A Study of Symmetric-Pair Antenna Arrays for Direction Finding and Communications

by

Wolfram Albert Ulrich Titze

A thesis submitted to the University of London for the Degree of

Doctor of Philosophy in Electronic Engineering

Department of Electronic & Electrical Engineering

UNIVERSITY COLLEGE LONDON

October 1992

ProQuest Number: 10044413

ProQuest.

ProQuest 10044413

Ich muß immer das Geschriebene wieder lesen, um mich zu überzeugen, daß es sich nicht so schwerfällig liest wie schreibt. [...] Manchmal denke ich, es ist Eitelkeit und Pedanterie, weil man die Nebensachen, die auch Mühe gemacht haben, nicht unerwähnt lassen will. Ich tröste mich damit, es sei auch wieder Gewissenhaftigkeit, indem man auch den Leser selbst urteilen lassen möchte.

Heinrich Rudolf Hertz, 1857-1894

(From a letter to his parents dated 15th January 1888)

Abstract

A novel type of phased array, called a symmetric-pair array, is described. Its antenna elements are arranged in pairs around a common phase centre, which yields easy-to-process biphase output signals from each pair. Symmetric-pair arrays are attractive for applications such as direction finding and beam forming and, due to the nature of the pair output signals, it is also possible to achieve gain with isotropic coverage on reception.

The properties of symmetric-pair arrays are examined, both for ideal conditions and in the presence of phase and amplitude errors. Graphical representations of the pair output signals are introduced to aid in the understanding of the properties of this type of antenna array.

The necessary algorithms and techniques to perform unambiguous direction finding with symmetric-pairs are developed. A detailed outline of a symmetric-pair direction finding system is given, and different options to process its baseband signals are discussed. The performance of the direction finder in low signal-to-noise environments is analysed using the mathematics of random processes, and the probability of detection as a function of false-alarm rate and angle of arrival is evaluated.

Beam forming techniques for symmetric-pair arrays are explained, and it is shown that this type of array allows savings in the control hardware due to its inherent symmetry. The problem of gain with isotropic coverage for reception is discussed, and it is demonstrated how it can be achieved with symmetric-pair arrays.

Computer simulations and experimental results are used throughout the thesis to confirm and support the presented theory.

Meinen Eltern,

die mich in jeglicher Hinsicht unterstützten.

ACKNOWLEDGEMENTS

CONTENTS

LIST OF PRINCIPAL SYMBOLS

SYMBOL	DEFINITION	UNIT
a_x	Wave incident upon port x	$\sqrt{W}$
A_x	Peak signal amplitude	V
b_x	Wave leaving port x	$\sqrt{W}$
b/w	Fractional bandwidth	
B_{IF}	Intermediate frequency bandwidth	Hz
B_n	Noise bandwidth	Hz
C	Scattering matrix, representing the imperfections of the antenna system	
d	Element spacing of a symmetric-pair	m
f	Frequency	Hz
F	Scattering matrix, representing the errors of the feed network	
f_c	Carrier frequency, centre frequency	Hz
f_{xy}	Element of the feed network error matrix	
F_{xy}	Submatrix of matrix F	
$F_X(x)$	Cumulative distribution function [$P(X \leq x)$]	
G	Gain of an antenna element,	dBi
	Gain of the array in isotropic mode	dBi
G_{AGC}	Linear gain of the AGC amplifier	
$G_{S/N}$	Gain in signal-to-noise ratio	dB
$g(\theta)_{xy}$	Element of the element pattern matrix	
G(f)	Power spectral density	W/Hz
$G(\theta)$	Scattering matrix, representing the element pattern of the array for the angle of incidence θ	
$G(\theta)_{xy}$	Submatrix of matrix G	
h_{xy}	Element of the hybrid matrix	
H	Scattering matrix of a hybrid	

I	Identity matrix	
l	Distance between two points on the surface of a sphere	m
l_x	Direction cosine	
L	Voltage loss relative to co-phasal summation	
L	Loss factor	dB
m	expectation, mean, first moment	
M	Scattering matrix, representing the mutual coupling between the antenna elements	
m_x	Direction cosine	
m_{xy}	Element of the mutual coupling matrix	
$\mathbf{M}_{xy}$	Submatrix, of matrix **M**	
n	Number of bits, Number of integrated samples	
N	Noise power, Number of elements	W
n_x	Direction cosine	
$n_x(t)$	Instantaneous noise voltage	V
p	Number of pairs in a symmetric-pair array	
p_c	Number of pairs for a collinear unambiguous arrangement	
P_d	Probability of detection	
P_{fa}	Probability of a false-alarm	
p(x)	Probability density function of variate X	
p(x,y)	Bivariate probability density function	
P(X<b)	Probability that variate X is smaller than b	
R	Scattering matrix, representing the imperfections of the reference signal path	
r_{xy}	Element of the reference signal path matrix	
$r_x(t)$	Instantaneous noise amplitude	V
R(τ)	Auto correlation function	
S	Signal power	W
S	Scattering matrix	
S/N	Signal-to-noise ratio	
$s_x(t)$	Instantaneous signal voltage	V
T	Threshold	V
V	Peak amplitude of the element excitation	V
V_{ae}	Output amplitude of a pair with amplitude errors	V

V_{eo}	Output amplitude of a pair with equal and opposite phase shifts	V
V_{pe}	Output amplitude of a pair with phase errors	V
V_{Δ}	Output amplitude of a pair with difference processing	V
V_{Σ}	Output amplitude of a pair with sum processing	V
$V_{\Sigma+\Delta}$	Output amplitude of a pair with sum + diff. processing	V
$V_{\Sigma-\Delta}$	Output amplitude of a pair with sum - diff. processing	V
ΔV	Amplitude difference between the elements of a pair	V
$\mathbf{y}$	Vector of the incident signal	V
$y_x(t)$	Cross-correlator input signal	V
$\mathbf{Z}$	Impedance matrix	
$z_x(t)$	Cross-correlator output signal	V
α	Azimuth angle, phase difference between feed lines	Radians
β	Cone width	Radians
γ	Angle between two element axes	Radians
Δ_{al}	Offset between the phase centres of the array and an individual pair along the element axes	m
Δ_{or}	Offset between the phase centres of the array and an individual pair orthogonal to the element axes	m
ε	Elevation angle	Radians
η	Noise power spectral density, if normalised to a resistance:	W/Hz V²/Hz
θ	Angle between the major axis of the bivariate ellipse and the z_{Σ} axis	Radians
θ	Angle of incidence off the element axis	Radians
θ_{eot}	Phase transition angle of a pair with equal and opposite phase shifts	Radians
$\theta_{\Delta t}$	Phase transition angle in the difference pattern	Radians
$\theta_{\Sigma t}$	Phase transition angle in the sum pattern	Radians
λ	Wavelength	m
λ_0	Quadrature correlation coefficient	
ρ	Reflection coefficient	
ρ	Correlation coefficient	
ρ_0	In-phase correlation coefficient	
σ	Standard deviation	
σ^2	Variance or second central moment	
ϕ	Phase difference between element and phase centre	Radians

ϕ_{ae}	Output phase of a pair with amplitude errors	Radians
ϕ_{cone}	Cone angle	
ϕ_e	Additional phase shift due to phase error at the pair	Radians
ϕ_{emax}	Maximum phase error at the elements	Radians
ϕ_{eot}	Phase difference for sign change of a pair with equal and opposite phase shifts	Radians
$\phi_n(t)$	Random noise phase	Radians
ϕ_{pae}	Output phase of a pair with phase and amplitude errors	Radians
ϕ_Δ	Phase difference for sign change in the diff. pattern	Radians
ϕ_Σ	Phase difference for sign change in the sum pattern	Radians
$\Delta\phi$	Applied phase shift at the elements of a pair	Radians
ω	Angular frequency	Hz
Ω	Solid angle	Steradians

LIST OF ABBREVIATIONS

ABBREVIATION	MEANING
acf	Autocorrelation function
AGC	Automatic gain control
AM	Amplitude modulation
cdf	Cumulative distribution function
CW	Continuous wave
DC	Direct current
DF	Direction finding
IF	Intermediate frequency
IIR	Infinite impulse response
LP	Low-pass
pdf	Probability density function
RF	Radio frequency
S/N	Signal-to-noise ratio

LIST OF TABLES

LIST OF FIGURES

18

1 INTRODUCTION

1.1. BACKGROUND

Phased arrays are nowadays well established work-horses in surveillance and communications and have undergone rapid development in their rather short history. Despite the fact that the scientific and engineering basis for the field of microwave transmission had been developed in the 1940's, the first microwave phased arrays did not became available until a decade later. This was due to the fact that electronic phase shifters and computers, two key components for a phased array, had previously not been available [Sta 74].

A phased array consists of a group of spatially distributed radiators which are fed from a single source and therefore coherently excited. For beamforming purposes, the phasing of the individual antenna elements is such that their contributions add constructively to a plane wavefront in a chosen direction. By controlling the phase and the amplitude of each radiator, it is possible to steer and form the beam pattern of the array. The direction and shape of the beam can therefore be electronically manipulated at microseconds speed without actually physically moving the array structure, which provides considerable beam agility. The antenna elements of the array act as interfaces that transduce guided waves into free space waves and vice versa. The theory of phased array beamforming and signal processing is well covered in the literature, and standard array structures have been established such as linear, circular, planar or conformal arrays ([Sta 74], [Mai 82], [Rud 83], [Hal 90], [Hay 92]).

Most phased arrays are periodic arrays with equally spaced elements, which makes it easier to devise design procedures. Periodic arrays with omnidirectional elements usually have spacings of d ≤ λ/2 to avoid grating

lobes. They also suffer from the effects of mutual coupling, which is caused by the close inter-element distance and changes the effective element pattern. As a result, the gain, driving-point impedance and effective aperture taper of the array can be a function of the scanning angle. All these effects have to be considered in the design. The cost of any array is approximately proportional to the number of elements, which means that the cost for a linear periodic structure is proportional to its length, and grows with the square of this quantity for a two-dimensional periodic array.

The second category of phased arrays are aperiodic or thinned arrays, which have unequally spaced elements. It is usually difficult to devise design procedures for this type of array, except for the class of random arrays [Ste 76]. A special case of random arrays, whereby the elements are uniformly randomly distributed within a sphere, has received considerable attention in the past, and Ender and Wilden actually constructed and characterised such an array which they called the crow's nest antenna ([Wil 90], [Wil 87], [End 81], [Lo 64], [Dze 78], [Lev 79]). The objective of any aperiodic array design is to reduce costs by thinning without sacrificing key array parameters like the mainlobe width or average sidelobe level. Due to the reduced number of elements, there are fewer degrees of freedom for the designer and hence less control of the radiation pattern, which makes it harder to meet the specifications. The hazard of grating lobes in aperiodic arrays due to inter-element distances of more than half a wavelength, is avoided by unequal spacing of the elements, and the larger inter-element spacing reduces the effect of mutual coupling.

The type of phased array investigated in this thesis is called a symmetric-pair array. Dependent on the application, this novel type of array belongs to the category of aperiodic or random aperiodic arrays. (However, it should be noted that certain periodic structures could also be treated as symmetric-pair configurations, but these special cases will not be considered here.) Aperiodic symmetric-pair configurations are used for direction finding, and random aperiodic symmetric-pair arrangements can be employed for beam forming and have significant attractions, if gain with isotropic coverage on reception is to be achieved.

1.2. OBJECTIVES

The concept of symmetric-pair arrays and possible applications of it have only recently been reported by Benjamin and Griffiths [Ben 89]. Such arrays consist of pairs of antenna elements that can be arranged in three dimensions around a common phase centre. Each of the antenna pairs can be regarded as an independent sub array, which substantially simplifies the analysis. Symmetric-pair arrays have output signals that are easy to process, and are therefore attractive for mobile communications and any kind of application where versatility at moderate hardware expense is required. The pair concept also allows the matching of the antenna configuration to the structural constraints of a building or vehicle.

The objective of this thesis is therefore to investigate the fundamental properties of symmetric-pair antenna arrays, both under ideal and imperfect conditions. Having established an understanding of this novel type of array, its potential for direction finding (DF) is examined in some detail. This leads to the development of a symmetric-pair DF system and the discussion of its predicted performance. Two possible applications of the symmetric-pair concept in communications are also investigated in this thesis. One of them is beam forming, and the second is gain with isotropic coverage on reception. It will be shown that the latter could not be implemented with other array structures.

1.3. STRUCTURE OF THE THESIS

The substance of the thesis is contained in the next five chapters. The properties of symmetric-pairs, different ways of processing and graphically representing their signals as well as the effects of phase and amplitude errors are covered in chapter 2. This lays the foundations necessary for the understanding of the different applications of this type of antenna array and is therefore essential reading.

Chapters 3 to 5 are devoted to direction finding with symmetric-pairs. The basic techniques used for direction finding with symmetric-pairs are outlined in chapter 3. To focus on the key principles, an idealised world

with a single infinitely strong CW carrier as received signal is assumed in this chapter. In chapter 4, realistic signals with a certain bandwidth and finite signal-to-noise ratio are introduced. After solution of the associated problems, a symmetric-pair direction finding system is developed, and different methods to process its baseband signal are discussed and compared. The performance of the proposed DF system in low signal-to-noise environments is assessed in chapter 5. This is done by taking the mathematics of random Gaussian noise at the antenna elements into account. Expressions for the probability of detection of the direction finder as a function of false-alarm rate and angle of arrival are finally derived in this chapter.

The applications of symmetric-pair arrays in communications are covered in chapter 6. Different beam forming techniques as well as the concept of gain with isotropic coverage for reception are introduced here. A summary and recommendations for future work are contained in chapter 7.

An additional note on the citation of references. All references can be found in a block before the appendices and they are grouped by chapters. In the text, they are referred to with the first three letters of the first author's surname and the year of publication. Since all references are (just about) from the 20th century, the digits representing the century have been omitted.

2 PROPERTIES OF SYMMETRIC-PAIRS

2.1. DEFINITION OF A SYMMETRIC-PAIR ARRAY

Symmetric-pair arrays consist of a number of antenna pairs of arbitrary inter-element spacing, arranged around a common phase centre in a two or even three dimensional configuration as shown in Fig. 2-1. The centre point of the two elements of each pair is called the phase centre of a pair. In principle, it is possible to use any kind of antenna elements for this type of array. The only constraint is that the elements of a pair must be identical and point in the same direction.
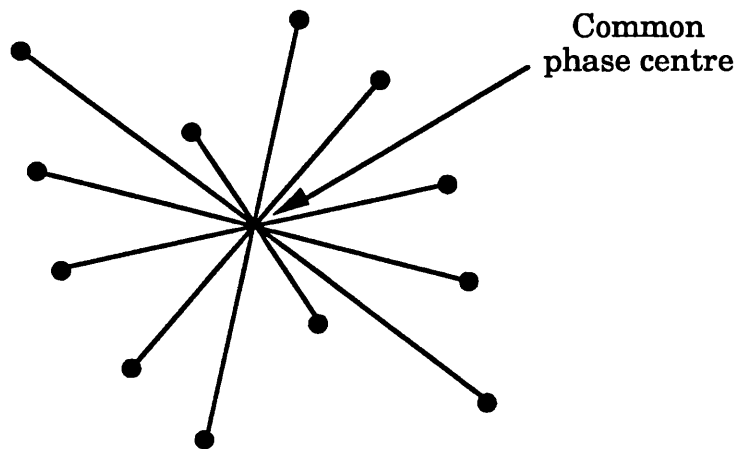
Figure 2-1: 3D array of random symmetric-pairs

Due to the common phase centre, each pair can be treated as an independent sub-array, which simplifies the calculation of the radiation pattern of the array and allows us to use the superposition principle. This applies only as long as no mutual coupling has to be taken into account. In the case of mutual coupling, possible simplifications depend on the symmetry of the array geometry.

2.2. PROCESSING OF THE SIGNALS OF A SINGLE SYMMETRIC-PAIR

To make full use of the information contained in the symmetry of the signals of each pair, processing on a pair basis is necessary. The two methods introduced in the following sections describe the fundamentals of symmetric-pair signal processing and are used in one way or another in all applications of this kind of antenna array.

2.2.1. Sum and difference processing

Because of the pair symmetry, an incident wavefront will always encounter one member of the pair, say M1, with a certain time (and hence phase) lead before that front reaches the phase centre. The other member, M2, will be excited with an equal phase lag (Fig. 2-2). With a definition of the angle of incidence θ as the angle between the element axis of the pair and the direction of propagation, counted clockwise from the reference element M1, the two elements will be excited at phases $-\phi$ and $+\phi$ relative to the phase centre:

$$\phi = \frac{\pi d}{\lambda} \cos\theta$$

...(2.1)

where λ is the wavelength and d the spacing between the elements.

The vector sum V_Σ of the signals of a pair, obtained with a lossless power combiner, is given by:

$$V_\Sigma = \frac{1}{\sqrt{2}}\left(Ve^{j(\omega t+\phi)} + Ve^{j(\omega t-\phi)}\right) = \sqrt{2}V\cos\phi \; e^{j\omega t}$$

...(2.2)

where V is the amplitude of the excitation of each element. The resultant is a real signal with positive or negative sign, dependent on whether the sum signal is in phase or anti-phase with the reference phase at the phase centre.

The vector difference V_Δ of a single pair is:

$$V_\Delta = \frac{1}{\sqrt{2}}\left(Ve^{j(\omega t+\phi)} - Ve^{j(\omega t-\phi)}\right) = \sqrt{2}V\sin\phi\ e^{j(\omega t+\pi/2)}$$

...(2.3)

This signal is orthogonal to the reference phase and can also be of either sign. All the sum outputs of an array of symmetric-pairs are either in phase or anti-phase to each other, and so are the difference signals.
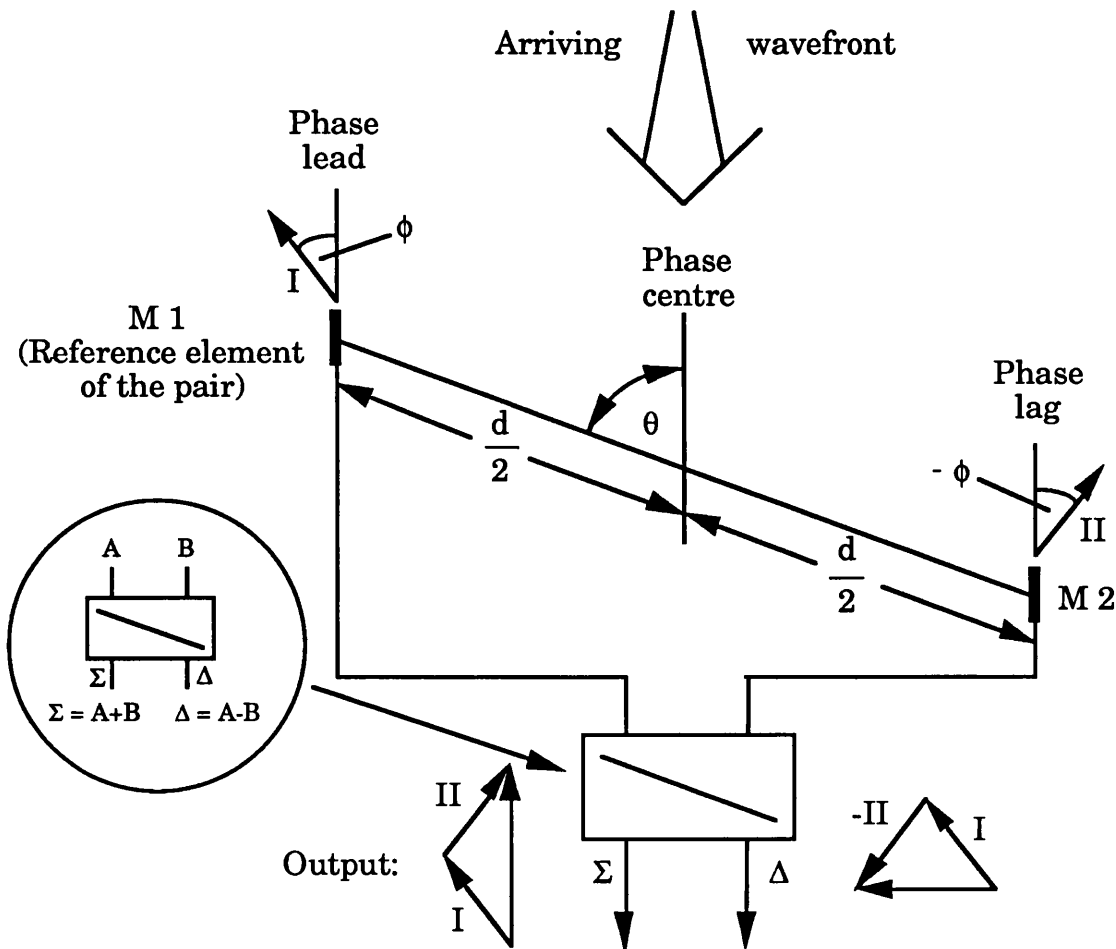
Figure 2-2: Symmetric-pair with sum and difference processing of the signals

A very simple way of testing whether signals are in phase or within $\pm\pi/2$ of the reference phase is to use a power combiner and to add the output of a symmetric-pair and the phase reference signal. (The latter can be provided by an extra element at the phase centre.) If their phase difference is less than $\pm\pi/2$ the combined signal will be larger than the reference signal alone, otherwise it will be smaller. In the case of the difference signals, a 90° phase shifter has to be inserted, to account for the orthogonality of the difference signal.

The standard device for generating the sum (Σ) and difference (Δ) vectors is a 180° hybrid. Hence, the two vectors of equal amplitude and unknown phase at the elements of any pair can be converted into two vectors of unknown amplitudes but known phases (except for the sign ambiguity) at the sum and difference outputs of the hybrid. There is no power wasted, and the combined power of the output signals is $2V^2$, as expected.

Before the sign of the sum or difference signal changes, the corresponding amplitude shrinks to zero. The angles of incidence at which the sign changes occur are called phase transition angles. The sign bit of the Σ output changes, if the phase difference ϕ_Σ between one element of the pair and the phase reference is:

$$\phi_\Sigma = \left(\frac{\lambda}{4} + n\frac{\lambda}{2}\right)\frac{2\pi}{\lambda} , \qquad m_\Sigma \geq n \geq 0$$

where n is an integer and $m_\Sigma = d/\lambda - 0.5$ rounded down to the next integer. The phase transition angles $\theta_{\Sigma t}$ for $0° \leq \theta \leq 90°$ are now defined as follows:

$$\theta_{\Sigma t} = \arccos\left(\frac{\lambda/4 + n\lambda/2}{d/2}\right) = \arccos\left(\frac{0.5 + n}{d/\lambda}\right) , \qquad m_\Sigma \geq n \geq 0 \qquad ...(2.4)$$

Since the Σ pattern of a symmetric-pair is symmetrical about boresight and the element axis, the phase transition angles for the other quadrants are easily calculated.

In the difference pattern, the sign changes for:

$$\phi_\Delta = \left(n\frac{\lambda}{2}\right)\frac{2\pi}{\lambda} , \qquad m_\Delta \geq n \geq 0$$

where n is an integer and $m_\Delta = d/\lambda$ rounded down to the next integer. For $0° \leq \theta \leq 90°$, the phase transition angles $\theta_{\Delta t}$ of the difference pattern are:

$$\theta_{\Delta t} = \arccos\left(\frac{n\lambda/2}{d/2}\right) = \arccos\left(\frac{n\lambda}{d}\right) , \qquad m_\Delta \geq n \geq 0 \qquad ...(2.5)$$

: calculated pattern   - - - - - : measured pattern
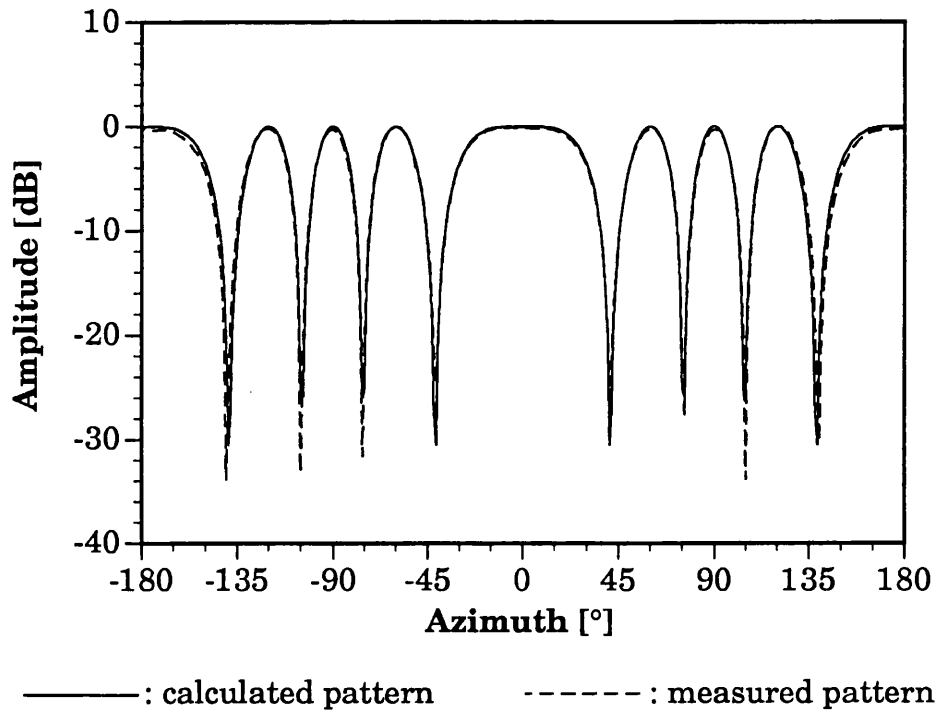
Figure 2-3: *Measured and calculated sum pattern of a symmetric-pair with 2λ spacing*

: calculated pattern   - - - - - : measured pattern
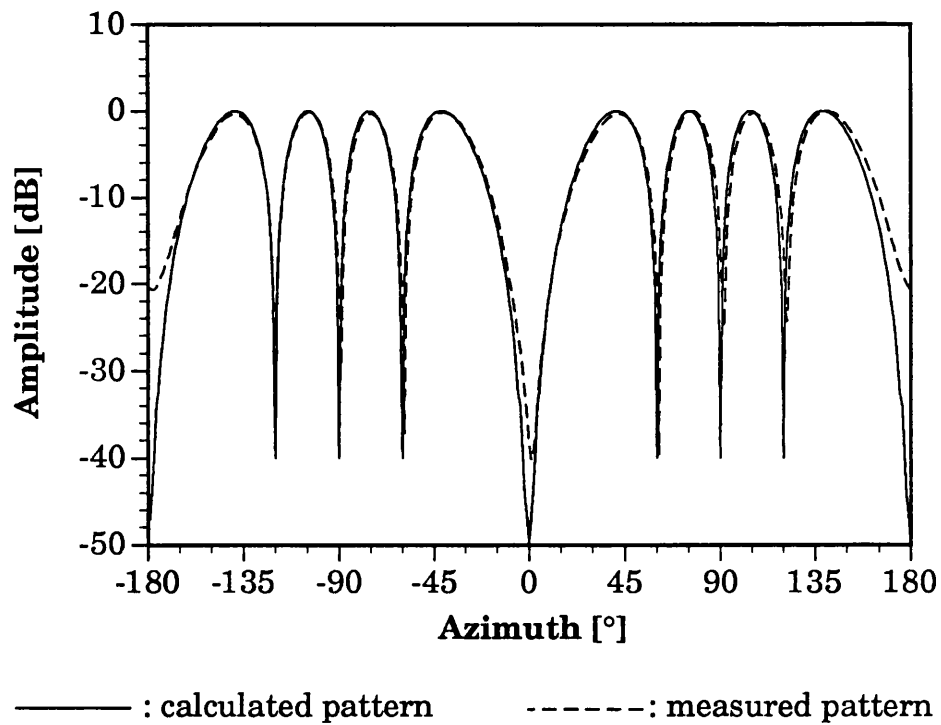
Figure 2-4: *Measured and calculated difference pattern of a symmetric-pair with 2λ spacing*

This pattern is also symmetrical about boresight and the element axes and has a phase transition at 90° and 270° irrespective of the element spacing. However, the polarity of the sign zones is now anti-symmetric about boresight.

