*, **285**, 137.
- [72] Rees, W.G.: 1984, *Ph.D. Thesis, University of Cambridge*.
- [73] Rees, W.G., Duffett-Smith, P.J.: 1983, *Mon. Not. R. astr. Soc.*, **205**, 759.
- [74] Rees, W.G., Duffett-Smith, P.J.: 1985, *Mon. Not. R. astr. Soc.*, **212**, 463.
- [75] Riley, J.M., Warner, P.J., Rawlings, S., Saunders, R., Pooley, G.G.: 1989, *Mon. Not. R. astr. Soc.*, **236**, 13p.
- [76] Rogers, A.E.E., Hinteregger, H.F., Whitney, A.R., Counselman, C.C., Shapiro, I.I., *et al.* : 1974, *Ap.J.*, **193**, 293.
- [77] Ryle, M.J.: 1960, *Inst. Elec. Eng.*, **6**, 14.
- [78] Schwab, F.R.: 1980, *Proc. Soc. Photo-Opt. Instrum. Eng.*, **231**, 18.
- [79] Schwab, F.R., Cotton, W.D.: 1983, *Astron. J.*, **88**, 5, 688.

- [80] Seiradakis, J.H., Graham, D.A.: 1980, *Astron. Astrophys.*, **85**, 353.
- [81] Seward, F.D., Zhen-Ru Wang: 1988, *Ap. J.*, **332**, 199.
- [82] Sivia, D.S.: 1987, *Ph.D. Thesis, University of Cambridge*.
- [83] Smith E.K. Jr., Weintraub, S.: 1953, *Proc. IRE*, **41**, 1035-1037.
- [84] Smith, F.G, Disney, M.J., Hartley, K.F., Jones, D.H.P., King, D.J., Wellgate, G.B., Manchester, R.N., Lyne, A.G., Goss, W.M., Wallace, P.T., Peterson, B.A., Murdin, P.G., Danziger, I.Z.: 1978, *Mon. Not. R. astr. Soc.*, **184**, 39p.
- [85] Spinks, M.J.: 1985, *Ph.D. Thesis, University of Cambridge*.
- [86] Spinks, M.J., Rees, W.G., Duffett-Smith, P.J.: 1986, *Nature*, **319**, 471.
- [87] Staelin, D.H.: 1969, *Proc. IEEE*, **57**, 724.
- [88] Tan, S.M.: 1986, *Ph.D. Thesis. University of Cambridge*.
- [89] Tappin, S.J.: 1984, *Ph.D. Thesis, University of Cambridge*.
- [90] Thompson, A.R., D'Addario, L.R.: 1982, *Radio Sci.*, **17**, 357.
- [91] Thompson, A.R., Moran, J.M., Swenson, G.W., Jr.: 1986, *Interferometry and Synthesis in Radio Astronomy*, Wiley-Interscience.
- [92] Toman, K.: 1976, *Radio Sci.*, **11**, 107-119.
- [93] Tsien, S.C., Duffett-Smith, P.J.: 1982, *Mon. Not. R. astr. Soc.*, **198**, 941.
- [94] Tuffs, R.J.: 1983, *Ph.D. Thesis, University of Cambridge*.
- [95] Tuffs, R.J.: 1986, *Mon. Not. R. astr. Soc.*, **219**, 13.
- [96] Twiss, R.W., Carter, A.W.L., Little, A.G.: 1960, *Observatory*, **80**, 153.
- [97] Van den Bergh, S.: 1971, *Astrophys. J.*, **165**, 457.
- [98] Van den Bergh, S., Kamper, K.: 1983, *Astrophys. J.*, **268**, 129.
- [99] Van Vleck, J.H., Middleton, D.: 1966, *Proc. IEEE*, **54**, 2-19.
- [100] Vasiljev, M.Yu, Volodin, Yu V., Gubanov, A.G., Dagkesamanskaja, R.D., Dagkesamanskii, G.I., Dobish, G.I., Ivanov, V.V., Izekov, B.K., Sukhodolskii, S.A., Frolov, V.A.: 1976, *Moscow Acad. Sci.*, preprint number 179.
- [101] Vasseur, G., Waldtenfel, P.: 1969, *J. Atmos. Terr. Phys.*, **31**, 885.

- [102] Walczowski, L.T., Smith, K.L.: 1985, *Mon. Not. R. astr. Soc.*, **212**, 27p.
- [103] Wilkinson, P.N.: 1983, *Proc. Int. Conf. VLBI Tech.*, Toulouse, Cent. Natl. Etude. Spat., p. 375.
- [104] Wilkinson, P.N., Spencer, R.E., Readhead, A.C.S., Pearson, T.J., Simon, R.S.: 1984, *VLBI and Compact Radio Sources*, IAU Symposium 110, p. 25, Eds. R. Fanti *et al.*
- [105] Woolard, E.W., Clemence, G.M.: 1966, *Spherical Astronomy*, Academic Press.