The phase transition angles coincide with nulls in the amplitude pattern of a pair. Because of the symmetry of both patterns about the element axis, no phase transition can occur at 0° or 180°, but there may well be zeros in the antenna pattern, dependent on the element spacing (for the sum pattern: d = odd multiple of $\lambda/2$, difference pattern: d = multiple of λ).

The calculated and measured sum and difference amplitude patterns of a symmetric-pair with 2 λ spacing are shown in Fig. 2-3 and Fig. 2-4 respectively. The prototype consisted of two monopoles on a ground plane with 2 λ spacing at 10 GHz. Fig. 2-4 demonstrates that the difference pattern has, as expected, nulls at 0° and 180° for a spacing of 2 λ. The two neighbouring grating lobes at each of the endfire positions therefore have the same sign.

2.2.2. Equal and opposite phase shifts at the elements

As in any antenna array, it is also possible to process the signals of symmetric-pairs with the help of phase shifters. However, to retain the symmetry of the arrangement and to keep the location of the phase centre of each pair unchanged, the phase shifts applied to each element of a pair must be equal and opposite (see Fig. 2-5) ([Fra 67],[Gly 68]). The output V_{eo} of a pair with such phase shifting networks and subsequent sum processing is equal to:

$$V_{eo} = \frac{1}{\sqrt{2}}\left(Ve^{j(\omega t+\phi+\Delta\phi)} + Ve^{j(\omega t-\phi-\Delta\phi)}\right) = \sqrt{2}V\cos(\phi+\Delta\phi)\,e^{j\omega t}$$

...(2.6)

where $\pm\Delta\phi$ is the applied phase shift at each element.

Eqn. 2.6 shows that the output signal of a symmetric-pair after the use of equal and opposite phase shifts is still bi-phasal, but phase

transition will now occur at different angles of incidence. The condition for phase transition is then:

$$\phi_{eot} = \left(\frac{\lambda}{4} + n\frac{\lambda}{2} - \Delta\phi\frac{\lambda}{2\pi}\right)\frac{2\pi}{\lambda} \ , \qquad\qquad m_1 \geq n \geq m_2$$

where:

$$m_1 = \frac{d}{\lambda} - \left(0.5 - \frac{\Delta\phi}{\pi}\right) \ , \quad \text{rounded down to the next integer}$$

$$m_2 = \frac{-d}{\lambda} - \left(0.5 - \frac{\Delta\phi}{\pi}\right) \ , \text{rounded up to the next integer}$$

ϕ is the phase difference between one of the elements and the phase centre and $\Delta\phi$ is the phase shift applied at the reference element M1 of a pair (see Fig. 2-2). Due to the phase shifts, the symmetry of the phase transition angle pattern about boresight is lost, whereas the symmetry about the element axis still remains. The above values for n define the phase transition angles θ_{eot} for $0° \leq \theta \leq 180°$:

$$\theta_{eot} = \arccos\left(\frac{\lambda/4 + n\lambda/2 - \Delta\phi \ \lambda/(2\pi)}{d/2}\right) = \arccos\left(\frac{0.5 + n - \Delta\phi/\pi}{d/\lambda}\right) \qquad ...(2.7)$$

where n is again an integer number between m_1 and m_2.
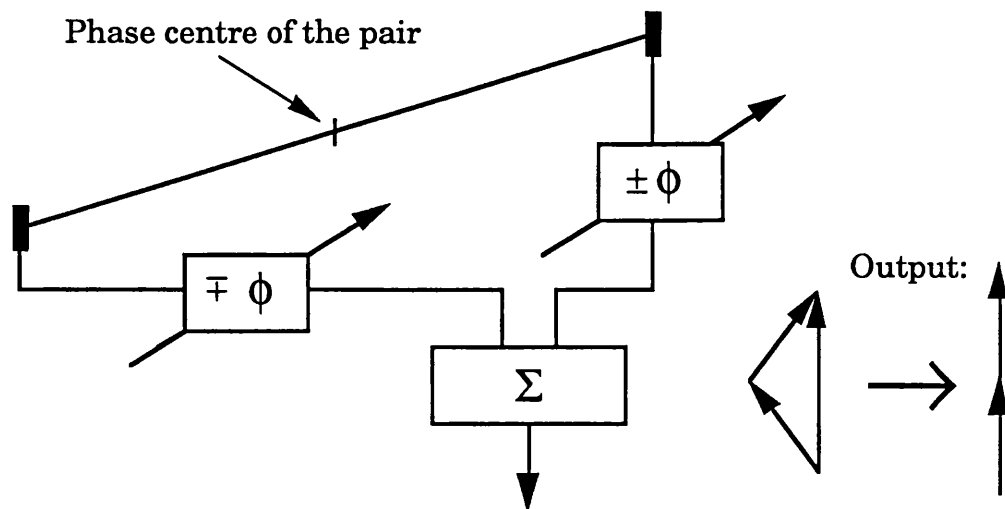
Figure 2-5: Symmetric-pair with equal and opposite phase shifts

Equal and opposite phase shifts with -90° at the reference element of the symmetric-pair yield the same pattern of phase transition angles as difference processing. The only difference is that the bi-phasal output signals are now still in phase or anti-phase to the reference signal and not orthogonal to it, as   is the case after difference processing.

If the applied phase shift at element M1 is positive, the pattern of the phase transition angles for $0° \leq \theta \leq 180°$ rotates clockwise, and the pattern for $180° \leq \theta \leq 360°$ rotates counter-clockwise. For negative phase shift at element M1 the directions of rotation are reversed. These relations can be easily visualised with the help of the graphical representations introduced in the next section.

It is convenient at this point to give a general formula for the phase transition angles (i.e. nulls) and the maxima of a symmetric-pair pattern that covers all types of processing. The conditions for a maximum can be derived by considerations similar to the ones for the phase transition angles. The respective values of n yield nulls and maxima for $0° \leq \theta \leq 180°$:

$$\theta = \arccos\left(\frac{a+n}{d/\lambda}\right) \qquad m_1 \geq n \geq m_2 \qquad \text{...(2.8)}$$

where:

$$m_1 = \frac{d}{\lambda} - a , \quad \text{rounded down to the next integer}$$

$$m_2 = \frac{-d}{\lambda} - a , \quad \text{rounded up to the next integer}$$

Processing technique	Values for a	
	Null	Maximum
Σ processing	0.5	0
Δ processing	0	-0.5
Σ processing with equal and opposite phase shifts	$0.5 - \dfrac{\Delta\phi}{\pi}$	$-\dfrac{\Delta\phi}{\pi}$

Table 2-1:   Values of constant a for the calculation of maxima and nulls

The value of the constant a is given in Table 2-1 and depends on the processing technique and whether the location of nulls or maxima is

calculated. It should be noted that any integer could be added in the definition of a, since the values for m_1 and m_2 would be automatically adjusted, so that eqn. (2.8) would still be correct.

2.3. GRAPHICAL REPRESENTATION OF THE SIGNALS OF SYMMETRIC-PAIRS

In the last section it has been shown that the output signals of symmetric-pairs are essentially always just in phase or anti-phase with each other. Since this is very important for symmetric-pair applications, two ways to represent this property graphically are introduced. For the sake of simplicity, the patterns of the antenna elements are assumed to be isotropic.

2.3.1. Single symmetric-pair but variable angle of arrival

This first representation is in its 2 dimensional version a cross section of the sign bit pattern of one symmetric-pair. Fig. 2-6a shows the cross section of the sum pattern of a pair with 4λ spacing, and Fig. 2-6b, represents the difference pattern of the same pair. Both pictures show clearly the symmetries of the two patterns about boresight and the element axis. It should also be noted that the sectors around boresight are rather narrow, whereas the sectors in the endfire areas of the pair are very broad. Each of the sectors contains a grating lobe of the amplitude pattern with a maximum in the middle of the sector and nulls at the borders. However, the sectors around the endfire area may just contain a fraction of a lobe dependent on the spacing of the elements. The endfire lobes do not exist at all for element spacings that produce a null in the endfire direction. It can be seen that for $0° \leq \theta \leq 180°$ the sum pattern consists of 2m + 1 lobes of alternate sign (Fig. 2-6a), and the difference pattern of 2m such lobes (Fig. 2-6b). Whereby m is defined as follows:

Sum pattern: m = d/λ - 0.5,   rounded up to the next integer
Difference pattern: m = d/λ,   rounded up to the next integer
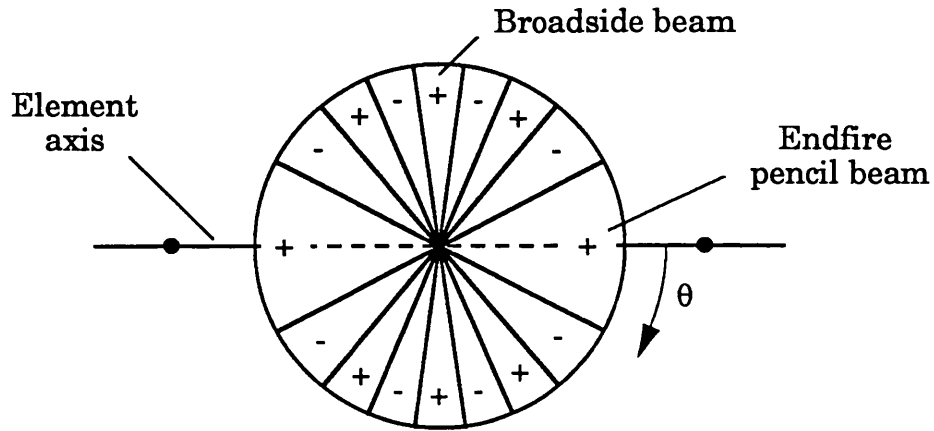
Figure 2-6a: Cross section of the sum sign bit pattern of a symmetric-pair with 4 λ spacing
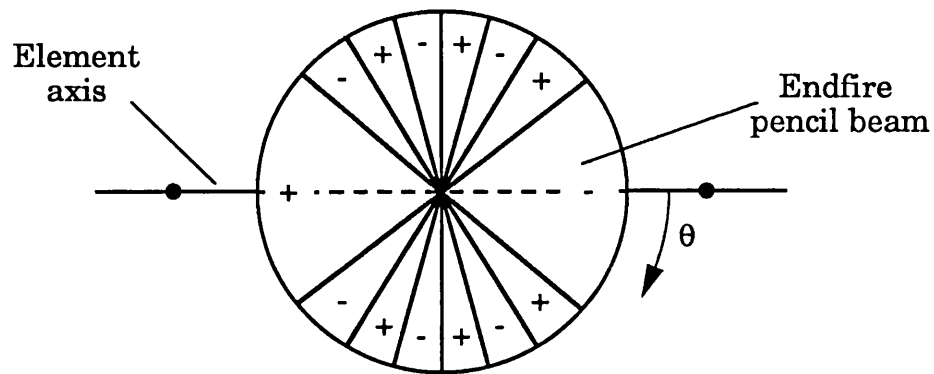
Figure 2-6b: Cross section of the difference sign bit pattern of a symmetric-pair with 4 λ spacing
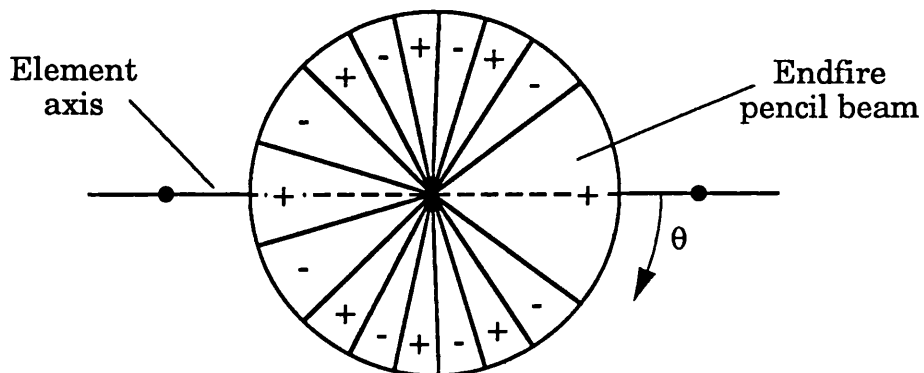
Figure 2-7: Cross section of the sum sign bit pattern of a symmetric-pair with 60° equal and opposite phase shift and 4 λ spacing

The impact of equal and opposite phase shifts on the sign bit pattern can be seen in Fig. 2-7. As predicted, there is no longer symmetry about boresight.

The actual 3 dimensional radiation pattern would be generated by revolving these fan patterns about the axis of the pair (Fig. 2-8). Thus, the sum pattern usually comprises one broadside disc, 2(m-1) hollow cones, and two solid-cone end-fire beams, all with common apex. The difference pattern has no broadside disc, and because of the nature of the difference signal, the two endfire cones are always of different sign.

Element axis

☐ = in phase (+)
▒ = anti-phase (-)
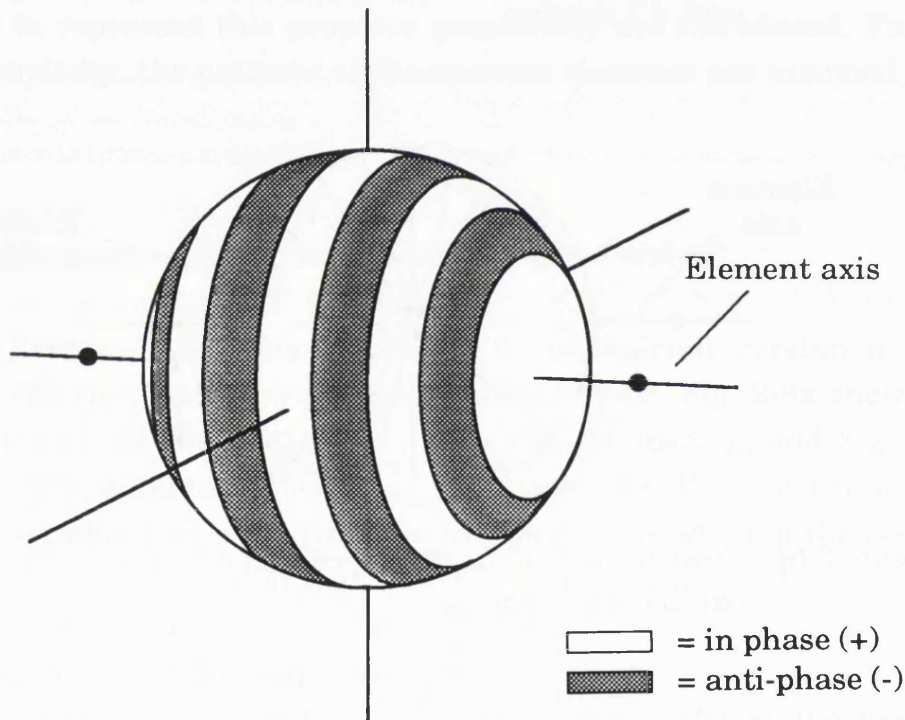
Figure 2-8:   3D Σ sign bit pattern of a symmetric-pair with 4 λ spacing

The broadside disc and the hollow cones cover the same solid angle Ω:

$$\Omega = 2\pi \int_{\theta_1}^{\theta_2} \sin\theta \, d\theta = 2\pi(\cos\theta_1 - \cos\theta_2) = \frac{2\pi\lambda}{d} \qquad \ldots(2.9)$$

where θ_1 and θ_2 are two neighbouring phase transition angles calculated with eqn. (2.4), eqn. (2.5) or eqn. (2.7), and the sphere is assumed to be a unity sphere. In the 3 dimensional representation of the signs of a

symmetric-pair, the phase transition angles cut the surface of a sphere into slices of equal thickness, which are centred on the element axis as shown in Fig. 2-8. Equal and opposite phase shifts just move these slices along this axis but do not change their thickness.

2.3.2. Fixed angle of arrival but arbitrary symmetric-pair configuration

The idea behind this second representation is to consider the wavefield in the direction of propagation as composed of a sequence of positive (+) and negative (-) zones of width $\lambda/2$. This has the advantage that it is possible to detect for a certain angle of arrival the sign bits of all pairs of a symmetric-pair array simultaneously. If the array is 3 dimensional, the positions of the elements have to be projected in the considered wavefront plane.

For the sum pattern of a symmetric-pair array, its phase centre lies in the middle of a positive zone (Fig. 2-9a), whereas for the difference pattern, the phase centre of the array is displaced by $\lambda/4$ (Fig. 2-9b). In both figures, the solid element of each pair is used to define the angle of incidence and the sign bit of the output signal. (The solid element corresponds to the reference element M1 in Fig. 2-2).

It is also possible to evaluate the sign bits of a symmetric-pair array, when equal and opposite phase shifts are used. The phase shifts can be represented as vectors, which start at the elements of each pair and are always parallel to the direction of propagation, since their influence is irrespective of the angle of incidence. The two vectors are antiparallel to each other and their length corresponds to the applied phase shift. If the phase shift is negative, the vector at the solid reference element is in the direction of propagation, for positive phase shifts, this vector is antiparallel. The sign of the output signal of each pair is determined by the sign of the zone in which the tip of the reference element vector lies. To illustrate the above matter, a two pair array with a phase shift of -120° and +90° at the pairs is shown in Fig. 2-10.

●: Reference element

(b) Difference sign zones   (a) Sum sign zones
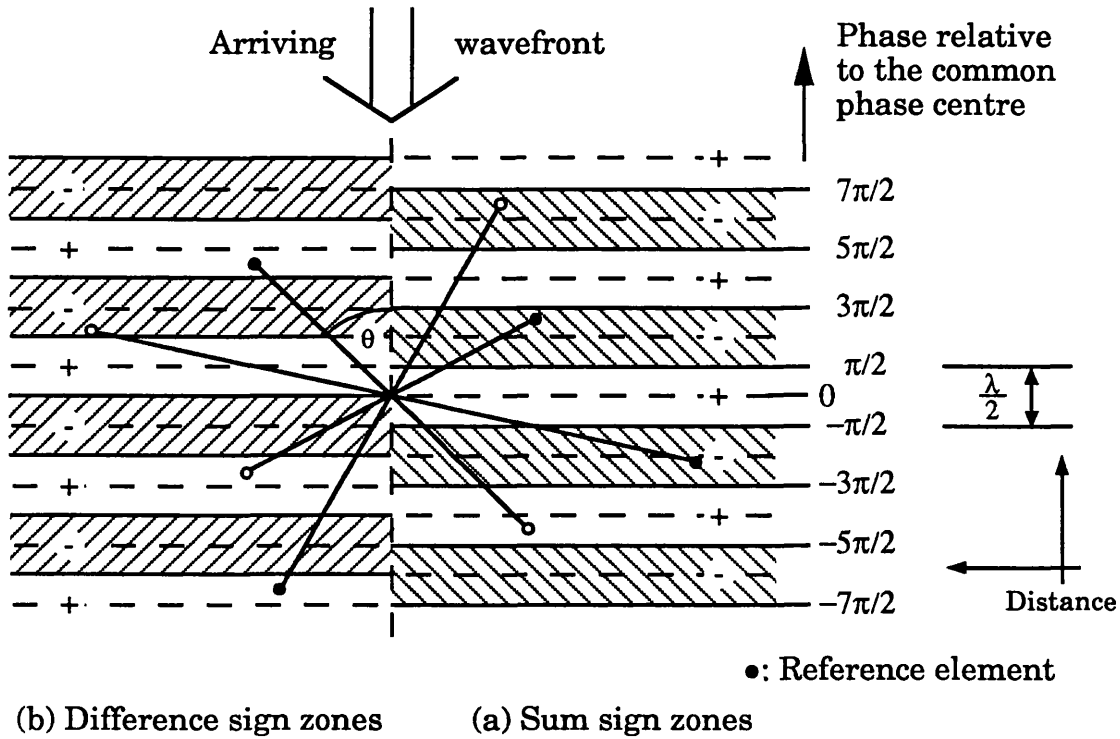
Figure 2-9: *Sign zones for the sum and difference patterns of*
a symmetric-pair array

●: Reference element
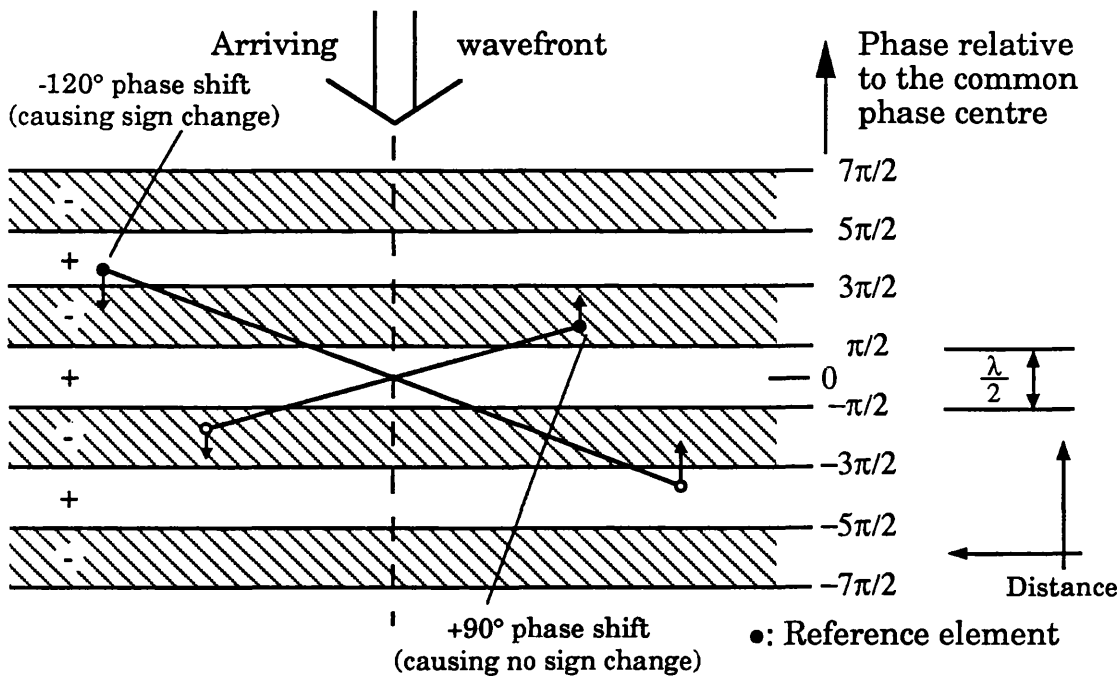
Figure 2-10: *Sign zones for a two pair array with equal and opposite*
phase shifts at the elements and sum processing

2.4. BANDWIDTH CONSIDERATIONS

The largest possible phase change $\Delta\phi$ at the element of a pair due to a frequency change of Δf can be expressed by slightly modifying eqn. (2.1):

$$\phi = \frac{\pi d}{\lambda}\cos\theta = \frac{\pi df}{c}\cos\theta$$

therefore:

$$\Delta\phi = \frac{\partial\phi}{\partial f}\Delta f = \frac{\pi d}{c}\cos\theta\,\Delta f \overset{\substack{\text{for max }\Delta\phi \\ |\cos\theta|=1}}{=} \frac{\pi d}{c}\Delta f$$

Using this relationship, the corresponding fractional bandwidth is given by:

$$b/w = \pm\frac{\Delta f}{f_c} = \pm\frac{\lambda_c}{d\,\pi}\Delta\phi \qquad \text{...(2.10)}$$

The 3 dB fractional bandwidth of a pair can be calculated for $\Delta\phi = \pi/4$ (the sum vector decreases from length $\sqrt{2}$ for fully co-phasal summation to 1). With eqn. (2.10), the fractional bandwidth is $\pm\lambda_c/4d$. For d equal to the array diameter D, this can be considered as the lower limit for the 3 dB bandwidth of a symmetric-pair array. For a phase change at the elements of $\Delta\phi = \pi/2$, the phase of a pair is reversed (relative to the phase centre). This will take place for a fractional frequency change of $\pm\lambda_c/2d$. The same frequency change incidentally defines the -3.9 dB bandwidth of a uniform linear array, where d would be the length of the array. It should be noted that the bandwidth calculations are only correct if fully co-phasal summation takes place at the centre frequency f_c, otherwise the bandwidth will be smaller, since the symmetry about the centre frequency is lost.

Eqn. (2.10) shows as well that, as expected, the bandwidth decreases with increasing spacing of the elements. However, the symmetric-pair principle as such works over virtually unlimited frequency range, provided, the received and transmitted signals are reasonably narrow band. In most applications, the limiting factor will be the bandwidth of the antenna elements.

2.5. EFFECTS OF PHASE AND AMPLITUDE ERRORS

So far in the analysis of symmetric-pairs, ideal antenna elements and processing hardware as well as no obscuration or mutual coupling from other pairs has been assumed. In practice, however, the amplitude and phase response of a pair could well be affected by such effects. The purpose of the following sections is to investigate as far as possible the effect of systematic phase and amplitude errors, i.e. errors that are not affected by the quality of the phase front of the received signal. The influence of these errors on both the accuracy of the phase transition angles of a pair and the shape of the output amplitude will be examined. There will only be a treatment of phase and amplitude errors for the sum signal of a pair, since the difference signal as well as signals of pairs with equal and opposite phase shifts show the same behaviour.

2.5.1. Phase errors only

Three different types of phase errors will be considered. The first one is caused by a change in element spacing due to hardware imperfections. This clearly changes the phase transition angles and the amplitude pattern according to eqn. (2.4) and eqn. (2.2) respectively.

The second type of phase error occurs due to different line lengths between the elements of a symmetric-pair and its summing network. This is a non symmetrical error and also changes the phase transition angles as well as the output phase, but can be easily corrected with phase shifters. If the phase difference between the two feed lines is α, the output signal of the summing network can be written as:

$$V_{pe} = \frac{1}{\sqrt{2}}\left(Ve^{j(\omega t + \phi + \alpha)} + Ve^{j(\omega t - \phi)}\right) = \sqrt{2}V\cos\left(\phi + \frac{\alpha}{2}\right)e^{j\left(\omega t + \frac{\alpha}{2}\right)}$$

...(2.11)

As can be seen, the output phase is offset by $\alpha/2$ due to the error, and the phase transition angles are shifted by a constant value. Element obscuration and mutual coupling yield the same problems at the final summing network. However, the effects of these types of error are hard to eliminate, since they are dependent on the angle of incidence.

In the third case, the element spacing is correct, but the phase centre of the array no longer coincides with the phase centre of the pair due to hardware inaccuracies. Ignoring the angular frequency term $e^{j\omega t}$ and taking the array phase centre as reference, the sum signal of the pair can then be written as:

$$V_{pe} = \frac{1}{\sqrt{2}}\left(Ve^{j(\phi+\phi_e)} + Ve^{j(-\phi+\phi_e)}\right) = \sqrt{2}V\cos\phi\,e^{j\phi_e}$$

...(2.12)

where ϕ_e is the phase error at the elements of a pair, which varies with the angle of incidence and is caused by the offset between the two phase centres. This offset consists of one component along and another orthogonal to the element axis (Fig. 2-11). The additional phase shift due to the error is therefore defined as follows:

$$\phi_e = \frac{2\pi}{\lambda}\left(\Delta_{al}\cos\theta + \Delta_{or}\sin\theta\right)$$

...(2.13)

where Δ_{al} is the offset along and Δ_{or} the offset orthogonal to the element axis. Δ_{al} is positive if the array phase centre is displaced away from the reference element of the pair, and Δ_{or} is positive for displacements of the phase centres towards 270° (negative for offsets towards 90°). The definition of the sign of the offsets is also shown in Fig. 2-11.

Examination of eqn. (2.12) shows that the amplitude pattern of a symmetric-pair remains unaffected by the offset between the phase centres, and the location of the phase transition angles is not affected either. However, dependent on the angle of incidence, a phase offset is added to the binary phase of the output signal. As long as the offset between the array and pair phase centres is less than $\lambda/4$, it is still possible to detect, whether the signal belongs to the in phase or anti-phase hemisphere (Figs. 2-12a-c). Larger offsets, as in Fig. 2-12d, entail a phase shift of the output signal which is larger than ±90°, which makes this distinction impossible.

In practice, it is often desirable to locate the offset between the two phase centres. The maximum phase error occurs when the geometrical phase centre of the pair and the phase centre of the array lie on a straight
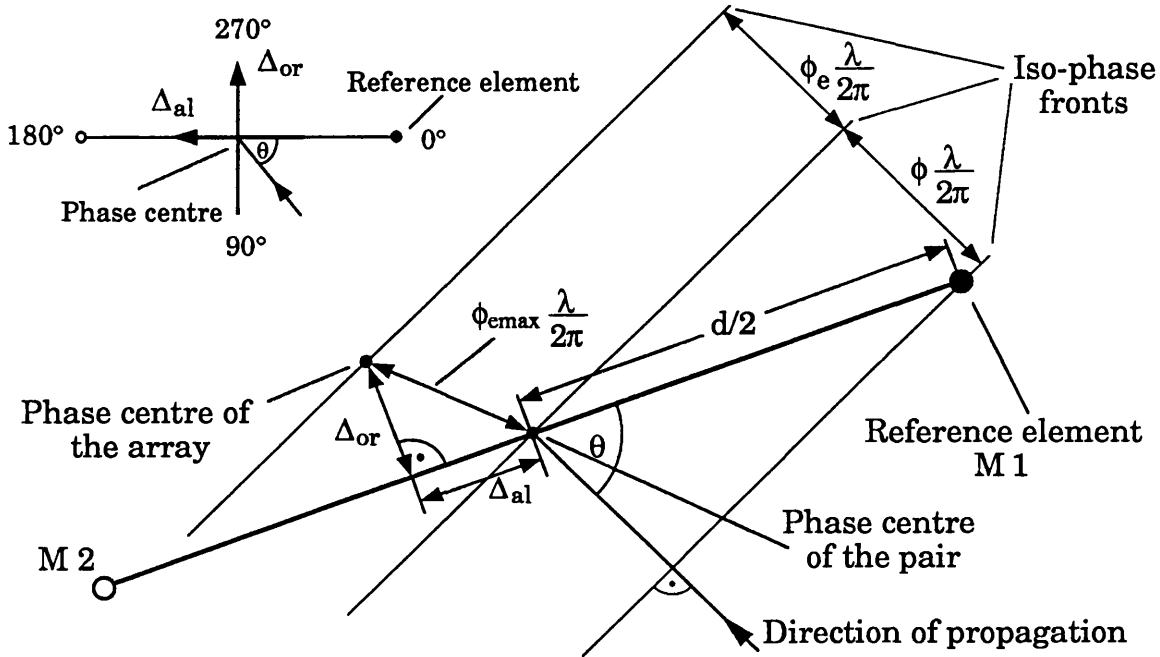
Figure 2-11: Definition of the offset of the phase centre

(a) $\Delta_{al} = 0\ \lambda;\ \Delta_{or} = 0\ \lambda$

(b) $\Delta_{al} = 0.2\ \lambda;\ \Delta_{or} = 0\ \lambda$

(c) $\Delta_{al} = 0\ \lambda;\ \Delta_{or} = 0.2\ \lambda$

(d) $\Delta_{al} = -0.2\ \lambda;\ \Delta_{or} = -0.2\ \lambda$
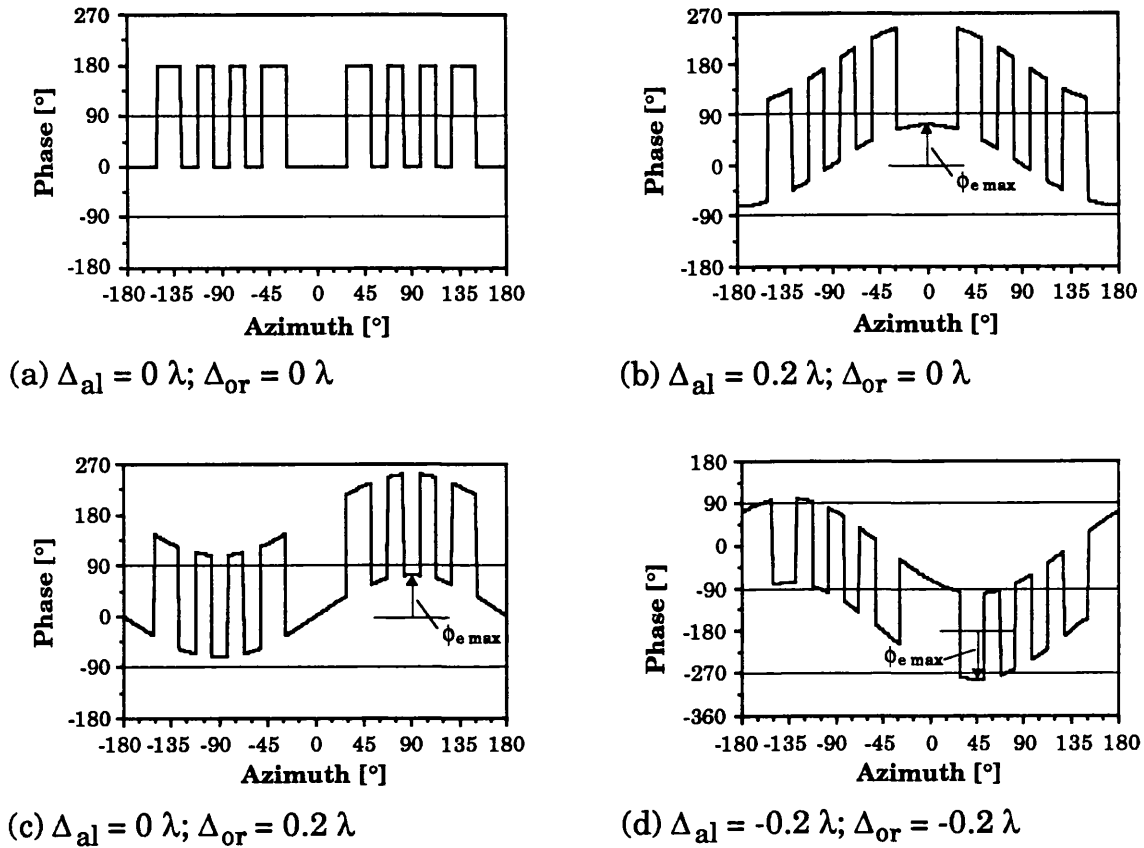
Figure 2-12: Phase of a symmetric-pair with 4 λ spacing and different phase errors

line with the direction of propagation of the signal. Knowing at which angle θ the maximum phase error occurs and the value and sign of the maximum phase error ϕ_{emax}, it is possible to calculate Δ_{al} and Δ_{or}. Using the fact that

$$\left(\phi_{emax}\frac{\lambda}{2\pi}\right)^2 = \Delta_{al}^{\ 2} + \Delta_{or}^{\ 2}\ ,$$

and rearranging it with the help of eqn (2.13), the offsets along and orthogonal to the element axis are given by:

$$\Delta_{al} = \phi_{emax}\frac{\lambda}{2\pi}\cos\theta$$

...(2.14a)

$$\Delta_{or} = \phi_{emax}\frac{\lambda}{2\pi}\sin\theta$$

...(2.14b)

Simple correction of phase errors is only possible if they are independent of the angle of incidence, which is only the case for certain phase errors of the second kind.

2.5.2. Amplitude errors only

Amplitude errors in a symmetric-pair configuration can be caused by obscuration effects or mutual coupling between the elements as well as by hardware imperfections, e.g. channel imbalances.

Assuming an amplitude difference of ΔV between the two elements, the sum signal of the pair is given by:

$$V_{ae} = \frac{1}{\sqrt{2}}\left((V+\Delta V)e^{j\phi} + Ve^{-j\phi}\right) = \sqrt{2}V\cos\phi + \frac{\Delta V}{\sqrt{2}}e^{j\phi}$$

...(2.15)

The amplitude of this signal is:

$$|V_{ae}| = \frac{1}{\sqrt{2}}\sqrt{(2V+\Delta V)^2\cos^2\phi + \Delta V^2\sin^2\phi}$$

...(2.16)

As a result, the nulls of the amplitude pattern are filled in and the maximum amplitude decreases with growing ΔV (Fig. 2-13). The phase of the pair is now:

$$\phi_{ae} = \arctan\left(\frac{\Delta V \sin\phi}{(2V + \Delta V)\cos\phi}\right) = \arctan\left(\frac{\Delta V}{2V + \Delta V}\tan\phi\right) \qquad ...(2.17)$$

It should be noted that the ambiguity of the arctangent has to be taken into account, in order to get the correct phase value. The phase pattern for different amplitude errors is shown in Fig. 2-14. The quasi-digital pattern of the phase for no amplitude error becomes more and more blurred at the phase transitions with increasing ΔV. However, owing to the nature of the tangent function in eqn. (2.17), all curves intersect when the argument ϕ is a multiple of $\pi/2$. As a result, the phase difference between the output signal of a pair and the reference phase is always exactly $\pi/2$ at the phase transitions angles. By deciding whether or not the phase of a pair lies within $\pm\pi/2$ of the reference phase, it is then still possible to determine the sign of the signal. Because the slope of the phase function is very steep at the transition angles, it is not necessary to measure precisely the phase difference in order to locate the phase transition angles with acceptable accuracy.

2.5.3. Phase and amplitude errors

The previous two sections showed that for correct pair spacing and no unsymmetrical phase errors, the location of the phase transition angles is not affected by reasonably small phase or amplitude errors, and it is still possible to determine the sign of the output signal. For simultaneous phase and amplitude errors, the effects on the output signal of a pair are simply superimposed. Hence the output amplitude of a pair with both errors is equivalent to the one of a symmetric-pair with only amplitude errors (eqn. (2.16)), because phase errors do not affect the amplitude, as shown in section 2.5.1. The phase of the signal is now:

$$\phi_{pae} = \phi_e + \arctan\left(\frac{\Delta V}{2V + \Delta V}\tan\phi\right) \qquad ...(2.18)$$

—— : no amplitude error,   — · — : ΔV = -0.2 (-2dB),   — · · — : ΔV = -0.7 (-10dB)
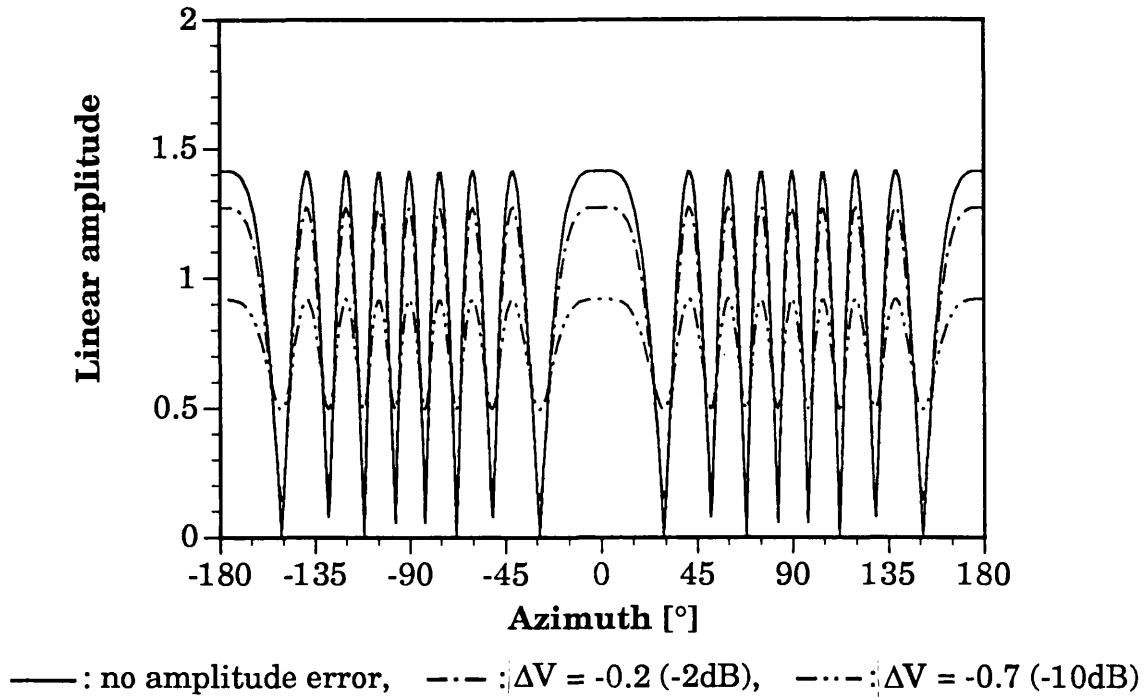
Figure 2-13: *Amplitude pattern of a symmetric-pair with 4λ spacing and different amplitude errors*

—— : no amplitude error,   — · — : ΔV = -0.2 (-2dB),   — · · — : ΔV = -0.7 (-10dB)
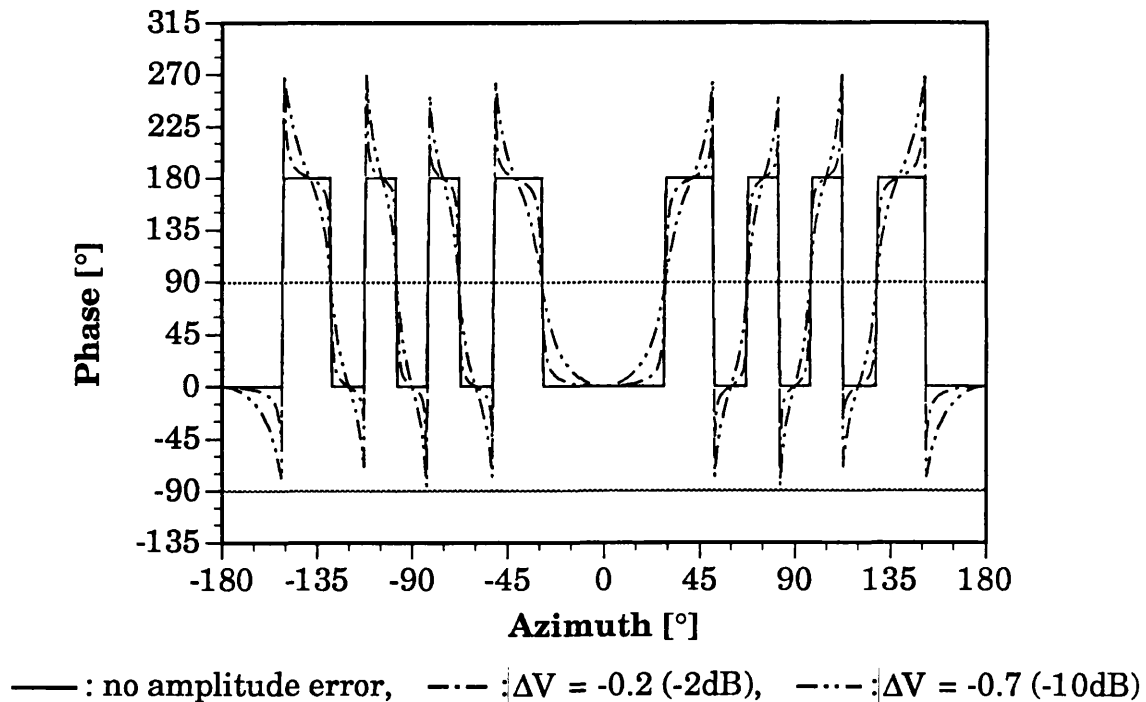
Figure 2-14: *Phase pattern of a symmetric-pair with 4λ spacing and different amplitude errors*
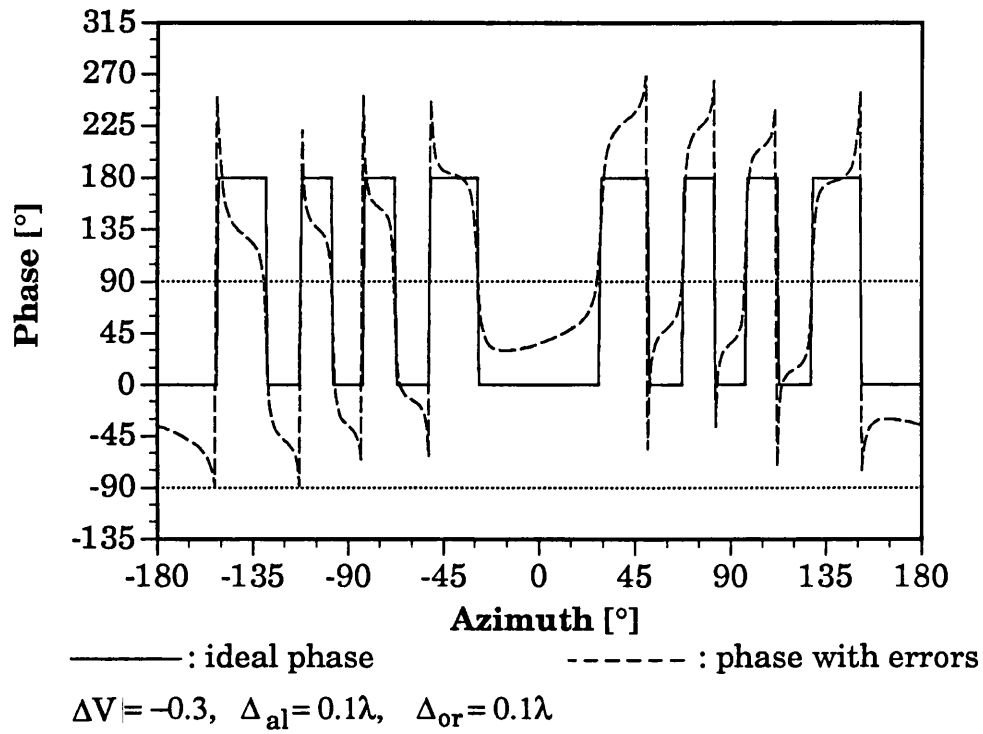
Figure 2-15:   *Phase pattern of a symmetric-pair with 4λ spacing and amplitude as well as phase errors*
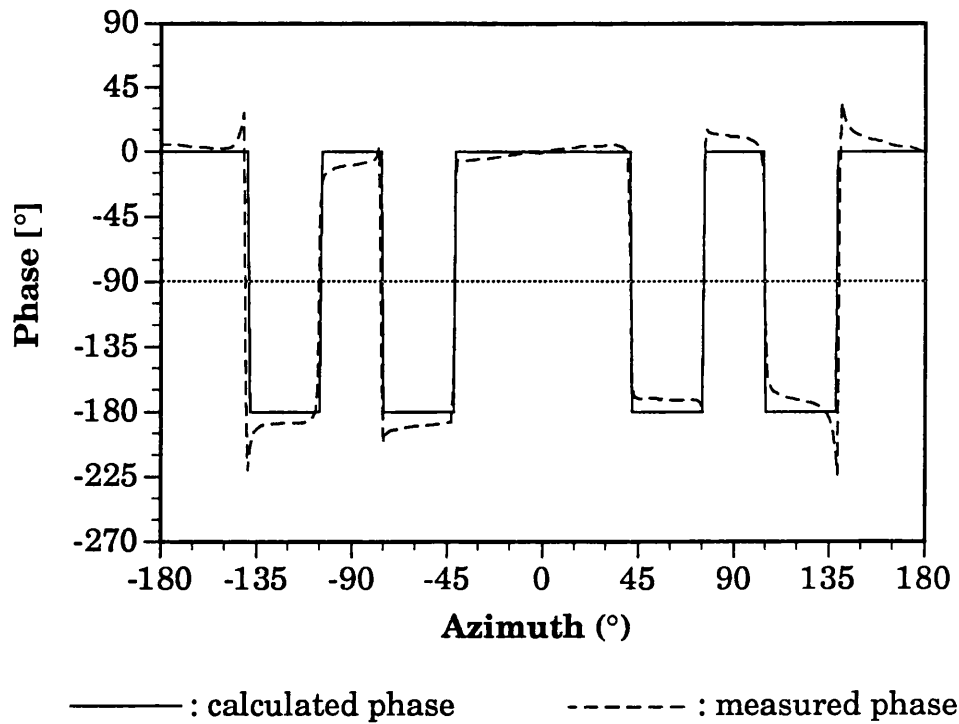
Figure 2-16:   *Measured and calculated phase pattern of a symmetric-pair with 2λ spacing*

Appendix 1 presents another way of deriving the phase of a symmetric-pair with amplitude and phase errors. As an example, the phase pattern of a symmetric-pair with $\Delta_{al} = \Delta_{or} = 0.1\lambda$ and $\Delta V = -0.3$ is shown in Fig. 2-15. Because of the superposition of both errors, the phase transition angles are not clearly defined any more. However, as long as the amplitude errors are small, the detected phase transition angles will virtually be identical with the calculated ideal angles. Fig. 2-16 shows the phase pattern of a symmetric-pair prototype with 2λ spacing. The measured pattern shows signs of small amplitude and phase errors. The phase error occurred because the phase centre of the pair was not identical with the middle of the turntable. Using eqns. (2.14a/b), the offset of the phase centre along the element axis was 0.25 mm, and the offset orthogonal to it -1.2 mm. The amplitude error was probably caused by the amplitude imbalance of the two channels of the hybrid used for this measurement (± 0.5 dB according to the data sheet).

3 DIRECTION FINDING WITH SYMMETRIC-PAIRS - THE PRINCIPLE

3.1. CURRENT DIRECTION FINDING TECHNIQUES

It is not the purpose of this section to review in detail all current direction finding (DF) principles, since this has been thoroughly covered in the relevant literature ([Get 89], [Gra 89], [Lip 87]). However, in order to appreciate the simplicity and advantages of a symmetric-pair direction finder, some basic ideas behind known DF systems will be summarised.

The purpose of any direction finding system is to establish the angle of arrival of a certain signal. This angle of arrival can be unambiguously described by the azimuth angle α, measured clockwise from north, and the elevation angle ε. The first DF systems were built shortly after the turn of the century and the history of direction finding is linked with the discovery of certain phenomena and the invention of some fundamental DF principles. According to Grabau and Pfaff [Gra 89], the landmarks in this development are the following:

(1) The discovery of directional properties of antennas.

(2) The subdivision of a DF system into an antenna and a signal processing part, whereby the latter does the actual calculation of the bearing (Bellini-Tosi principle 1908 [Bel 08]).

(3) The invention of the Adcock system which is a polarisation error free direction finder ([Adc 19], [Smi 26]).

(4) The development of the Watson-Watt two channel DF system [Wat 26].

(5) The invention of the Wullenwever direction finder as forerunner of today's circular and phased array systems [Rin 56]. (Not Wullenweber, as it is often erroneously called in the English literature.)

(6) The application of digital signal processing in DF.

The last step is still under rapid development, and it should be added that the use of superresolution techniques in recent years has opened new fields which offer scope for many new developments ([Hil 90], [Nic 87], [Joh 86]). Random arrays can theoretically also be used for direction finding ([Don 76], [Kas 78]), however, there is to the author's knowledge at present no commercial product available based on this technique.

It is also interesting to notice that two important developments for direction finding were reported in the same journal within months of each other, and the combination of both solved some of the most notorious problems at the time. Watson Watt developed the first instantaneous direction reading DF system for the investigation of thunderstorms and lightning, i.e. signals of very short duration whose bearing could not be taken with the up to then commonly used radiogoniometers [Wat 26]. However, he used two crossed loops, which left his system susceptible to polarisation errors. Smith-Rose on the other hand, who gave his paper before the Wireless Section of the IEE two months later in May 1926, managed to explain the cause of the polarisation errors (night errors) and solved the associated problems by using two crossed antenna pairs, as previously employed by Adcock [Smi 26]. However, he still worked with a goniometer to establish the angle of arrival and had therefore no instantaneous display of the angle of arrival.

All current direction finders use the amplitude or/and phase information provided by the antenna system to calculate the angle of arrival. Furthermore, all systems obtain the data required for the calculation of the bearing by highly accurate measurements of these parameters, which often demands very sophisticated hardware. In contrast, the antenna system of a symmetric-pair direction finder combined with some simple additional hardware already delivers a quasi binary signal, which is simple to extract and process. Similar to interferometers, symmetric-pair DF systems need rather large apertures

to achieve high accuracy and are capable of calculating the angle of arrival of a signal both in azimuth and elevation.

3.2. UNAMBIGUOUS 3D DIRECTION FINDING WITH SYMMETRIC-PAIRS

Since the polarity of the output signal of a symmetric-pair is dependent on the angle of incidence, is it obvious to use this information for the determination of the signal's angle of arrival. The sign bit pattern of a given symmetric-pair geometry will change for signals from different angles of incidence. Assuming an unambiguous arrangement, each angle of incidence would then produce a unique sign bit pattern. Using just the sum and difference information provided from an array with p pairs, one is able to distinguish between 2^{2p} different combinations. However, dependent on the array geometry, not all of these patterns will correspond to a plane wavefront.

For direction finding, it is essential that the estimated direction of arrival is unambiguous. In the case of symmetric-pairs, this means that the cones of one pair with the same sign must be distinguishable (see Chapter 2, Fig. 2-8). A convenient way of resolving these ambiguities is to use a set of collinear spaced symmetric-pairs (Fig. 3-1). The set of ambiguous cones can now be reduced to one unique cone each. Since the sum and difference signals of a pair provide 2 bits of information, they select 1/4 of the total solid angle. The correct cone can be defined by supplementing the relevant pair with p_c-1 coarse-resolution pairs where:

$$p_c = 1 + \log_4\left(\frac{d}{\lambda}\right)$$

...(3.1)

p_c is always rounded up to the next integer. The spacing of each consecutive pair would be 1/4 of the preceding one, and that of the central pair $\le \lambda$ to resolve the left/right ambiguity of the sum pattern. In fact, this last remaining ambiguity is solved by the anti-symmetry of the difference pattern. Fig. 3-2 shows the sign bit patterns of 3 collinear symmetric-pairs
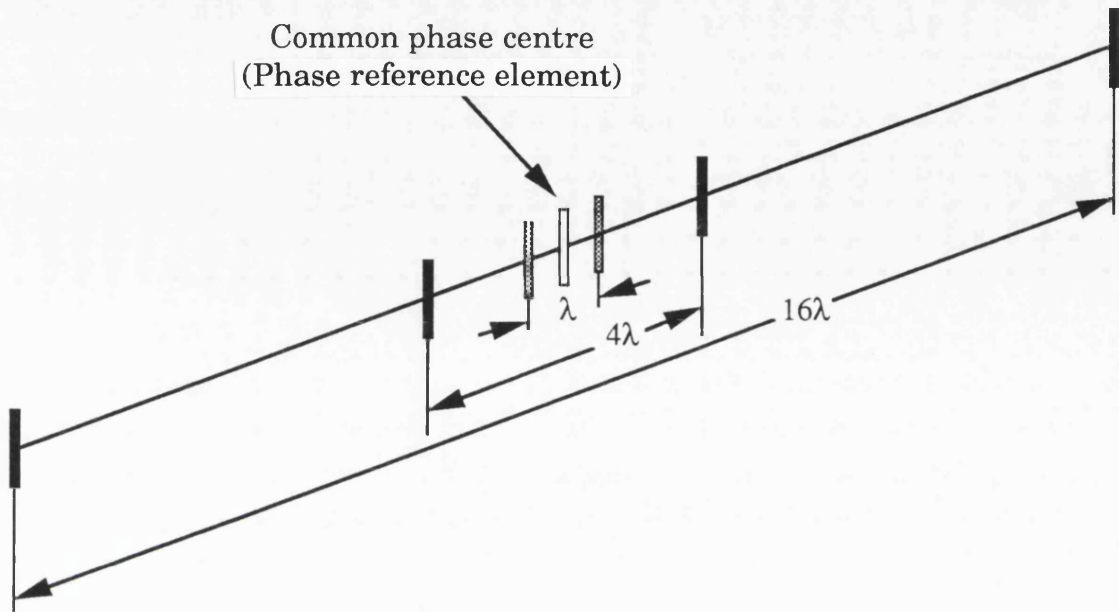
Figure 3-1: Collinear unambiguous symmetric pair arrangement
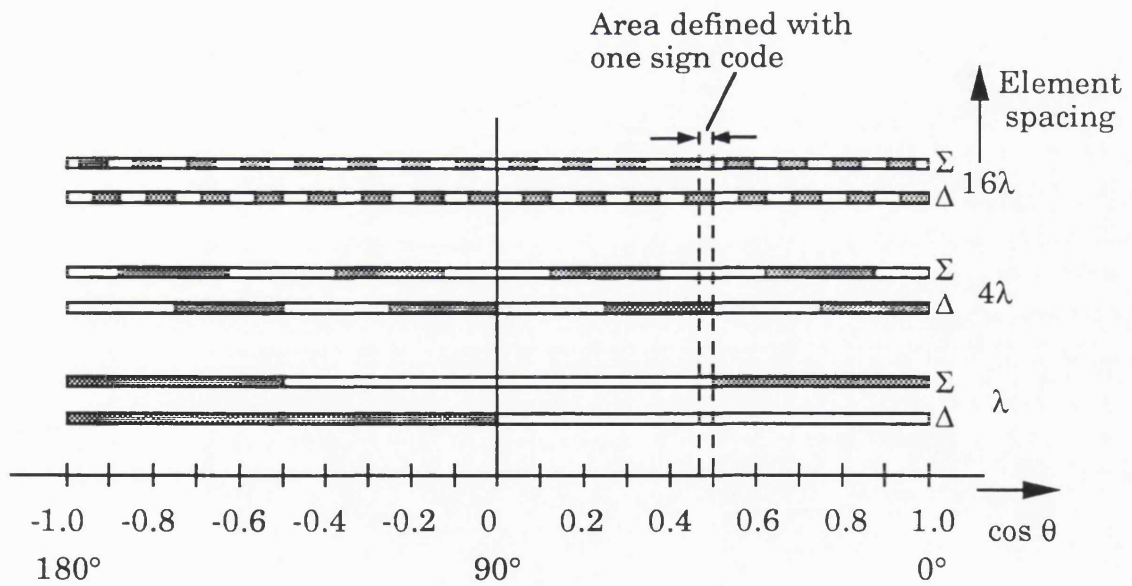
where: θ = Angle off endfire position

$\qquad$ = Positive sign bit

$\qquad$ = Negative sign bit

Figure 3-2: Sign bits of 3 collinear symmetric pairs for n=2 (quadrant selection using sum and difference)

with Σ and Δ processing. The average cone width for such an arrangement is defined as:

$$\beta = \frac{\pi}{4(d/\lambda)} = \frac{\pi\,\lambda}{4\,d}$$

...(3.2)

With d = 16 λ, the calculated average cone width is 2.8°.

Three mutual orthogonal sets of collinear pairs are used for direction finding. They produce three cones of common apex, but on mutually orthogonal axes, and these cones intersect to form a single pencil beam in the required direction. Its cross section is represented by the intersection of three annular zones on the surface of a sphere (Fig. 3-3). Due to the symmetry inherent to the symmetric-pair direction finding principle, it is possible to provide almost constant accuracy over the entire 3D space (see section 3.5.).
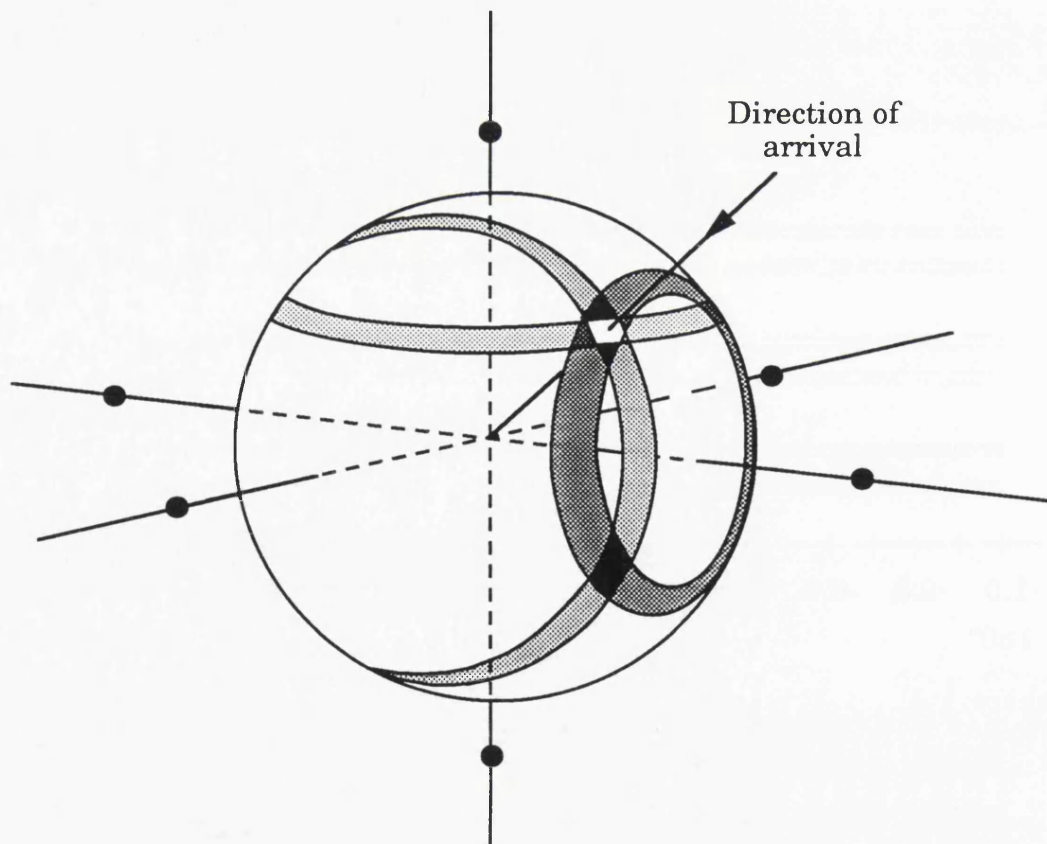
Figure 3-3: Intersection of the cones of three unambiguous mutually
orthogonal symmetric-pair arrangements

It should also be mentioned that two orthogonal axes are already sufficient for unambiguous DF over one hemisphere, since the remaining ambiguity is always in the other hemisphere, as can be seen in Fig. 3-3.

In practice, there may be considerable problems to build a system with three mutual orthogonal pairs with spacings $\leq \lambda$ (needed for ambiguity resolution) plus the reference element in the centre of the array without having considerable obscuration due to the elements themselves and their feed networks. This problem can be avoided by relocating at least one of the collinear axes. Each axis can work independently as long as it has a *phase* reference element. Knowing the exact location of all three axes, it is possible to take this into account for the calculation of the angles of arrival. If the signal source is far enough away, the error due to the distribution of the axes will be negligible. Table 3-1 shows the maximum DF error as a function of the spacing of two phase centres and the distance between the DF array and the source.

Max DF error due to the offset [arcsin (1/r)]	0.1°	0.2°	0.3°	0.4°	0.5°	1.0°
$r = \dfrac{\text{Distance between source and DF array}}{\text{Distance between the two phase centres}}$	573	286	191	143	115	57

Table 3-1:   Maximum DF error due to dislocated collinear axes

3.3. IMPROVEMENT OF THE DIRECTION FINDING ACCURACY FOR A GIVEN ARRAY GEOMETRY

The accuracy of this type of DF system can obviously be improved by adding extra antenna pairs with larger spacings. However, quite often the dimensions of the resulting system would simply be too large, or it would not be possible to achieve the mechanical rigidity of the structure required for reliable operation. It is therefore desirable to have techniques available that enhance the DF accuracy of a symmetric-pair array without enlarging the array size. Two possible options will be discussed in the following sections; the first method uses equal and opposite phase shifts, whereas the second technique specifically exploits properties of the Σ and Δ signals.

3.3.1. Equal and opposite phase shifts

The idea behind this technique is that equal and opposite phase shifts at the elements of a symmetric-pair shift the phase transition angles and hence the bit pattern of the pair (see chapter 2). The accuracy of the DF system can therefore be improved by measuring and processing the signs of a symmetric-pair for different well defined phase shifts. Fig. 3-4 shows an arrangement, which allows the simultaneous extraction of 4 bits of DF information from one pair. The first two bits of information are given by the signs of the Σ and Δ signal for $0°$ phase shift, and the third bit can be obtained by combining the Σ and Δ of the $+45°$ phase shifted signals with an *Exclusive*-OR operation. To extract the fourth bit, the sums and differences of the $+22.5°$ and $-22.5°$ shifted signals must be combined using three *Exclusive*-OR operations, as shown in Fig. 3-4. The corresponding sign bit patterns are given in Fig. 3-5. As can be seen, positive phase shifts at the reference element move the sign bit pattern to the left, and negative phase shifts move it to the right.

The extraction of DF information from one pair can either be done by parallel or sequential processing. The former allows to take bearings of signals which are present for only a short period of time. Since the signal is binary, 2^{n-1} measurements are needed to extract n bits of information from *one pair* with this kind of processing. Each positive and negative sector of the pair output signal will be subdivided into 2^{n-1} subsections, and the polarity of the extracted n bits can be obtained with simple logic, as shown in Fig. 3-5. However, the drawback of parallel processing is that considerably more hardware is needed to provide all the required signals simultaneously. By using $180°$ hybrids instead of power combiners to process the signals, as in Fig. 3-4, the number of required fixed phase shifters can be halved. This is because the bit patterns at the output of a hybrid are equivalent to those generated for a phase shift of $\Delta\phi$ (Σ port signal) and $\Delta\phi - \pi/2$ (Δ port signal) at the reference element of the pair.

If time is no limiting factor, intelligent sequential processing can be used to reduce the number of measurements. After each measurement, the required phase shift must be calculated, to subdivide the subsector identified through the previous measurement. Using this approach, only n measurements are necessary to extract n bits and the hardware
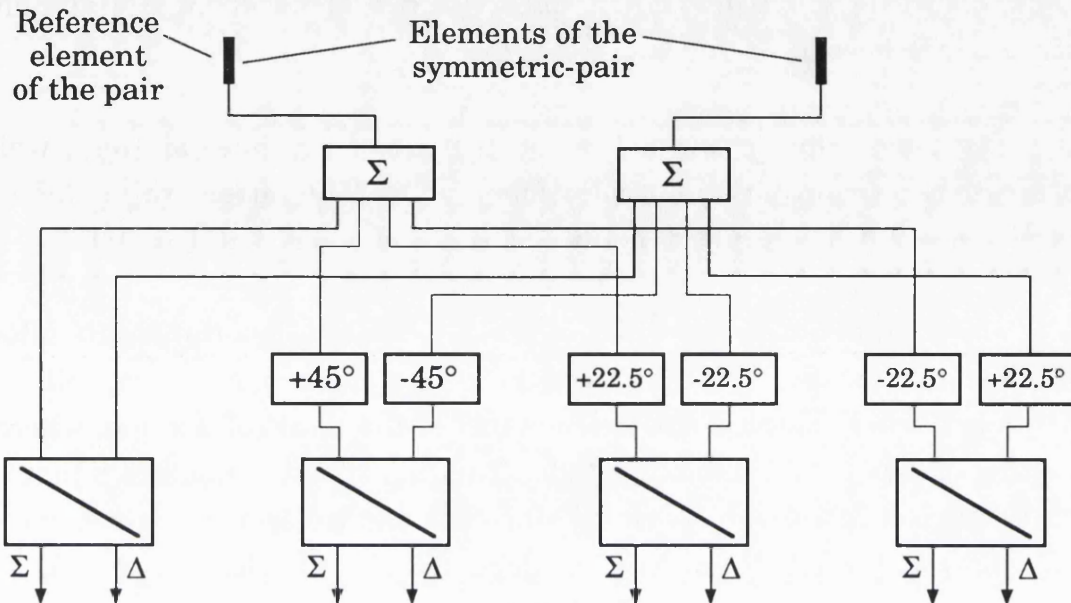
Figure 3-4: Processing hardware to enhance the DF accuracy of a single symmetric-pair
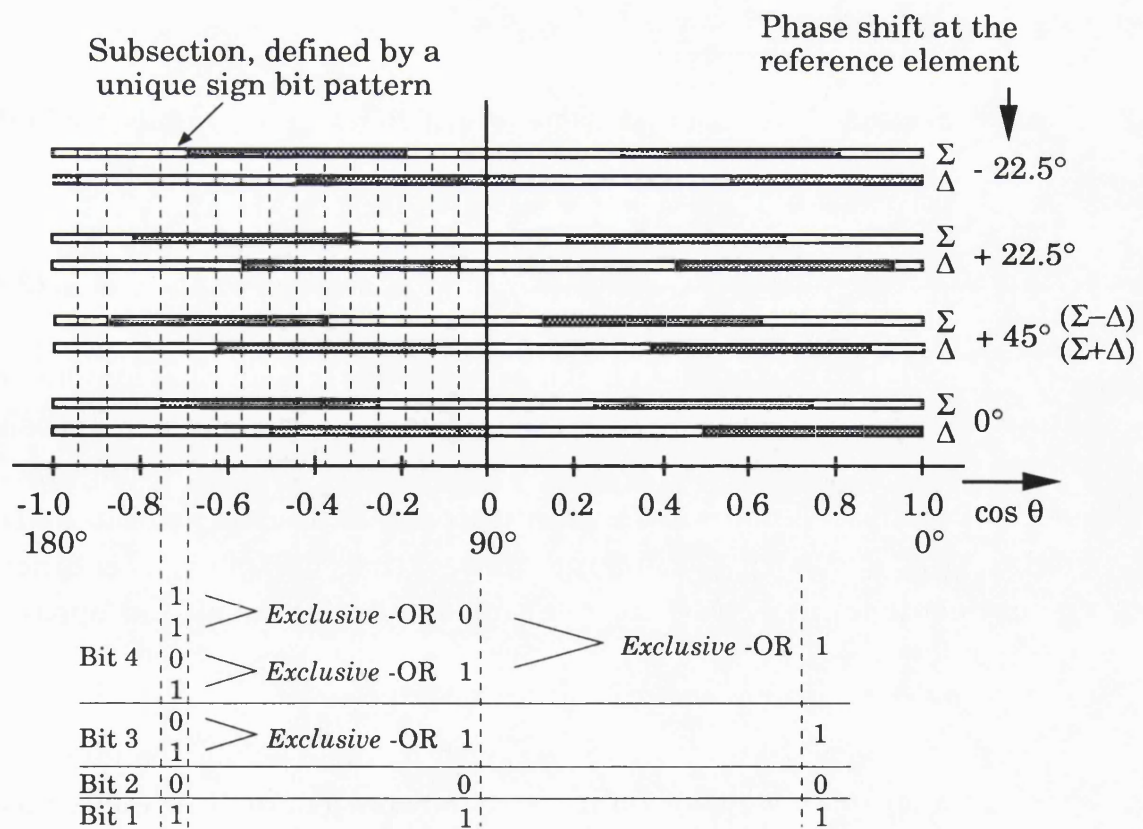
Figure 3-5: Sign bits of a symmetric-pair with Σ and Δ processing and equal and opposite phase shifts to extract 4 bits of DF information (2 λ spacing)

expenditure is much less. Just one hybrid and a set of digital phase shifters are needed.

The phase shifters can either be switched line phase shifters, which are rather narrow band, or vector modulators. The latter are broadband, but have the disadvantage of an insertion loss of about 6 dB [Bre 90].

The extraction of more bits of DF information from one pair allows larger differences in the spacing of consecutive pairs in a collinear arrangement. Each pair now selects $1/2^n$ of the total solid angle, where n is the number of bits extracted. (This is strictly speaking just an approximation for large spacings, since the endfire sectors cannot always be subdivided into 2^{n-1} subsections, dependent on the element spacing.) As a result, the spacing of each supplementary pair is now $1/2^n$ of the preceding one. Thus the number of pairs p_c for a collinear unambiguous configuration is:

$$p_c = 1 + \log_2 n \ (d/\lambda) = 1 + (1/n) \log_2 (d/\lambda) \qquad \qquad ...(3.3)$$

The average subsection or cone width for a given number of bits becomes now:

$$\beta = \frac{\pi \lambda}{d \ 2^n} \qquad \qquad ...(3.4)$$

The last two equations obviously only depend on the number of extracted bits and not on the processing method, be it equal and opposite phase shifts or enhanced Σ and Δ processing. The following table summarises the relationship between the number of extracted bits of DF information and the spacing of consecutive pairs in a collinear arrangement. It also gives the relevant details for equal and opposite phase shift processing.

It should be noted that the values given in Table 3-2 for the inter-pair spacing in an unambiguous collinear arrangement are theoretical ideal spacings. In practice, however, the spacing between successive pairs will be less to cater for imperfections and additional ambiguities due to the employed processing method, as will be discussed in chapter 4.

Number of bits extracted	Spacing for an unambiguous collinear arrangement	Settings of the phase shifter at the reference element	Hybrids for parallel processing
2	1/4 of the preceding pair	$0°$	1
3	1/8 of the preceding pair	$0°, +45°$	2
4	1/16 of the preceding pair	$0°, +45°,$ $+22.5°, -22.5°$	4

Table 3-2: Relationship between the number of bits extractable from one symmetric-pair and the corresponding inter-pair spacing

3.3.2. Enhanced Σ and Δ processing

Two techniques for enhanced Σ and Δ processing will be introduced in this section. Both of them only allow to extract a maximum of three bits of DF information, whereas the previous technique has theoretically no limitations.

The first method exploits that the difference and the sum of the Σ and Δ signals of a pair is equivalent to the output of a pair with equal and opposite phase shifts of $\pm 45°$ at the reference element (recall eqn. 2.6). With eqn. (2.2) and eqn. (2.3) from chapter 2, and ignoring the respective angular frequency terms, the Σ - Δ signal can be written as:

$$V_{\Sigma-\Delta} = \frac{1}{\sqrt{2}}\left(\sqrt{2}V\cos\phi - \sqrt{2}V\sin\phi\right) = \sqrt{2}V\cos(\phi + \pi/4)$$

...(3.5a)

and the Σ + Δ signal is then equal to:

$$V_{\Sigma+\Delta} = \frac{1}{\sqrt{2}}\left(\sqrt{2}V\cos\phi + \sqrt{2}V\sin\phi\right) = \sqrt{2}V\cos(\phi - \pi/4)$$

...(3.5b)

These two signals can be generated by feeding the Σ and Δ output of a symmetric pair into another hybrid after compensating the inherent 90° phase difference. Fig. 3-5 shows the corresponding bit pattern and Fig. 3-6 the block diagram. Due to the absence of the 45° phase shifters, the system is
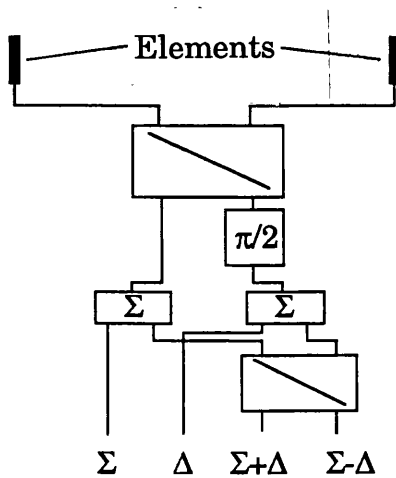
Figure 3-6: Generation of
Σ, Δ, Σ+Δ and Σ-Δ signals

now very broadband, and the frequency range of the hybrid or the bandwidth of the antenna elements remain the limiting factors.

The second processing technique uses the *amplitude* information provided by the Σ and Δ signals of the hybrid. The pattern defined by the sign of $|\Sigma| - |\Delta|$ is similar to that of the Σ pattern but of half the zone width. A third bit of information can now be extracted (beside the Σ and Δ bit), which allows to select 1/8 of the total solid angle with a single pair. This technique is also very broadband, and it should be noted that neither amplitude nor symmetrical phase errors affect the position of the sign transition of the third bit. With eqn (2.16) and its equivalent for the difference signal, it can be shown that the transition angles remain the same for amplitude errors, though, the amplitude values at the crossover points depend on the magnitude of the amplitude error.

3.4. ALGORITHMS FOR DIRECTION FINDING WITH SYMMETRIC-PAIRS

3.4.1. Outline of the procedure

The binary information provided by a symmetric-pair DF array contains the elevation and azimuth angle of the signal under consideration. The following section will show how to calculate these two parameters from the bit pattern.

A convenient way of extracting the angular information is to process the signals of two unambiguous axes at a time. The intersection of two cones defines the angle of arrival with one ambiguity, as previously illustrated in Fig. 3-3. This remaining ambiguity can be resolved with the additional information provided by a third axis. The minimum number of axes for unambiguous 3D direction finding is therefore three.

In order to calculate the angle of arrival, the following procedure must be adopted:

(1) Identify for each collinear axis the cone (or subsection) that is defined by the binary information of the symmetric-pairs (see Fig. 3-5).

(2) Calculate for each axis the angles $\phi_{conex,y}$ which describe the boundaries of each cone (Fig. 3-7).

(3) Choose a pair of axes and test whether the rims of the two relevant cones can intersect with each other.

(4) Calculate the elevation and azimuth angles for each possible intersection in a local coordinate system, which is defined with the help of the two axes under consideration. Determine dummy intersection angles, if less than four intersections are possible. (This happens for signals with low elevation relative to the plane defined by the two axes.)

(5) Transform the intersection angles into the reference coordinate system.

(6) Repeat steps (3) to (5) for the other possible permutations and resolve the last remaining ambiguity by comparing the calculated angles of arrival.

(7) Choose the intersection which covers the smallest solid angle and derive the angle of arrival from it.
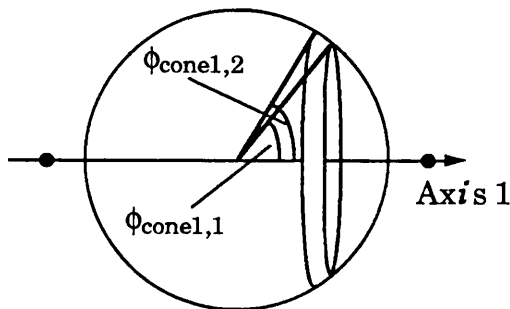
The first index denotes the axis, whereas the second index refers to the boundary. In the text, the second index is mostly omitted, since the relevant paragraphs describe in general the intersection of the rims of two cones from different axes, which excludes possible ambiguities.

Figure 3-7: Definition of the cone angles

The simplest way to get good accuracy is to choose the intersection of the narrowest cones to define the angle of incidence, as described in step 7 of the DF procedure. The area defined for the angle of incidence with this technique can sometimes be further reduced by exploiting the information provided by the cones from the remaining axes. However, this is very much dependent on the angle of arrival and the array geometry, and the improvement in accuracy is normally not such that the additional computing power needed for the relevant calculations would be justified.

Some of the steps of the DF algorithm outlined above are self-explanatory and can be implemented using formulae presented so far in this thesis. However, steps 3, 4 and 5 require more explanation and are therefore described in some detail in the next sections. To keep the algorithms as general as possible, any orientation of the axes of the DF array is allowed in the following analysis. In the case of three mutually orthogonal axes, some of the calculations will be significantly simplified.

3.4.2. Test for possible intersection of the rims of two cones

To examine whether it is possible to intersect the rims of two different cones, both cone angles per axis and the angle between the two axes must be known. The former angles have already been obtained in step two of the DF algorithm, whereas the latter one remains to be calculated.

The position of each axis in the DF reference coordinate system can be described in spherical DF coordinates. (This means that azimuth and elevation are defined as necessary for direction finding, see section 3.1. As a consequence, the cartesian system in the reference system will be a left hand system with x pointing to north and z to 90° elevation):

axis 1: α_1, ε_1
axis 2: α_2, ε_2

where:
$0° \leq \alpha < 360°$ and $-90° \leq \varepsilon \leq 90°$

With

$$x_k = \cos \alpha_k \cos \varepsilon_k \qquad \qquad ...(3.6a)$$

$$y_k = \sin \alpha_k \cos \varepsilon_k \qquad \qquad ...(3.6b)$$

$$z_k = \sin \varepsilon_k \qquad \qquad ...(3.6c)$$

the coordinates of the two axes can be transformed into cartesian coordinates. The index k describes the kth axis. The x, y, z coordinates describe the points where the axes pierce the surface of a unit sphere. The distance l between these two points is given by:

$$l = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \qquad ...(3.7)$$

Using a formula for triangle calculations [Bro 83], the wanted angle γ between the two axes under consideration is equal to:

$$\gamma = 2 \arccos \sqrt{1 - \frac{l^2}{4}} \qquad ...(3.8)$$

All necessary parameters for the intersection test are now known. Tables 3-3a and 3-3b show the conditions to be satisfied by the cone angles in order to allow intersection of the cones. The hatched areas in the pictures describe the hemisphere in which the cone of each axis lies, and the region, defined by the intersection of the two hatched areas defines where the intersection will take place. The sketches in the last column of each table describe what would happen if the conditions for the intersection of two cones were not fulfilled. Intersection of the two relevant cones is impossible if none of the conditions listed in both tables can be fulfilled.

3.4.3. Calculation of the cone-cone intersections in the local coordinate system

After identification of the possible intersections, it is necessary to calculate their coordinates. In order to do this, it is convenient to introduce a local coordinate system that allows the description of the intersections in terms of a local azimuth and elevation angle. This coordinate system is

Angle between the axes: $\gamma \leq 90°$	Conditions for intersection	Situations which arise if the conditions are not fulfilled
⟋ : cancel out		
$\phi_{cone1} \leq 90°$ and $\phi_{cone2} \leq 90°$ Axis 1, γ, Axis 2	$\phi_{cone1} + \phi_{cone2} \geq \gamma$ and $\left|\phi_{cone1} - \phi_{cone2}\right| \leq \gamma$	ϕ_{cone1}, γ, ϕ_{cone2} ... ϕ_{cone2}, ϕ_{cone1}, γ
$\phi_{cone1} > 90°$ and $\phi_{cone2} \leq 90°$ Axis 1, γ, Axis 2	$\left(180° - \phi_{cone1}\right) + \phi_{cone2} \geq 180° - \gamma$	$180° - \phi_{cone1}$, γ, ϕ_{cone2}
$\phi_{cone1} > 90°$ and $\phi_{cone2} > 90°$ Axis 1, γ, Axis 2	$\left(180° - \phi_{cone1}\right) + \left(180° - \phi_{cone2}\right) \geq \gamma$ and $\left|\left(180° - \phi_{cone1}\right) - \left(180° - \phi_{cone2}\right)\right| \leq \gamma$	$180° - \phi_{cone2}$, γ, $180° - \phi_{cone1}$ ... $180° - \phi_{cone2}$, γ, $180° - \phi_{cone1}$
$\phi_{cone1} \leq 90°$ and $\phi_{cone2} > 90°$ Axis 1, γ, Axis 2	$\phi_{cone1} + \left(180° - \phi_{cone2}\right) \geq 180° - \gamma$	ϕ_{cone1}, γ, $180° - \phi_{cone2}$

Table 3-3a: Conditions for the possible intersection of the rims of two cones for $\gamma \leq 90°$
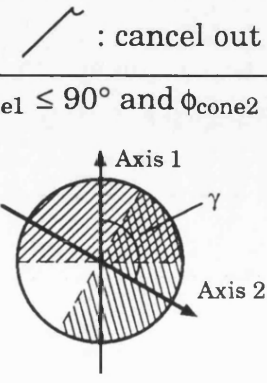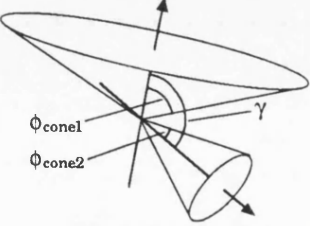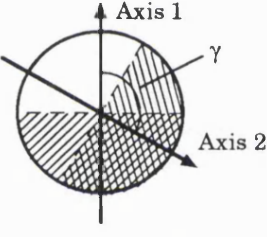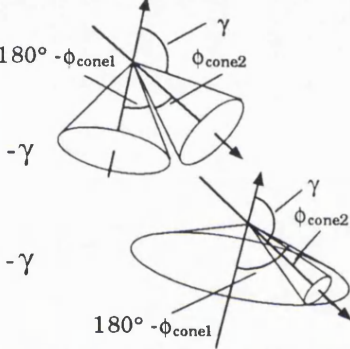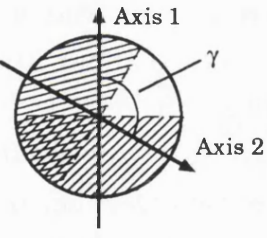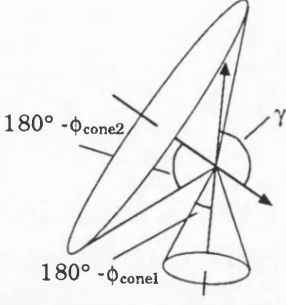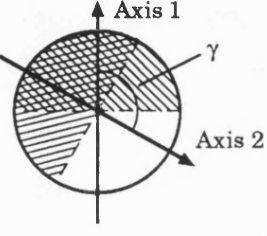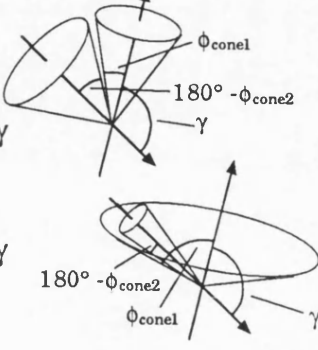
Angle between the axes: $\gamma > 90°$	Conditions for intersection	Situations which arise if the conditions are not fulfilled
$\diagup$: cancel out		
$\phi_{cone1} \leq 90°$ and $\phi_{cone2} \leq 90°$	$\phi_{cone1} + \phi_{cone2} \geq \gamma$	
$\phi_{cone1} > 90°$ and $\phi_{cone2} \leq 90°$	$\left(180° - \phi_{cone1}\right) + \phi_{cone2} \geq 180° - \gamma$ and $\left|\left(180° - \phi_{cone1}\right) - \phi_{cone2}\right| \leq 180° - \gamma$	
$\phi_{cone1} > 90°$ and $\phi_{cone2} > 90°$	$\left(180° - \phi_{cone1}\right) + \left(180° - \phi_{cone2}\right) \geq \gamma$	
$\phi_{cone1} \leq 90°$ and $\phi_{cone2} > 90°$	$\phi_{cone1} + \left(180° - \phi_{cone2}\right) \geq 180° - \gamma$ and $\left|\phi_{cone1} - \left(180° - \phi_{cone2}\right)\right| \leq 180° - \gamma$	

Table 3-3b: Conditions for the possible intersection of the rims of two cones for $\gamma > 90°$

given by the two element axes ξ and η which define the azimuth plane, and the azimuth angle is counted clockwise from the ξ axis. The elevation is as usual defined relative to the azimuth plane. The derivation of the elevation and azimuth angles of the intersections in the local coordinate system is given in Appendix 2. The local azimuth angle α_ℓ is then defined as follows (eqn. (A2.10)):

$$\alpha_\ell = \arctan\left(\frac{\cos(\phi_{cone2}) - \cos(\phi_{cone1})\cos(\gamma)}{\sin(\gamma)\cos(\phi_{cone1})}\right)$$

...(3.9)

where ϕ_{cone1} and ϕ_{cone2} are the cone angles of the two intersecting rims and γ the angle between axes ξ and η. To get the correct azimuth angle, the ambiguity of the arctangent has to be taken into account. The two possible elevation angles $\varepsilon_{\ell 1/2}$ per intersection are given by (eqn. (A2.8)):

$$\varepsilon_{\ell 1/2} = \pm\arcsin\sqrt{\left(\sin(\phi_{cone1})\right)^2 - \left(\frac{\cos(\phi_{cone2}) - \cos(\phi_{cone1})\cos(\gamma)}{\sin(\gamma)}\right)^2}$$

...(3.10)

When the incident signal has very low elevation relative to the plane defined by the two axes, up to three out of four possible intersections might not be possible, as shown in Fig. 3-8. This entails most of the time a rather large cross sectional area of the two cones, which makes them unusable for accuracy evaluations. However, the intersections can still be used to resolve ambiguities if dummy intersection angles are introduced (see Fig. 3-*8*). The following rules define how to choose the correct cone angle as dummy intersection angle:

(1) Three intersections were possible:
   Dummies are the cone angles which are just involved in one of the three intersections (or *not* in two intersections).

(2) Two intersections were possible:
   Dummies are the cone angles which were only involved in one intersection (or *not* in two and *not* in no intersection).

(3) One intersection was possible:
   Dummies are the cone angles which were involved in the possible intersection (or *not* involved in no intersection).

Typical examples for the intersection of cones

4 intersections possible:
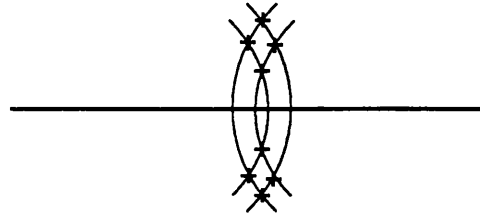
3 intersections possible:
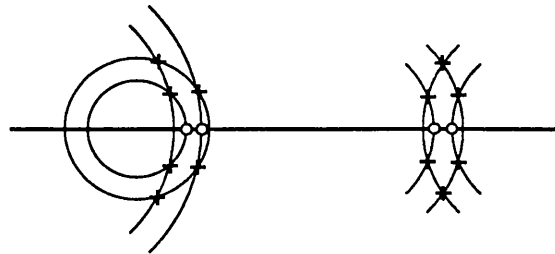
2 intersections possible:
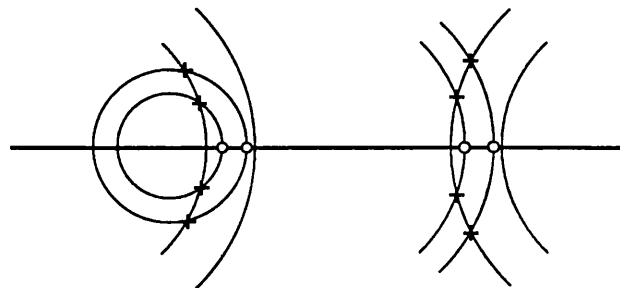
1 intersection possible:
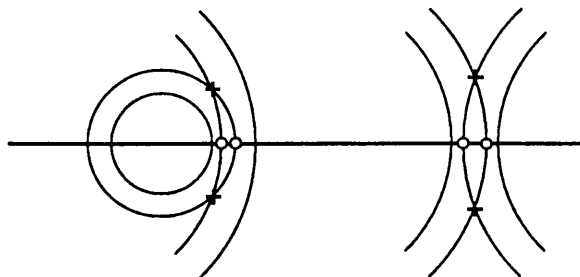
+ : possible intersection   o : dummy intersection

Figure 3-8: Possible intersections of cones

With these additional dummy angles, the areas defined by the intersection of two cones have approximately the shape of a triangle, a quadrangle or a pentagon (see Fig. 3-8). It should also be noted that the intersections of an endfire beam and a cone cannot be used for accuracy evaluations. This is due to the unfavourable shape of these intersections.

Having calculated the coordinates of all possible intersections and worked out the necessary dummy angles, it is now possible to transform the data into the reference system.

3.4.4. Coordinate transformation into the reference system and calculation of the angle of arrival

For this transformation, a local coordinate system must be defined, which is not necessarily cartesian but oblique. Using the two element axes as ξ and η axes (as in the previous section), the third axis ζ can be calculated by computing the normalised vector product. ζ is orthogonal to ξ and η, and eqns. (3.11a-c) describe the third axis in the cartesian reference system. The coordinates x_3, y_3, z_3 describe the point where the ζ axis pierces the surface of a unit sphere:

$$x_3 = (y_1 \, z_2 - z_1 \, y_2) \, / \sin\gamma \qquad\qquad ...(3.11a)$$
$$y_3 = (z_1 \, x_2 - x_1 \, z_2) \, / \sin\gamma \qquad\qquad ...(3.11b)$$
$$z_3 = (x_1 \, y_2 - y_1 \, x_2) \, / \sin\gamma \qquad\qquad ...(3.11c)$$

where $x_{1/2}, y_{1/2}, z_{1/2}$ are the coordinates of the vectors of the two element axes in the cartesian reference system and γ the angle between them. The three axes form a left-hand coordinate system. Having defined the oblique coordinate system, it is possible to express the coordinates of the intersections in the ξ, η, ζ system:

$$\xi_i = \cos(\alpha_\ell) \, \cos(\varepsilon_\ell) \qquad\qquad ...(3.12a)$$
$$\eta_i = \cos(\gamma - \alpha_\ell) \, \cos(\varepsilon_\ell) \qquad\qquad ...(3.12b)$$
$$\zeta_i = \sin(\varepsilon_\ell) \qquad\qquad ...(3.12c)$$

The index i describes the ith intersection.

For the coordinate transformation from the local oblique system to the cartesian reference system, all direction cosines must be known [Bel 31]. With eqns (3.6) to (3.8), it is possible to calculate the angles between the coordinate axes which are required for the determination of the direction cosines. The equations for the transformation into the cartesian reference system are the following:

$$x_i = \frac{(m_2 n_3 - m_3 n_2)\xi_i + (l_3 n_2 - l_2 n_3)\eta_i + (l_2 m_3 - l_3 m_2)\zeta_i}{DET}$$

...(3.13a)

$$y_i = \frac{(m_3 n_1 - m_1 n_3)\xi_i + (l_1 n_3 - l_3 n_1)\eta_i + (l_3 m_1 - l_1 m_3)\zeta_i}{DET}$$

...(3.13b)

$$z_i = \frac{(m_1 n_2 - m_2 n_1)\xi_i + (l_2 n_1 - l_1 n_2)\eta_i + (l_1 m_2 - l_2 m_1)\zeta_i}{DET}$$

...(3.13c)

with:

$$DET = l_1 m_2 n_3 + m_1 n_2 l_3 + n_1 l_2 m_3 + n_1 m_2 l_3 + m_1 l_2 n_3 + l_1 n_2 m_3$$

The variables l_x, m_x and n_x are direction cosines as defined in Table 3-4.

Direction cosine between axis		reference system		
		x	y	z
local oblique system	ξ	l_1	m_1	n_1
	η	l_2	m_2	n_2
	ζ	l_3	m_3	n_3

Table 3-4: Definition of the different direction cosines

The intersection of the two cones is now expressed in the cartesian reference system and can be transformed into spherical DF coordinates as follows:

$$\varepsilon_i = \arcsin(z_i)$$

...(3.14a)

$$\alpha_i = \arcsin\left(\frac{x_i}{\sin(\varepsilon)}\right)$$

...(3.14b)

or:

$$\alpha_i = \arccos\left(\frac{y_i}{\cos(\varepsilon)}\right)$$

...(3.14c)

Dependent on the values of x_i and y_i, eqn (3.14b) or (3.14c) must be used respectively.

The possible regions for the angle of arrival can now be identified by applying this procedure to all feasible intersections of the cones of the two collinear axes under consideration. The remaining ambiguity is resolved with the information provided by a second pair of axes. The centre of the thus defined region for the angle of arrival is then the value of the bearing.

3.5. SYSTEM INHERENT LIMITATIONS ON THE DIRECTION FINDING ACCURACY

As long as the systematic errors can be kept low, all conventional DF systems have for a single signal, in theory, high accuracy in azimuth and, provided the applied DF technique allows it, also in elevation. Due to its quasi binary signals, an error-free symmetric-pair system has certain unavoidable tolerances in the azimuth and elevation bearings. However, by choosing an appropriate array configuration, these system inherent inaccuracies can be kept very small. The purpose of this section is to investigate these intrinsic DF errors of an ideal symmetric-pair direction finder.

After resolution of the ambiguities, the angle of arrival is established in a symmetric-pair DF system by intersecting the two most favourable cones, as described in section 3.2. The shape of this intersection defines the corresponding azimuth and elevation errors. These system inherent errors can be described with the following parameters:

(1) Solid angle of the area defined by the intersection of the two cones.

(2) Maximum width of this region in azimuth. (The maximum azimuth error is then defined as ± half the width, see also Fig. 3-12.)

(3) Maximum width of this region in elevation. (The maximum elevation error is then defined as ± half the width, see also Fig. 3-12.)

(4) RMS value of the maximum azimuth and elevation error as defined in (2) and (3).

In fact, solid angle and maximum azimuth and elevation error define the inaccuracies quite comprehensively, while the RMS error parameter just gives supplementary information.

Due to the nature of symmetric-pairs, it is obvious that the highest accuracy will be achieved for signals coming from directions that are normal to a plane, defined by two of the collinear axes. (In an orthogonal system, this will be the direction of the third axis.) For these angles of incidence, two very narrow cones intersect and give high accuracy, whereas the third cone is just used to resolve the last ambiguity. If the angles of arrival are close to the plane of the two collinear axes, another set of axes should be used to achieve high accuracy.
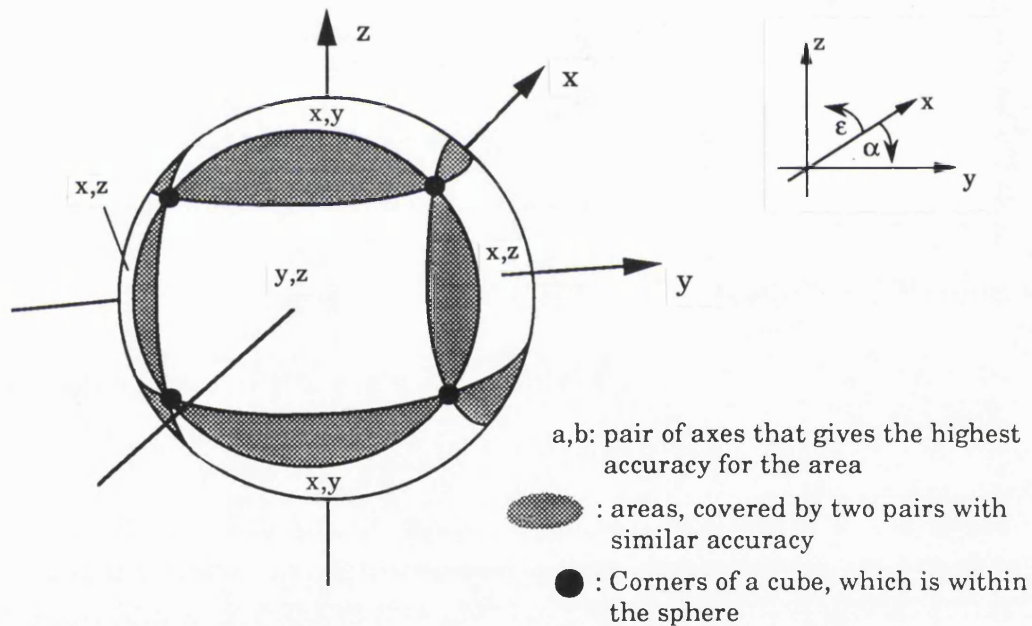
Figure 3-9 *Areas where pairs of axes of an orthogonal system yield the highest accuracy*

For the case of three mutually orthogonal axes, Fig. 3-9 shows which pair of axes gives the best accuracy for which angles of incidence. The points used to define the border between the different areas are the corners of a cube which lies within the sphere. Table 3-5 presents the angles of arrival that are covered by each pair. The azimuth and elevation angles of the corners of the cube can be calculated using the relevant geometrical relationships as illustrated in Fig. 3-10.

Areas covered \ Pair of axes	x,y	y,z	x,z
α	0° to 360°	315° (-45°) to 45°	45° to 135°
ε	35.3° to 90°	-35.3° to 35.3°	-35.3° to 35.3°
α	0° to 360°	135° to 225°	225° to 315°
ε	-90° to -35.3°	-35.3° to 35.3°	-35.3° to 35.3°

Table 3-5:   Areas covered by a pair of axes for accuracy evaluations

$$\Rightarrow \varepsilon = \arctan\frac{1}{\sqrt{2}} = 35.26°$$
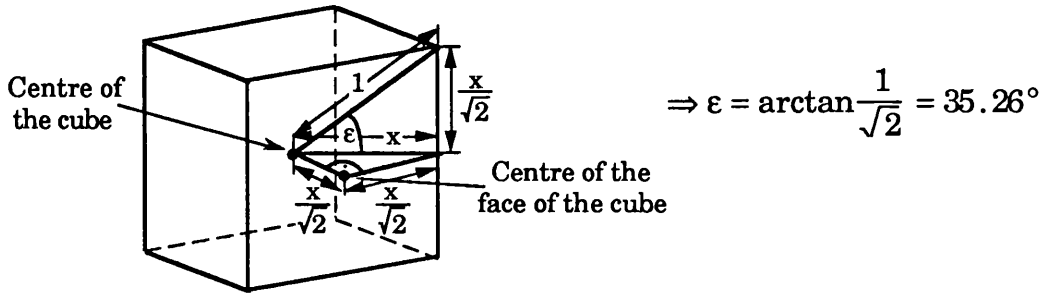
Figure 3-10: Geometrical relationships within the cube

3.5.1. Accuracy of an orthogonal system

The first configuration examined is a set of three mutually orthogonal axes with a maximum element spacing of 32 λ and simple sum processing. (The same result could be obtained with smaller maximum spacings if more sophisticated processing was used.) Since the spherical coordinate system that defines the azimuth and elevation angle does not support the 3D symmetry of the arrangement as a cartesian system would do, different pairs of axes yield different graphs.

Figs. 3-11a-d show plots of the four parameters that describe the system-inherent errors for angles of incidence with azimuth angles between 0° and 360° and elevation angles between 30° and 90°. The information provided by the x and y axis can be used to calculate the inaccuracies for this range of angles of arrival (see Table 3-5). The maximum azimuth and elevation errors can be evaluated from the coordinates of the intersections of the rims of the two cones. Since the azimuth information loses its relevance with increasing elevation, the azimuth error data *of all plots in this document* are tapered with a cos(ε)

term to take this into account. The following surface integral must be solved to calculate the exact solid angle covered by each intersection:

$$\int_{y_1}^{y_2}\int_{x_1}^{x_2} \frac{1}{\sqrt{1-y^2-x^2}}\, dx\, dy$$

Where $x_{1/2}$ and $y_{1/2}$ are the projections of the cone angles on the relevant element axes (e.g.: $x_1 = \cos(\phi_{conel,1})$). After some manipulation (see Appendix 3), the following solution for the integral can be obtained:

$$\int_{y_1}^{y_2}\int_{x_1}^{x_2} \frac{1}{\sqrt{1-y^2-x^2}}\, dx\, dy =$$

$$y_2 \arcsin\left(\frac{x_2}{\sqrt{1-y_2^2}}\right) + \frac{x_2}{2}\arcsin\left(\frac{x_2^2-1-2y_2^2}{\left|x_2^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_2^2+1\right)\left(1-y_2^2\right)-2x_2^2}{\left(1-y_2^2\right)\left|x_2^2-1\right|}\right) - $$

$$y_1 \arcsin\left(\frac{x_2}{\sqrt{1-y_1^2}}\right) + \frac{x_2}{2}\arcsin\left(\frac{x_2^2-1-2y_1^2}{\left|x_2^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_2^2+1\right)\left(1-y_1^2\right)-2x_2^2}{\left(1-y_1^2\right)\left|x_2^2-1\right|}\right) - $$

$$y_2 \arcsin\left(\frac{x_1}{\sqrt{1-y_2^2}}\right) + \frac{x_1}{2}\arcsin\left(\frac{x_1^2-1-2y_2^2}{\left|x_1^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_1^2+1\right)\left(1-y_2^2\right)-2x_1^2}{\left(1-y_2^2\right)\left|x_1^2-1\right|}\right) + $$

$$y_1 \arcsin\left(\frac{x_1}{\sqrt{1-y_1^2}}\right) + \frac{x_1}{2}\arcsin\left(\frac{x_1^2-1-2y_1^2}{\left|x_1^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_1^2+1\right)\left(1-y_1^2\right)-2x_1^2}{\left(1-y_1^2\right)\left|x_1^2-1\right|}\right) $$

...(3-14)

where: $|...|$ means absolute value

This is a rather lengthy expression and in order to save computing time, the calculation of the solid angle was somewhat simplified for the simulation program. Since the areas of intersection are very small for large spacings, they can be considered as being plane and the bounds can be approximated by straight lines between the intersections of the cones. This simplifies the calculation and also allows us to evaluate the solid angle defined by the intersection of two cones which are not generated by orthogonal axes. Furthermore it is possible to calculate with this approach the area of intersection for cases where less than four cone
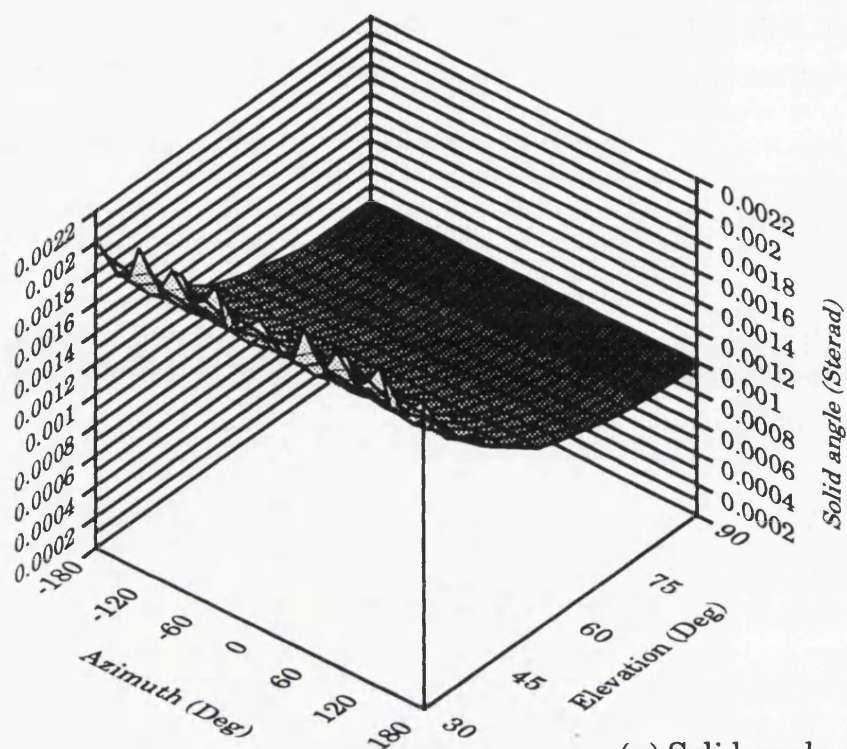
(a) Solid angle covered
by each intersection
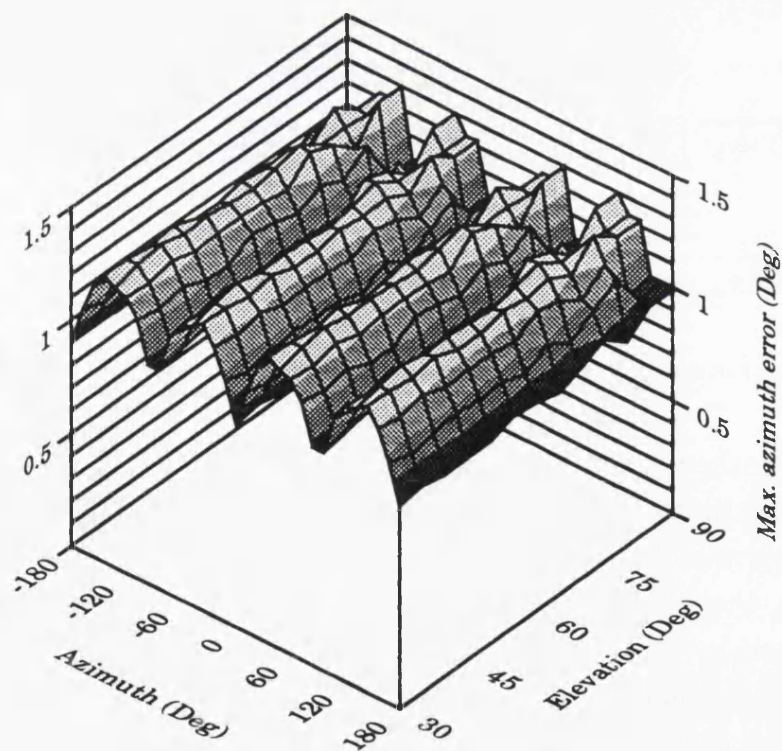
(b) Maximum azimuth error

*Figure 3-11: System inherent DF errors for $0° \leq \alpha \leq 360°$ and $30° \leq \varepsilon < 90°$
using the x and y axes for the accuracy evaluations*
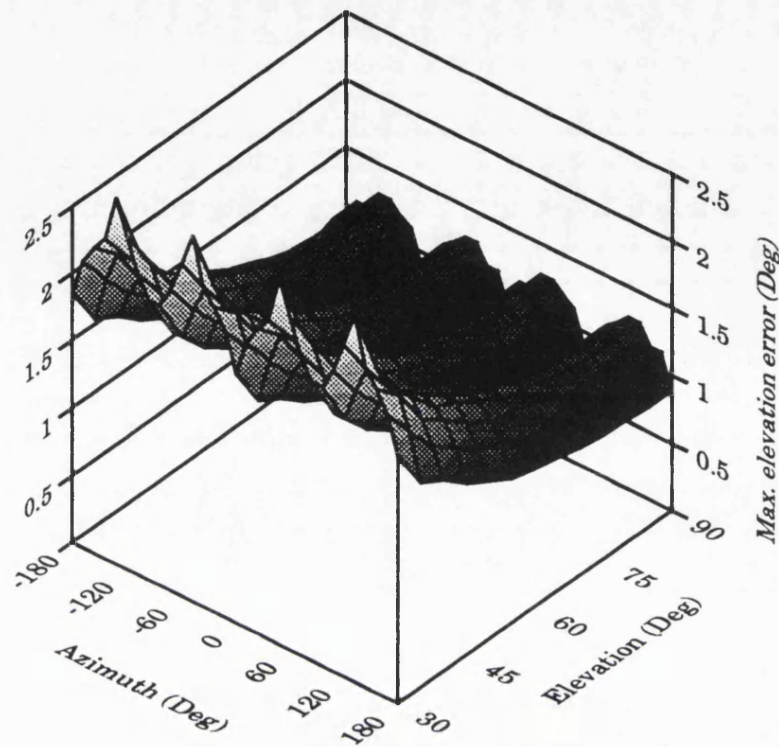
(c) Maximum elevation error
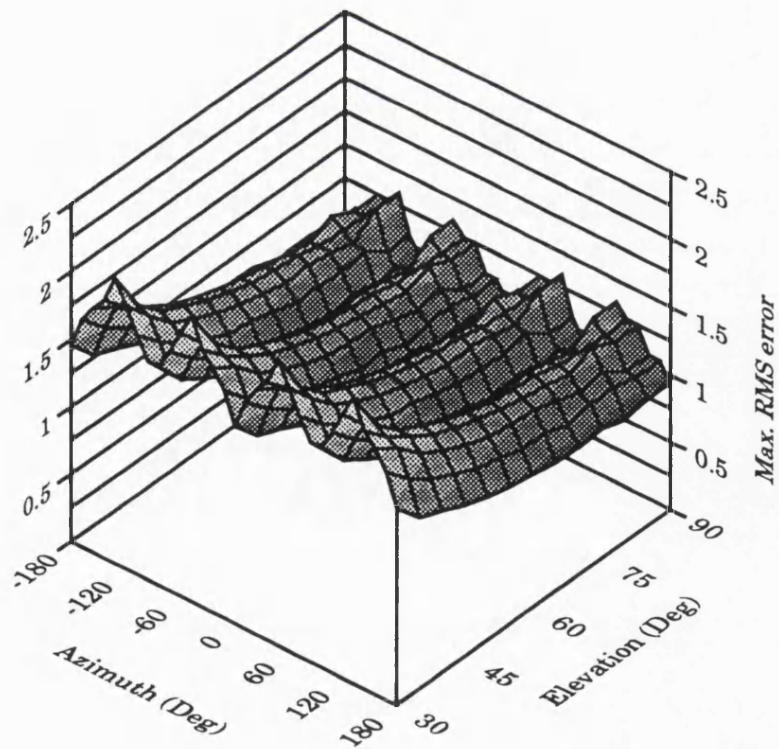
(d) Maximum RMS error

Figure 3-11: System inherent DF errors for $0° \leq \alpha \leq 360°$ and $30° \leq \varepsilon < 90°$ using the x and y axes for the accuracy evaluations

intersections are possible. A polygon can be defined with the additional dummy angles, and its area is calculated by subdividing it into triangular sectors whose areas can be readily evaluated. For a spacing of 32 λ of the outer pairs and sum processing, the difference between the exact and approximated value of the solid angle for four possible intersections was less than 2%, which confirms the validity of the simplified computations.

The solid angle covered by each intersection remains virtually constant for different azimuth angles at constant elevation, as shown in Fig. 3-11a. To explain this, it is useful to imagine the contour of the sphere projected into the plane which contains the two element axes (see Fig. 3-12). The sign zones of the two orthogonal pairs cut the surface of the sphere into squares (or oblongs, if they have different spacings) of equal size, and all angles of incidence with the same elevation angle ε lie on a circle with radius cos ε. Thus, the areas defined by the intersection of the cones for a certain elevation angle have virtually the same gradient on the surface of the sphere, which results in approximately the same solid angle. Comparison of Fig. 3-11b and Fig. 3-11c shows that the azimuth and elevation error have maxima and minima for the same azimuth angles, which is due to the contour of the intersections. Fig. 3-13 illustrates how their approximate shape changes for different azimuth angles from a square to a diamond and back again to a square, all with the same area. This explains why the solid angle covered by the intersection remains virtually constant, despite the fact that the maximum azimuth and elevation error grow and decrease simultaneously.
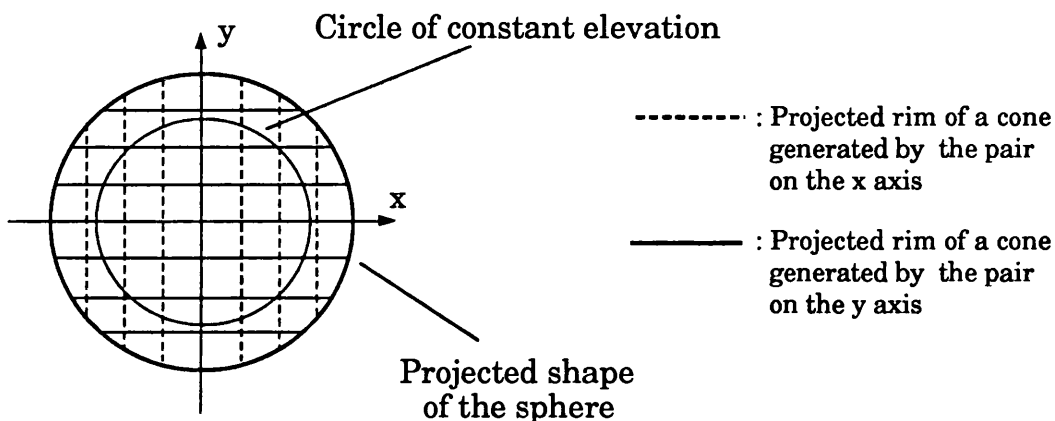
Figure 3-12: Projection of the sphere into the plane of the two element axes
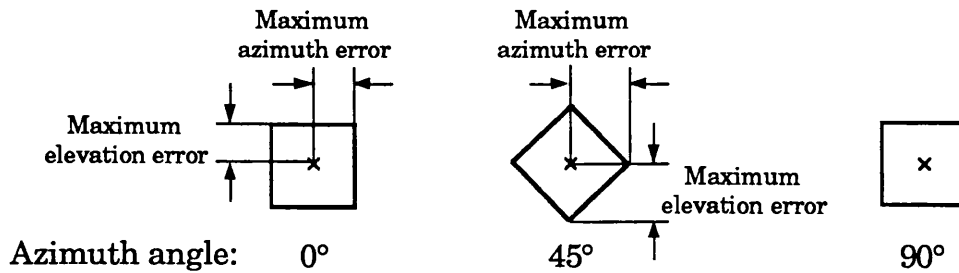
Figure 3-13: Shape of the intersection areas

In Figs. 3-14a-d, the four error parameters are shown for azimuth angles from 45° to 135° and elevation angles between -45° and +45°. This time, the x and z axes are used for the accuracy calculations, and the largest spacing of the pairs of each axis was again 32 λ. The graph for the solid angle (Fig. 3-14a) shows nicely that the best accuracy (or smallest solid angle) is obtained for a direction normal to the two axes, which is in this case along the y axis ($\alpha = 90°$, $\varepsilon = 0°$). It should be noted that the constant value of the maximum azimuth error (Fig. 3-14b) for an azimuth angle of 90° is due to the fact that this represents the width of the boresight disc generated by the pair on the x axes, which obviously remains constant for different elevation angles. The maximum elevation error (Fig. 3-14c) is in this arrangement equivalent to the thickness of the cones generated by the pair on the z axis and is therefore independent of the azimuth angle. The cones are narrow for elevation around 0° and grow for increasing elevation angles.

Using the y and z axes for accuracy evaluations, similar plots to the ones for the x and z axes. are obtained This is because the axes of the two systems just differ by a 90° offset in azimuth.

These simulations also show that the maximum variation in the accuracy of an orthogonal symmetric-pair DF system is only $\sqrt{2}$ for azimuth and elevation errors, and a factor of 2 for the corresponding solid angle over the entire 3D space. The system yields the worst accuracy at points which are represented by the corners of the cube within the sphere (Fig. 3-9).

It should be mentioned, that the algorithms derived in section 3.4 are used in the computer simulations. The programs have been written in Fortran and run on a Macintosh™ computer.
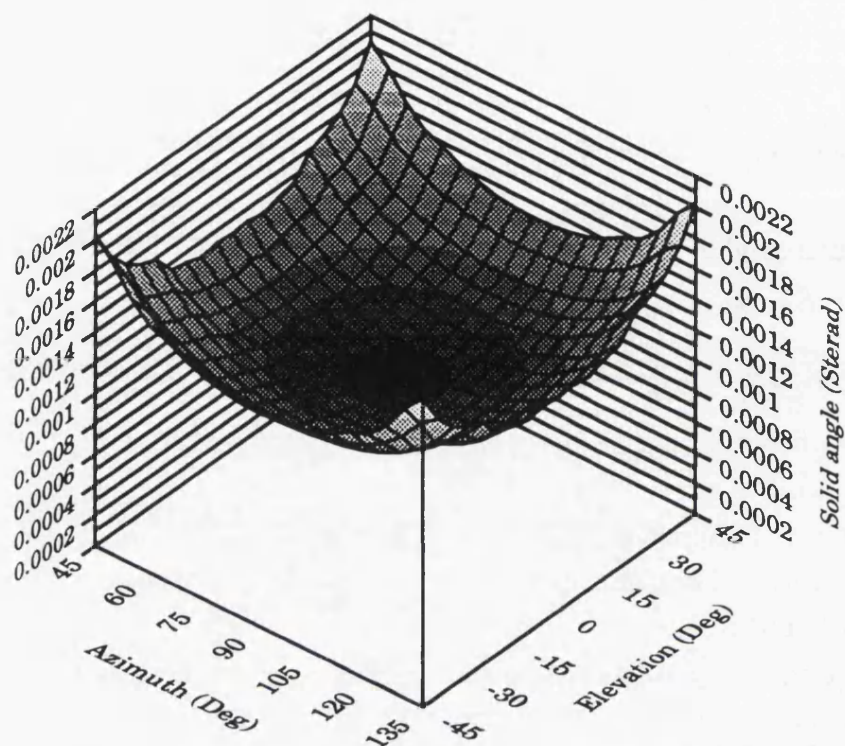
(a) Solid angle covered
by each intersection
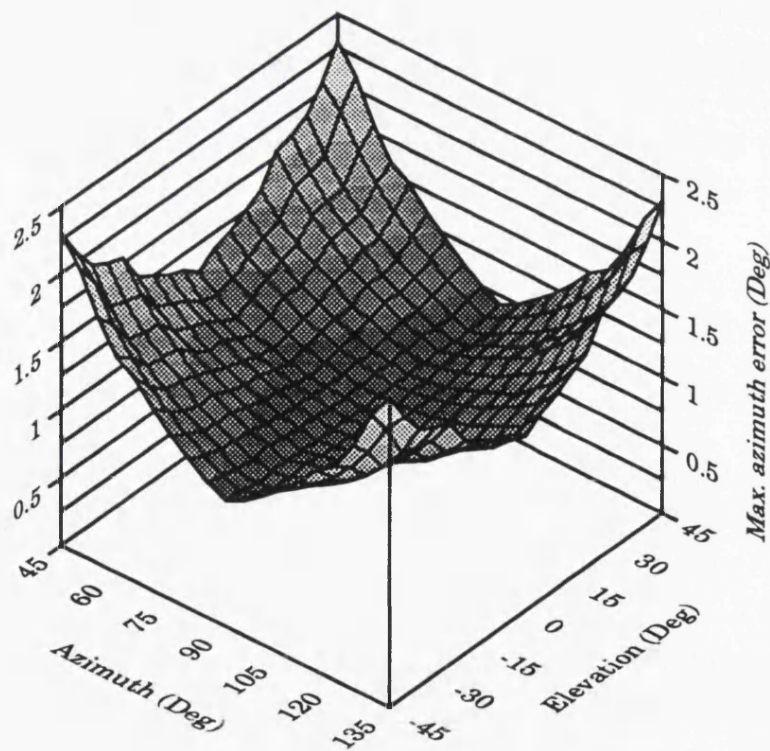
(b) Maximum azimuth error

Figure 3-14: System inherent DF errors for $45° \leq \alpha \leq 135°$ and $-45° \leq \varepsilon \leq 45°$
using the x and z axes for the accuracy evaluations
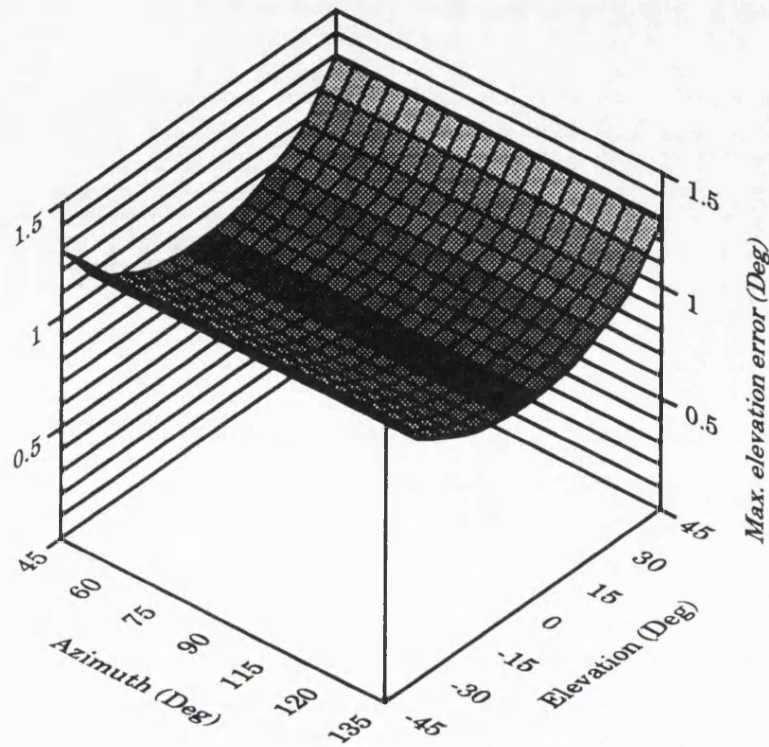
(c) Maximum elevation error
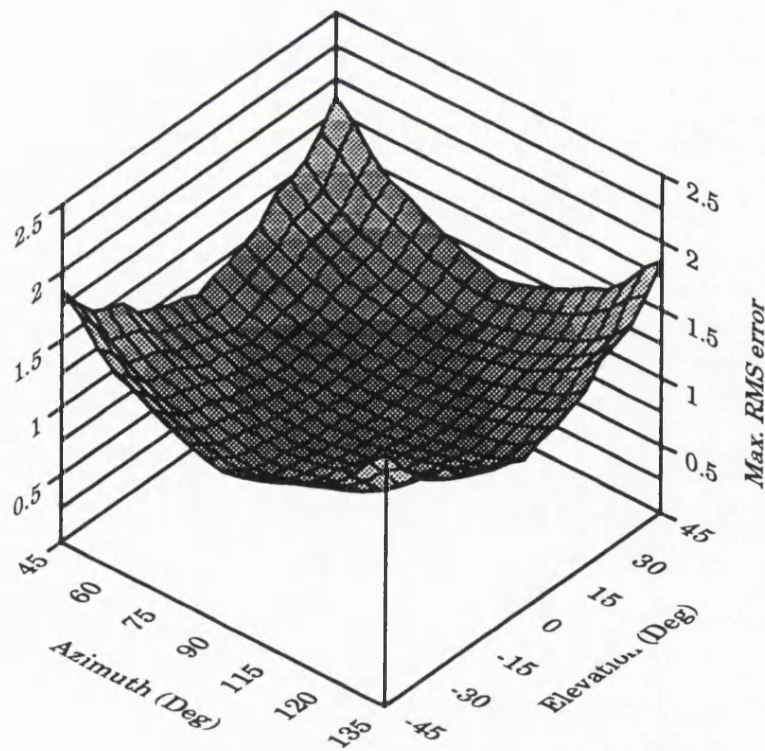
(d) Maximum RMS error

Figure 3-14: System inherent DF errors for $45° \le \alpha \le 135°$ and $-45° \le \varepsilon \le 45°$ using the x and z axes for the accuracy evaluations

3.5.2. Accuracy of an oblique system

A system with three orthogonal axes allows symmetrical and uniform coverage of the 3 dimensional space. The next symmetrical arrangement in 3D with more than 3 axes would be a system with 7 axes, where the additional four axes are the diagonals of a cube, whose centre is the origin of a cartesian system. However, this arrangement will just enhance the overall accuracy if more complicated processing is used, i.e. more than just the information of the most favourable pair of axes is used for the accuracy calculations. A system with three mutually orthogonal axes is otherwise the optimum configuration.

In order to demonstrate this, simulations of the system inherent errors of a symmetric-pair system with an angle between the two element axes of less than 90° were carried out. Since the main objective is to demonstrate the tendency of the errors, smaller sets of data have been used. Figs. 3-15a-f show the change of the maximum azimuth error as a function of the angle between the two element axes. Both axes lie in the azimuth plane, whereas one of them is at 0° azimuth and the pair spacing is again 32 λ. It can clearly be seen the metamorphosis from a quadrant symmetrical error to an error that is just symmetrical to a hemisphere and steadily growing for smaller angles between the axes. If the two axes coincide ($\gamma = 0°$), the azimuth error would be equivalent to the cone width of the two pairs. The maximum azimuth error at 0° and 180° would be the endfire beamwidth of the pair and the error at ±90° would be the beamwidth of the broadside disc. Figs. 3-16a and 3-16b show the increase in the solid angle covered by each intersection. The elevation error for different γ is presented in Figs. 3-17a and 3-17b. For the unrealistic case of $\gamma = 0°$, the elevation error would be ±90°, which would make any discrimination in elevation impossible.

The computer simulations show clearly that any spacing between the element axes other than 90° would entail a less accurate DF system, if just the information provided by two collinear axes was used for the determination of the angle of incidence. The recommended structure for a 3D DF system is therefore one with three mutually orthogonal axes, because this preserves the simplicity of the symmetric-pair approach and keeps the processing hard- and software at a reasonable level.

(a) γ = 90°

(b) γ = 75°

(c) γ = 60°

(d) γ = 45°

(e) γ = 30°
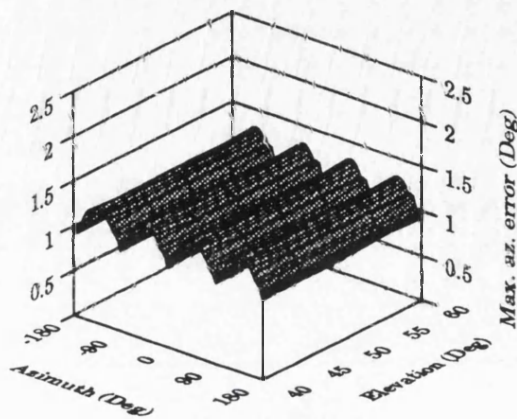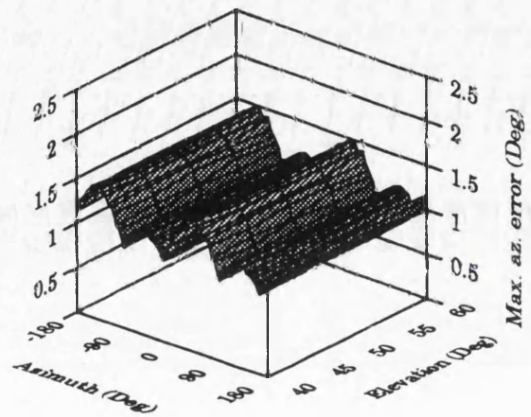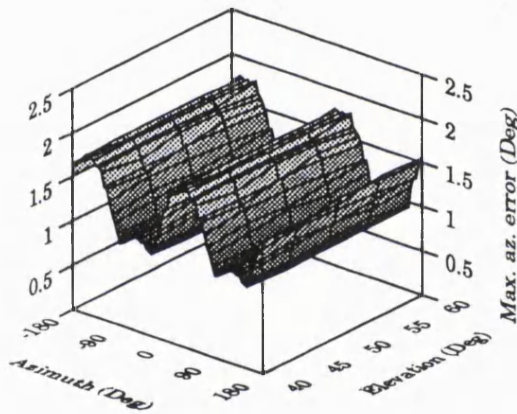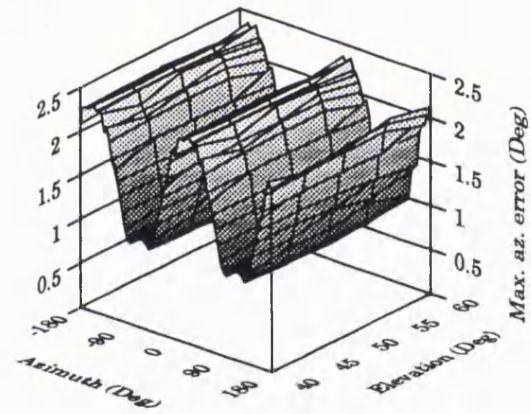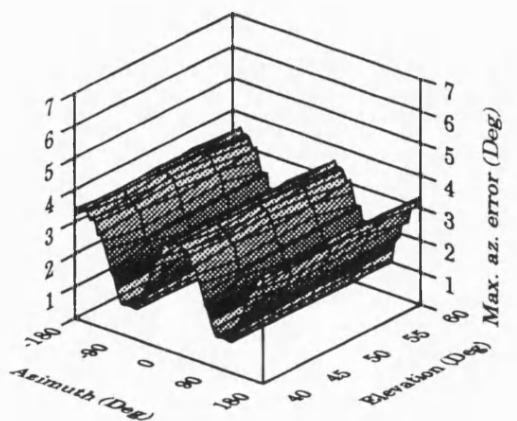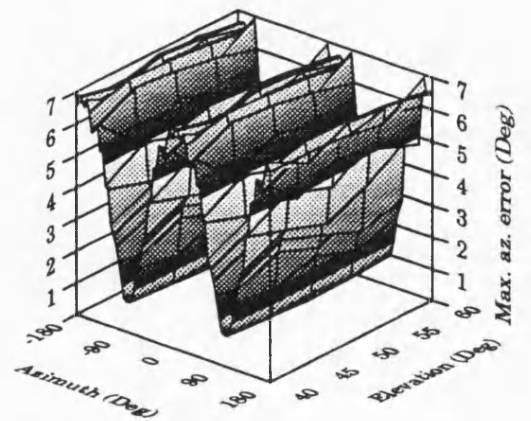
(f) γ = 15°

Figure 3-15:   Maximum azimuth error for different angles γ between the
element axes (both axes in the azimuth plane)
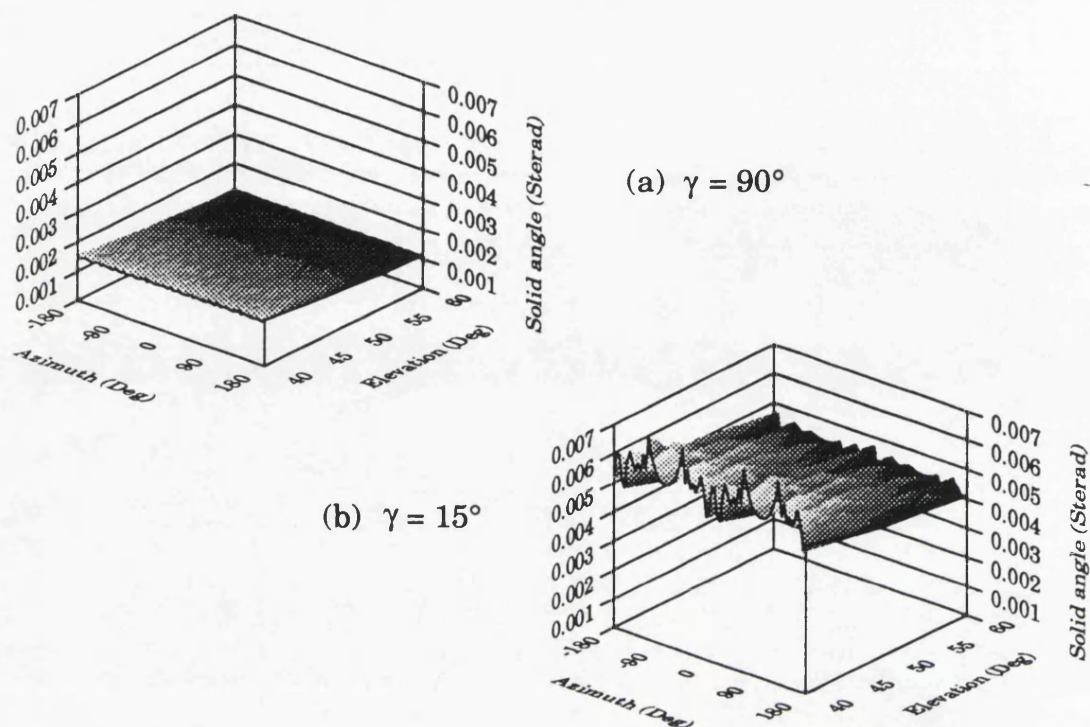
Figure 3-16: Solid angle covered by the intersection for different angles γ between the element axes (both axes in the azimuth plane)
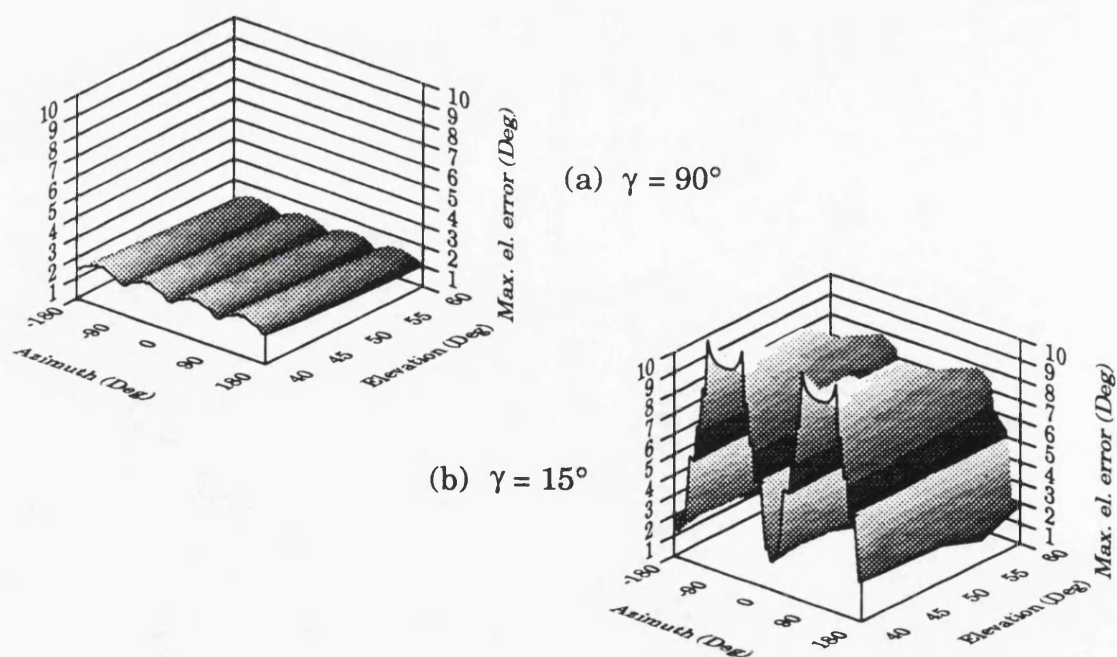
Figure 3-17: Maximum elevation error for different angles γ between the element axes (both axes in the azimuth plane)

3.6. EFFECTS OF MUTUAL COUPLING, ELEMENT OBSCURATION AND HARDWARE IMPERFECTIONS ON THE DF ACCURACY

The influence of system inherent errors on the DF accuracy of an ideal symmetric-pair direction finder was investigated in the previous section. However, additional problems caused by hardware imperfections, mutual coupling between antenna elements and obscuration by neighbouring elements are present in a real system. This section tries to assess the impact of these systematic errors on the DF accuracy of the system.

Scattering parameters ['S-parameters'] are used in the error model, because the necessary data can conveniently be obtained in this format [Gri 90]. Each of the error sources is represented by a separate matrix as shown in Fig. 3-18.
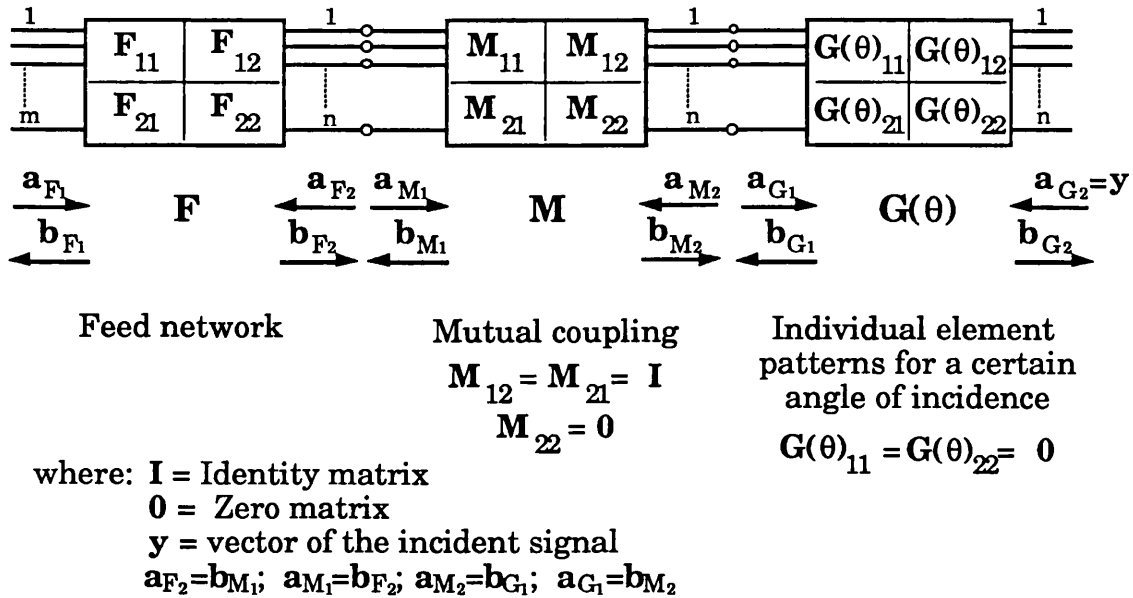
Figure 3-18: Matrix representation of the systematic errors of a symmetric-pair direction finder

Matrix F is a $(m+n)x(m+n)$ matrix, where m is the number of outputs of the feed network and n is the number of antenna elements. This matrix represents the imperfect feed network. The effects of mutual coupling are taken into account in matrix M, which is of size $(2n)x(2n)$. It should be noted that the mutual coupling coefficients are solely dependent on the

array geometry and are independent of the angle of incidence of any signal. The actual element response of each individual active element in the array (which includes obscuration by neighbouring radiators) is represented by the matrix $G(\theta)$, which is also of the size $(2n) \times (2n)$. The values of this matrix depend on the angle of incidence θ. The vectors a_x and b_x represent the waves travelling to and from a certain set of ports. The incident signal is given by the vector $y = [1 \quad e^{\psi_1(\theta)} \quad e^{\psi_2(\theta)} \ldots \quad e^{\psi_{n-1}(\theta)}]^t$ where the first term represents the normalised phase and amplitude at the reference element.

3.6.1. The structure of the matrices F, M and G(θ)

The form of the matrix F obviously depends on the method chosen to process the signals of a pair. In this analysis, parallel Σ and Δ processing of the signals of each pair is assumed. The structure of the relevant scattering matrix of a 180° hybrid is equal to:

$$
H = \left(
\begin{array}{cc|cc}
h_{11} & h_{12} & h_{13} & h_{13} \\
h_{12} & h_{22} & h_{23} & -h_{23} \\
\hline
h_{13} & h_{23} & h_{33} & h_{34} \\
h_{13} & -h_{23} & h_{34} & h_{44}
\end{array}
\right)
$$

with:
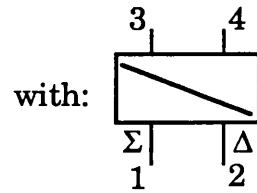
Exploiting the reciprocity of the hybrid, the following parameters can be obtained:

h_{xx} = reflection coefficients

h_{13} = transmission coefficient between port 1 and port 3 or 4

h_{23} = transmission coefficient between port 2 and port 3 or 4

h_{12} = Isolation between port 1 and 2

h_{34} = Isolation between port 3 and 4

The general structure of the matrix of the feed network for Σ and Δ processing is therefore:

$$
\mathbf{F} = \left(
\begin{array}{c|c}
\begin{array}{c}
r_{11} \\
\begin{array}{cc} h^l_{11} & h^l_{12} \\ h^l_{12} & h^l_{22} \end{array} \qquad 0 \\
\qquad 0 \qquad \ddots \\
\qquad \begin{array}{cc} h^p_{11} & h^p_{12} \\ h^p_{12} & h^p_{22} \end{array}
\end{array}
&
\begin{array}{c}
r_{12} \\
\begin{array}{cc} h^l_{13} & h^l_{13} \\ h^l_{23} & -h^l_{23} \end{array} \qquad 0 \\
\qquad 0 \qquad \ddots \\
\qquad \begin{array}{cc} h^p_{13} & h^p_{13} \\ h^p_{23} & -h^p_{23} \end{array}
\end{array}
\\ \hline
\begin{array}{c}
r_{21} \\
\begin{array}{cc} h^l_{13} & h^l_{23} \\ h^l_{13} & -h^l_{23} \end{array} \qquad 0 \\
\qquad 0 \qquad \ddots \\
\qquad \begin{array}{cc} h^p_{13} & h^p_{23} \\ h^p_{13} & -h^p_{23} \end{array}
\end{array}
&
\begin{array}{c}
r_{22} \\
\begin{array}{cc} h^l_{33} & h^l_{34} \\ h^l_{34} & h^l_{44} \end{array} \qquad 0 \\
\qquad 0 \qquad \ddots \\
\qquad \begin{array}{cc} h^p_{33} & h^p_{34} \\ h^p_{34} & h^p_{44} \end{array}
\end{array}
\end{array}
\right)
$$

...(3-16)

where r_{xy} are the elements of the scattering matrix **R** which takes the imperfections of the reference signal path through the feed network into account, and $h_{xy}{}^k$ are the scattering parameters of the kth hybrid.

Mutual coupling will mainly affect the pairs with small spacings. However, the symmetry of the structure will decrease the coupling effects, as it was shown for an interferometer DF system with different arrangements of the elements by Zhu [Zhu 90]. It is therefore sufficient just to include the two inner pairs of each axis in the calculations, and to assume that the coupling effects with the elements of all other pairs are negligible. As an example, the mutual coupling matrix for a 2D arrangement will be investigated, and a similar matrix can be obtained for the 3D case. If the pattern of each element is omnidirectional in the plane defined by the element axes and the array consists of passive radiators, the following assumptions for the coupling coefficients can be made due to the symmetry of the arrangement (see Fig. 3-19):

$$m_{11}$$

$$m_{22} = m_{33} = m_{44} = m_{55}$$

$$m_{66} = m_{77} = m_{88} = m_{99}$$

$$m_{1010} = m_{1111} = \ldots \quad = m_{nn}$$

$$m_{12} = m_{21} = m_{13} = m_{31} = m_{14} = m_{41} = m_{15} = m_{51}$$

$$m_{24} = m_{42} = m_{43} = m_{34} = m_{35} = m_{53} = m_{52} = m_{25}$$

$$m_{23} = m_{32} = m_{45} = m_{54}$$

$$m_{16} = m_{61} = m_{17} = m_{71} = m_{18} = m_{81} = m_{19} = m_{91}$$

$$m_{68} = m_{86} = m_{87} = m_{78} = m_{79} = m_{97} = m_{96} = m_{69}$$

$$m_{67} = m_{76} = m_{89} = m_{98}$$

$$m_{26} = m_{62} = m_{48} = m_{84} = m_{37} = m_{73} = m_{59} = m_{95}$$

$$m_{29} = m_{92} = m_{28} = m_{82} = m_{46} = m_{64} = m_{47} = m_{74} =$$

$$m_{38} = m_{83} = m_{39} = m_{93} = m_{56} = m_{65} = m_{57} = m_{75}$$

$$m_{27} = m_{72} = m_{49} = m_{94} = m_{36} = m_{63} = m_{58} = m_{85}$$

These coefficients are measured with *all other elements* of the array terminated in matched loads.
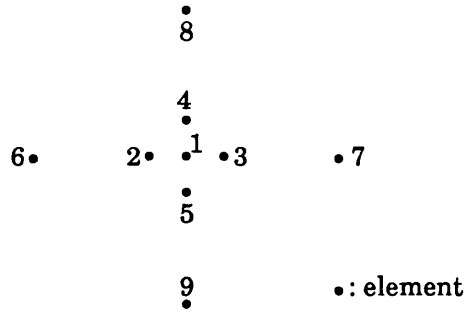
Figure 3-19: *Arrangement of the central pairs for mutual coupling*
           calculations

For an n element passive array, the sub matrix M_{11} of the coupling matrix **M** has then the following structure:

...(3.17)

M_{12} and M_{21} are identity matrices and M_{22} is a zero matrix. Dependent on the operating bandwidth of the DF system and the type of antenna elements, it may even be possible to consider only the coupling between the two inner pairs and the reference element, which would simplify the sub matrix M_{11}.

The reflection sub matrices $G(\theta)_{11}$ and $G(\theta)_{22}$ of $G(\theta)$ are both zero matrices and the individual element patterns for a certain angle θ are taken into account by the two transmission sub matrices $G(\theta)_{12}$ and $G(\theta)_{21}$. The elements of these matrices are zero apart from the main diagonal elements which represent the individual element responses for a certain angle. For passive elements, $G(\theta)_{12}$ equals $G(\theta)_{21}$.

For the sake of clarity, two matrices were used to describe the effects of mutual coupling and differing element responses. However, due to the simplicity of some of the sub matrices, these two matrices can be combined in matrix C, which now describes all imperfections of the antenna system:

$$C = \left(\begin{array}{c|c} M_{11} & G(\theta)_{12} \\ \hline G(\theta)_{21} & G(\theta)_{22} \end{array} \right)$$

...(3.18)

3.6.2. Estimation of the DF error due to system imperfections

To get an idea of how much the accuracy of the DF system is affected by systematic errors a worst case calculation was done. The system considered has two orthogonal axes, and only the coupling between the inner pairs of each axis and the reference element is taken into account. The elements were assumed to be vertical dipoles. Since it is not known how much the pattern of each element suffers from obscuration effects, ideal omnidirectional antenna responses are assumed for these calculations.

The values for the mutual impedance of two parallel dipoles as given by Kraus are used as coupling coefficients between the elements [Kra 88]. These numbers represent a reasonable approximation for the worst case calculation. Furthermore, it is assumed that the dipoles are reasonably

matched, so that their self impedance (terminal impedance for the isolated radiator) is close to 50Ω. The values for the hybrid parameters are taken from data sheets and are in this example the worst case figures for an X-band hybrid.

The evaluation of the errors due to mutual coupling and an imperfect feed network is shown in appendix 4. The calculations give the maximum value by which the phase of the signal of each antenna element could be changed due to systematic errors. Considering the signals of one pair, this phase value can be split into two components. A certain portion of the phase offset is equal and in the same direction for both elements, which is caused by the coupling between the elements of the pair and all other array elements. This effect causes an offset of the Σ and Δ processed signals relative to the reference phase. It does not affect the accuracy of the DF system as long as this offset is smaller than $\pm 90°$ (see chapter 2.5.1.). The other part of the phase offset can act as an opposite phase shift or also as an equally oriented component for the signals of each pair. The former would affect the location of the phase transition angles. However, it cannot be predicted whether the second portion of the phase offset acts entirely as an opposite phase shift or not, because this depends on the phase response of all components. As a worst case for the DF accuracy, it is therefore assumed that the second type of phase offset is purely equal and opposite. The coupling coefficients and the resulting maximum errors for an arrangement with two orthogonal pairs spaced by λ and grouped around a reference element are shown in Table 3-6.

It can be seen that the maximum opposite phase error, which causes a change of the phase transition angles and hence the DF error, is just 2.98° or 4.32° dependent on the reflection coefficient of the elements. Since these values are absolute worst case figures, they are in good accordance with the results quoted by Zhu [Zhu 90]. He calculates a maximum DF error of 2.5° for a square four element dipole interferometer without optimisation and 0.5 λ element spacing. The rather large maximum phase error at the reference element does not affect the bearing accuracy of the system, since it is still possible to decide whether the signals are in phase or anti-phase to each other as long as amplitude errors can be kept low. The maximum amplitude differences of the signals at the element

ports of the hybrid are relatively low, which in turn hardly affects the system accuracy, as previously discussed in sections 2.5.2. and 2.5.3.

	Maximum reflection coefficient at the elements	
	$\mid m_{xx} \mid = 0.1$	$\mid m_{xx} \mid = 0.2$
Worst case coupling coefficients:	$\mid m_{xx} \mid = 0.1$ $\mid m_{12} \mid = 0.143$ $\mid m_{24} \mid = 0.048$ $\mid m_{23} \mid = 0.017$	$\mid m_{xx} \mid = 0.2$ $\mid m_{12} \mid = 0.169$ $\mid m_{24} \mid = 0.056$ $\mid m_{23} \mid = 0.021$
Maximum opposite phase error:	$2.98°$	$4.32°$
Maximum deviation of the phase of the Σ or Δ processed signal from the ideal reference phase:	$3.48°$	$5.95°$
Maximum deviation of the reference phase:	$14.96°$	$19.1°$
Maximum amplitude difference at the element ports of the hybrid:	0.90 dB	1.82 dB

Table 3-6:   Effects of systematic errors on the DF accuracy

It can be seen that the maximum opposite phase error, which causes a change of the phase transition angles and hence the DF error, is just 2.98° or 4.32° dependent on the reflection coefficient of the elements. Since these values are absolute worst case figures, they are in good accordance with the results quoted by Zhu [Zhu 90]. He calculates a maximum DF error of 2.5° for a square four element dipole interferometer without optimisation and 0.5 λ element spacing. The rather large maximum phase error at the reference element does not affect the bearing accuracy of the system, since it is still possible to decide whether the signals are in phase or anti-phase to each other as long as amplitude errors can be kept low. The maximum amplitude differences of the signals at the element

~~ports of the~~ hybrid are relatively low, which in turn hardly affects the system accuracy, as previously discussed in sections ~~2.5.2. and 2.5.3~~.

A symmetric-pair DF system is therefore relatively robust against errors introduced by mutual coupling and an imperfect feed network. This is due to the symmetry of the arrangement and due to the binary character of the output signals of each pair. These qualities decrease the impact of systematic errors at the expense of a relatively large aperture size of the system.

It should be recalled that mutual coupling just has an impact on the phase transition angles of the inner pairs. As a result, just bearings around the relevant phase transition angles suffer from mutual coupling and the accuracy of the system is not affected for all other angles of arrival.

3.7. PERFORMANCE UNDER MULTI-SIGNAL CONDITIONS

The ability of a DF system to cope with multiple co-channel signals depends on its design and the angular separation of the signals. If the signals are less than a beamwidth apart, conventional methods cannot resolve them and superresolution techniques must be used [Hil 90]. Conventional multibeam DF systems can handle multi-signal conditions if the signals are more than a beamwidth apart [Lip 87]. However, the drawback of these systems is that they require a large number of receiver-channels to allow instantaneous coverage of the entire field-of-view. Null steering systems can also be used to cope with more than one incident wave [Cve 88].

A symmetric-pair DF system, is just able to cope with one signal at a time per frequency channel. Hence, it is only possible to determine the angle of arrival simultaneously for a number of signals, when they are distinguishable by their frequency and enough channels are available to process them.

For signals from different directions which cannot be distinguished by their frequency and/or modulation pattern, there is at present virtually

no possibility of taking the bearing of two sources accurately without any *a priori* knowledge of the location of one of the transmitters. If the location of one of the sources is known, its contribution to the output signal of each pair could be calculated and subtracted from it. It would then be possible to determine the angle of arrival of the other signal. However, it will involve substantial computing power in order to achieve this. The effect of additional signals on the same frequency is a change of phase transition angles. If the angle of incidence of the wanted signal is close to a subsector boundary, the shift of the phase transition angles (and hence the subsector boundaries) due to the presence of another signal can yield false bearings.

If the amplitudes of the two signals are very different, the performance of the system is limited by the beamwidth and the grating lobe pattern of each pair. It may well happen that the weaker signal, received through the maximum of a lobe, masks the strong signal, which comes from a direction which is close to the minimum of the grating lobes. Thus it cannot be guaranteed that a symmetric-pair array can separate signals which are more than a beamwidth apart. This, however, is a restriction from which conventional beam-former arrays also suffer. They cannot necessarily separate the signals because the stronger signal received through the sidelobes may mask the weaker signal received through the mainlobe of the combined array response.

3.8. PRACTICAL APPLICATIONS OF SYMMETRIC-PAIR DIRECTION FINDING SYSTEMS

As has been shown in the previous sections, symmetric-pair antenna arrays are capable of performing fully unambiguous direction finding with high accuracy over the entire 3 dimensional space. The configuration required consists of three mutual orthogonal collinear axes.

For a successful application of the symmetric-pair principle in DF it is important to know the polarisation of the expected signals to be able to choose the appropriate type of antenna elements. Since obscuration effects from neighbouring elements should be kept to a minimum and there is only limited space available, elements with large apertures are not suited for the centre pairs, even if the third axis is dislocated.

Symmetric-pair DF arrays with two axes give highest resolution in the direction orthogonal to the two element axes. This property entails two possible applications for this type of DF array:

(1) Mounted under the wings and fuselage of an aircraft, the system could be readily used for surveillance of signals transmitted from the ground.

(2) Using an upright system with two axes, this kind of direction finder could be used as a landing aid at airports. Properly aligned, it gives the best accuracy for the correct direction of approach, as is required for fine corrections to aircraft direction of flight. The cruder accuracy for all other directions is sufficient, since it is just used to guide the aircraft onto the wanted course.

4 DEVELOPMENT OF A SYMMETRIC-PAIR DIRECTION FINDING SYSTEM

4.1. SCOPE OF THE PROBLEMS

The purpose of the previous chapter has been to outline the basic techniques used for direction finding with symmetric-pair arrays. To focus on the key principles, an idealised world with an infinitely strong CW signal had been assumed. In practice, however, the signals under consideration will have a certain bandwidth and only limited strength. To achieve satisfactory performance of the DF system, these properties have to be considered in the design. Hence, the following problems must be solved:

(1) As shown in chapter 3, the phase transition angles of a symmetric-pair are given by the nulls of its grating lobe pattern. As a result, unreasonably high signal-to-noise ratios would be required to determine the sign of a received signal for angles of arrival close to grating lobe nulls.

(2) Due to the frequency dependence of the phase transition angles (see chapter 2, eqn. 2.8), strong broadband signals with angles of arrival close to any phase transition angle would cause permanent change of the polarity of the pair output (flickering). This would make it impossible to determine the correct sign.

In view of the above problems, the following additional requirements for the direction finding system can be specified:

(1) The system should operate safely above an acceptable minimum signal-to-noise ratio, and it should indicate when the signal-to-noise ratio of the received signal is too low.

(2) The system should be able to cope with broadband signals.

91

4.2. INTRODUCTION OF THE ' DON' T KNOW'STATE

The exploration of the binary character of the processed symmetric-pair output is sufficient to determine the angle of arrival, as long as the presence of a strong CW signal can be assumed. However, as outlined above, it is sometimes impossible to attribute one of the two states (in phase or anti-phase) to the output signal of a pair. To solve this problem, the bi-phasal output must be converted into a bipolar baseband signal (see chapter 4.4). A bipolar threshold will then be used at baseband, so that the system only responds in the presence of a signal and measurements are not corrupted by noise. When the output amplitude of a symmetric-pair drops below the predefined bipolar threshold levels (see Fig. 4-1), the polarity of the signal cannot be determined, and the user does not know whether the pair output is in phase or anti-phase with the reference signal. This third possibility, the so called 'don't know' state, covers the areas around grating lobe nulls, i.e. low output amplitudes. The '+' and '-' states in Fig. 4-1 indicate the polarity of strong signals. Since the threshold level is usually fixed, the width of the 'don't know' area depends on the magnitude of the received signal. Weaker signals may not be strong enough to cross the threshold at all, which means that their bearings cannot be taken.
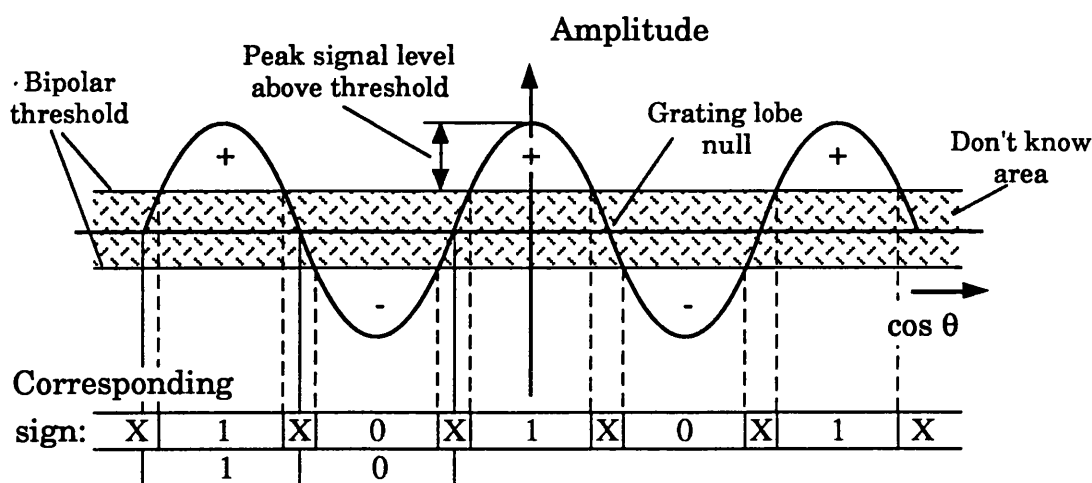
Figure 4-1:   Description of the 'don't know' state

The idea behind this additional 'don't know' state is to permit extraction of the same number of DF bits as under ideal conditions, but to

reduce the minimum required signal-to-noise ratio to acceptable levels for signals incident around phase transition angles. This objective can be achieved by exploring the fact that no definitive answer, i.e. it is not known whether the signal is in phase or anti-phase, can be useful information, since it indicates that the angle of arrival of the received signal is around a grating lobe null. It will later be shown that a more rigorous application of the 'don't know' bit also allows us to deal with broadband signals around phase transition angles.

The 'don't know' state technique can be applied in all cases where two or more bits of information are extracted from a symmetric-pair. As shown in section 3.3, 2^{n-1} measurements are usually required to obtain n bits of DF information from one symmetric-pair. Assuming the bipolar threshold is set high enough, so that it cannot be crossed by noise, a signal is present when at least one of the 2^{n-1} measurements gives a determinate answer ('+' or '-'). Consequently, the minimum number of measurements required for appropriate use of the 'don't know' state is two, which is equivalent to the extraction of two bits of information. If only one bit was extracted from a pair, it would be impossible in the event of a 'don't know' to cross-check for the presence of a signal, since only one measurement is taken.

As previously indicated, the width of the 'don't know' zone depends on the strength of the received signal. The weaker the signal, the wider the zone. In all cases, the 'don't know' zone is centred around a grating lobe null, which is equivalent to a phase transition angle. To make use of the 'don't know' information in calculations, an unambiguous set of possible zone widths has to be defined whereby, dependent on the angle of arrival and the signal strength, always only one of them is applicable at a time. This condition is met if the zone width is a multiple of the subsection width, which in turn is defined by the number of bits extracted from a symmetric-pair (see section 3.3, Fig. 3-5). The smallest 'don't know' zone is therefore one subsection, the widest 2^{n-1} - 1 subsections wide, where n is the number of extracted bits. Fig. 4-2 illustrates this principle for a symmetric-pair with 2 λ spacing, from which 3 bits of DF information are obtained. Fig. 4-2a shows the bit pattern for a infinitely strong signal, i.e. the grating lobe nulls do not pose any problem. In Fig. 4-2b, the 'don't know' zone is one subsection, in Fig. 4-2c two and in Fig. 4-2d three subsections wide. It can

$\Sigma + \Delta$
$\Sigma - \Delta$ **No**
Δ **'don't know'**
Σ

$(S_T = \infty \text{ dB})$

Figure 4-2a

$\Sigma + \Delta$
$\Sigma - \Delta$ **Always one**
Δ **'don't know'**
Σ

$(S_T = 8.34 \text{ dB})$

Figure 4-2b

$\Sigma + \Delta$
$\Sigma - \Delta$ **Always two**
Δ **'don't knows'**
Σ

$(S_T = 3.01 \text{ dB})$

Figure 4-2c

$\Sigma + \Delta$
$\Sigma - \Delta$ **Always three**
Δ **'don't knows'**
Σ

$(S_T = 0.69 \text{ dB})$
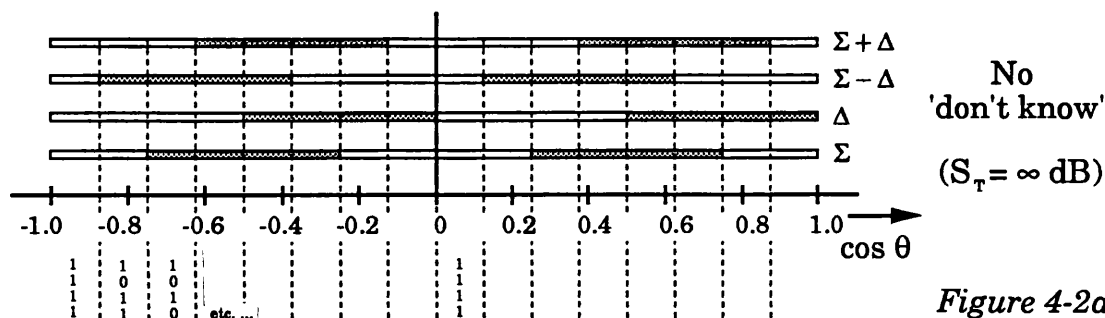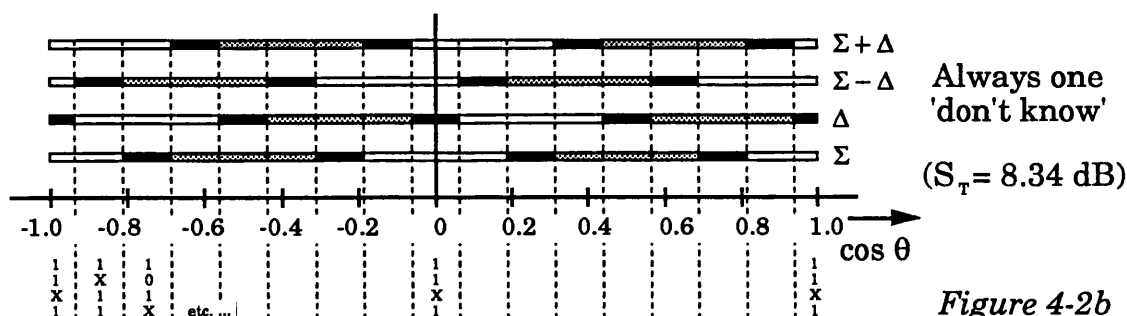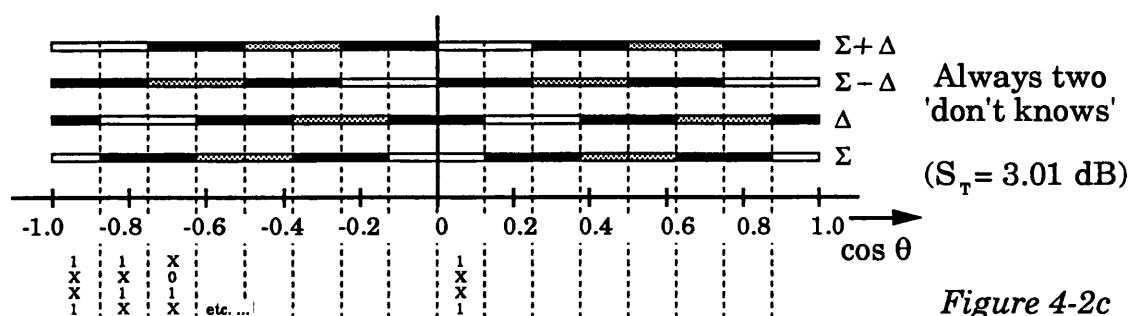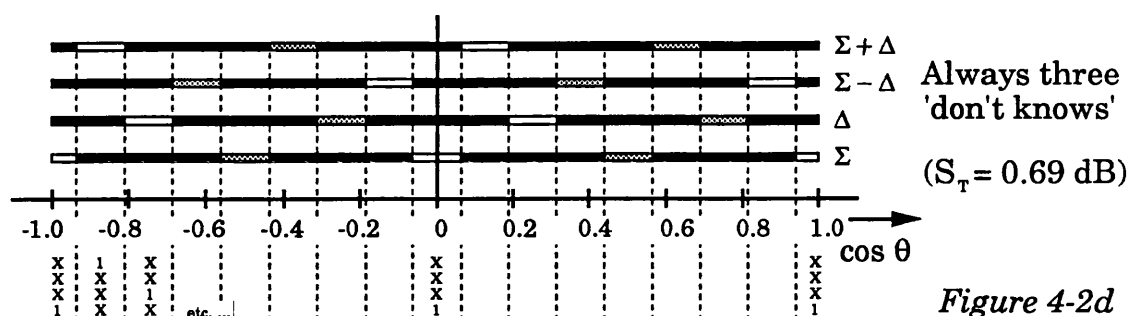
Figure 4-2d

where: ▭ = 1 = Positive sign bit (+)
         ▧ = 0 = Negative sign bit (-)
         ▬ = x = 'Don't know' state

Figure 4-2: 'Don't know' zones for a symmetric-pair with 2λ spacing and n=3

be seen that in cases where the 'don't know' zone is an odd multiple of the subsection width, the boresight subsection ($\cos \theta = 0$) is centred around boresight, for even multiples, a section boundary is at boresight.

The different widths of the 'don't know' zones can be related to the necessary signal strength relative to the threshold. Eqn. 4.1 gives the required peak signal level above threshold for which exactly i 'don't knows' occur during 2^{n-1} measurements taken to extract n bits. (The peak signal level is the maximum output of a symmetric-pair, see Fig. 4-1.)

$$S_{Ti} = -20 \text{Log}\left(\sin\left(i \frac{\pi}{2^n} \right) \right) \text{[dB]}$$

...(4.1)

where: $1 \le i \le 2^{n-1} - 1$

Eqn. 4.1 yields results for signal levels where the actual 'don't know' zone is an exact multiple of a subsection (Fig. 4-2). For intermediate signal levels, the extent of the zones varies between i and i+1 times the subsection width, and either i or i+1 'don't knows' will appear, dependent on the angle of arrival. As an example, Fig. 4-3 shows the true width of the 'don't know' zones for a pair with 2 λ spacing, n = 3 and a peak signal level above threshold of 5 dB. It can clearly be seen that always either one or two 'don't know' states will be recorded. As expected, two adjacent zones have together the width of one subsection.
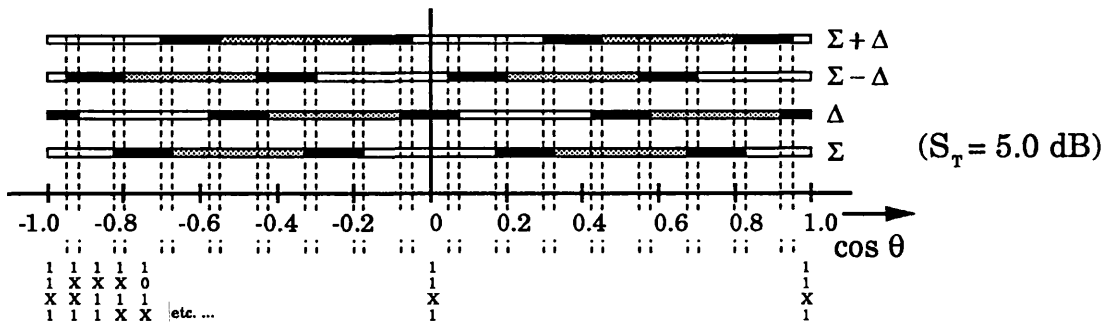
Figure 4-3   True 'don't know' pattern for a pair with 2λ spacing, n=3 and $S_T = 5$ dB

In cases where no 'don't know' events occur, it is possible to obtain the n bits of DF information using simple binary logic, as shown in section 3.3. As

soon as 'don't knows' occur, this is impossible and look-up tables must be used to relate the measured result to the required DF data.

To handle broadband signals with a symmetric-pair DF system, it is useful to recall the principle behind the 'don't know' state. As show above, it is possible to identify the correct subsection as long as at least one out of the 2^{n-1} measurements yields a definitive answer. This very measurement indicates the subsection with the largest amplitude. Hence, it is possible to discard the outputs of all other subsections, which are of smaller amplitude and may be above or below the threshold, since they are not essential for unambiguous identification.

Weak broadband signals that produce at least one 'don't know' event, do not pose any problem to the system, since the pair output, which could cause the flickering of the sign, has by definition the smallest amplitude and is therefore well below the threshold. In the case of strong broadband signals, the polarity of at least the weakest output has to be ignored to give accurate results. To implement it, either flexible instead of fixed thresholds are required or alternatively, methods involving amplitude measurements or amplitude comparisons can be applied.

4.3. UNAMBIGUOUS SPACING OF COLLINEAR ANTENNA PAIRS

To resolve the ambiguities of a symmetric-pair with large spacing, additional pairs with smaller spacings are needed. The pair with the smallest spacing produces unambiguous readings, and the spacing of all other pairs is such that each symmetric-pair resolves the additional ambiguities of the subsequent one. This kind of grouping of symmetric-pairs is called a collinear arrangement, since all antenna elements lie on one axis (see section 3.2). In this section, it will be discussed by how much the predicted ideal spacing between adjacent pairs must be reduced to allow for the effects of imperfections, interference and the newly introduced 'don't know' state.

The introduction of the 'don't know' state affects the spacing of the centre pair, which is ideally 1 λ. As demonstrated in Fig. 4-2, a subsection is centred around boresight in cases where the extent of the 'don't know' zone

is an odd multiple of the subsection. This entails an unwanted ambiguity for the central pair of a collinear arrangement, since both endfire subsections ($\cos\theta = \pm 1$) produce the same code for the ideal spacing of λ. To avoid this ambiguity, the spacing has to be reduced to:

$$d_{central} = \left(1 - \frac{1}{2^n}\right)\lambda$$

...(4.2)

where: n = number of extracted DF bits

As shown in section 3.3, successive symmetric-pairs in a collinear arrangement have spacings that ideally differ by a factor of $1/2^n$. If this factor is only slightly reduced to allow for imperfections, ambiguities occur. Fig. 4-4 illustrates this for the case of n = 2; the corresponding subsections are numbered for simplicity.
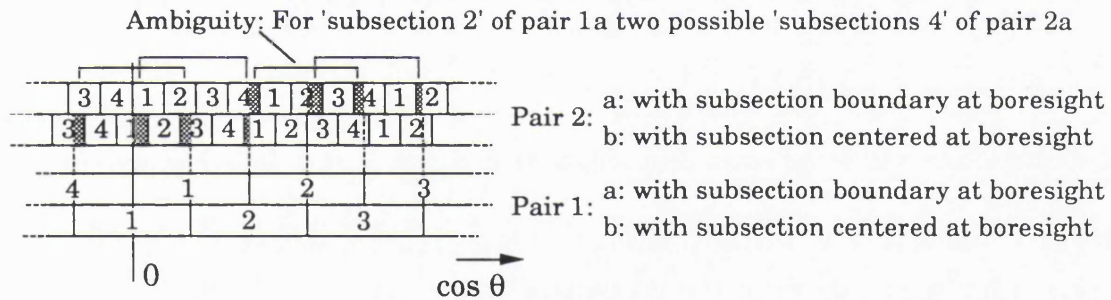
Figure 4-4:   Ambiguities due to inter-pair spacing between $1/(2^n - 1)$ and $1/2^n$

To eliminate these ambiguities, the inter-pair spacing factor must be reduced to $1/(2^n - 1)$. The spacing of the symmetric-pairs of an unambiguous collinear arrangement is therefore:

$$d_x = d_{central}\left(2^n - 1\right)^{x-1} = \frac{\left(2^n - 1\right)^x}{2^n}\lambda$$

...(4.3)

where:

d_x = the spacing of the xth pair

$x \geq 1$

These spacings must be calculated for the highest operating frequency of the array, and there will be no ambiguities at lower frequencies. It may

be necessary to reduce the actual spacings even further by a small amount, to introduce an additional factor of safety. Experiments with a prototype system will show whether this will be required.

Table 4-1 gives the theoretical average subsection widths of an unambiguous symmetric-pair direction finder with quadrant or octant selection and for different spacings. The values have been calculated using eqn. 4.3 and eqn. 3.4. The average subsection width is slightly worse than the average accuracy, since only the narrower cones will be used for accuracy evaluations (see section 3.5).

Pair No.	n=2 (quadrant selection)		n=3 (octant selection)	
	Spacing	Average cone width	Spacing	Average cone width
1	0.75 λ	60.00°	0.875 λ	25.71°
2	2.25 λ	20.00°	6.125 λ	3.67°
3	6.75 λ	6.67°	42.875 λ	0.53°
4	20.25 λ	2.22°		
5	60.75 λ	0.74°		

Table 4-1:   Average subsection width for n = 2, 3 and different spacings

4.4. LAYOUT OF THE DF SYSTEM

There are various possibilities to design a direction finding system which uses the symmetric-pair principle. Dependent on the budget and speed requirements, parallel or sequential processing can be used. The former is more expensive, since identical IF and baseband circuitry has to be built for every antenna pair. Its advantage is that a DF system based on parallel processing is much faster than its counterpart with sequential processing. It also has to be decided how many bits of DF information are to be obtained from each pair. The more bits are extracted, the wider is the spacing of successive pairs, and less symmetric-pairs are consequently needed to achieve a certain accuracy. On the other hand, the number of measurements, and hence the amount of data to be processed, rises with the number of extracted bits in powers of two, as shown in section 3.3.

Symmetric-Pair

Phase reference
element

Switches to access
the pairs sequentially

Downconversion to
a convenient IF

IF amplifiers and filters

Generation of the Σ, Δ,
Σ+Δ, Σ–Δ and reference
signal

Conversion of the bi-phasal
IF signal into a bipolar DC
baseband signal using phase
detectors

Integration, to enhance
performance for low S/N

Bipolar threshold
detector bank

Reading of the bipolar detector outputs,
calculation of the angle of arrival

Figure 4-5   Block diagram of the proposed symmetric-pair direction finder

The chosen configuration is based on the assumption that time is not a
crucial factor for the system performance, and since it is the proposal for a

prototype, hardware expenditure should be kept to a minimum. The suggested DF system uses therefore sequential processing. Three bits of DF information are extracted from each pair, since this seems a reasonable compromise between necessary processing hardware and number of pairs required to achieve an average accuracy of the order of one degree (see Table 4-1 in previous section). Fig. 4-5 shows the block diagram of the proposed DF system. Two multi-port switches select the relevant symmetric-pair, before the pair and reference signals are downconverted to a convenient IF, using a common local oscillator. Before the actual signal processing, the IF signals are amplified and bandpass filtered. The dynamic range of the system is mainly determined by the dynamic range of the three identical AGC IF amplifier stages. The signal processing is performed at IF and baseband, and standard components are used to generate the Σ, Δ, $\Sigma+\Delta$, and $\Sigma-\Delta$ signals. Four phase detectors, which are each fed with one of the four processed pair signals as well as the reference signal, convert the bi-phasal IF signal into a bipolar DC baseband signal. The polarity of the baseband signal indicates, whether the processed output of the symmetric-pair is in phase or anti-phase to the reference signal. The baseband signal is then either low pass filtered, sampled or integrated to enhance the system performance for low signal-to-noise ratios. The merits and disadvantages of these three options will be discussed in the next section. A bipolar threshold detector bank finally generates the three states '+', '-' and 'don't know'. This information is collected by a computer, which calculates the angle of arrival.

The above version of the system generates the three possible baseband states using only hardware. However, it would also be possible to sample the baseband signal at the phase detector output and implement the remaining steps with the computer. This option would be required if the threshold setting method discussed in section 5.3.2 was adopted or if sampling was used to enhance the system sensitivity.

4.5. METHODS TO ENHANCE SYSTEM SENSITIVITY

There are several possibilities to improve the sensitivity of the suggested DF system in order to enhance its performance in low signal-to-noise environments. One option is to make the IF filters as narrow as possible to reduce the noise power in the system. The limitations of this

technique are that the filters must be wide enough to let the received signal pass through, and dependent on the implementation, it might not be easy to build very narrow band filters at IF. Crystal filters may have to be used instead of lumped components filters. Apart from improving the system sensitivity at IF level, it is always possible to enhance it with appropriate techniques at baseband. Three possible means of doing this are discussed in the following subsections.

4.5.1. Baseband low-pass filtering

The signal-to-noise ratio at baseband can be improved by low-pass filtering. Low-pass filters are needed in any version of the system to remove the double frequency components at the phase detector output. If their only purpose is to cut off the unwanted sideband, i.e. their bandwidth is of the order of the IF filter bandwidth or larger, the low-pass filters are called zonal filters. These filters do not affect the signal-to-noise ratio at the phase detector output. The low-pass filters considered here are not zonal filters and their bandwidth is therefore smaller than the IF bandwidth.

The following analysis derives the signal-to-noise ratio at the output of a phase detector with zonal filters before the effect of an additional narrow low-pass filter is considered. The two phase detector inputs consist both of signal and noise components. The signal term is given by:

$$s_x(t) = A_x \cos(\omega_0 t + \phi_x) \qquad \qquad ...(4.4)$$

where x is the index of the phase detector input, A_x the peak amplitude of the signal and ϕ_x a constant phase term. The noise is assumed to be zero mean, narrow band Gaussian noise with a two sided power spectral density of $\eta/2$. Envelope and phase representation has been chosen for the calculation; the amplitude r(t) has a Rayleigh distribution, whereas the phase $\phi_n(t)$ is uniformly distributed:

$$n_x(t) = r_x(t) \cos(\omega_0 t + \phi_{nx}(t)) \qquad \qquad ...(4.5)$$

Both input noises are assumed to be uncorrelated. The output of the phase detector can be written as:

$$s_{out}(t) = (s_1(t) + n_1(t))(s_2(t) + n_2(t)) \qquad ...(4.6)$$

Substituting eqn. (4.4) and eqn (4.5) into eqn. (4.6) and ignoring the higher order terms, the following expression can be obtained:

$$s_{out} = \frac{A_1 A_2}{2} \cos(\phi_1 - \phi_2) + \frac{A_1 r_2(t)}{2} \cos(\phi_1 - \phi_{n2}(t)) +$$

$$\frac{A_2 r_1(t)}{2} \cos(\phi_2 - \phi_{n1}(t)) + \frac{r_1(t) r_2(t)}{2} \cos(\phi_{n1}(t) - \phi_{n2}(t)) \qquad ...(4.7)$$

The normalised power at the phase detector output, i.e. the power that would be dissipated in a 1 Ω resistor [Tau 86], is defined as:

$$S_{out} = \lim_{T \to \infty} \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} [s_{out}]^2 dt \qquad ...(4.8)$$

Since this calculation is purely concerned with power relationships $r_x(t)$ can be replaced by its long term RMS value [Rob 82]:

$$\sqrt{2} \, \overline{r_x(t)} = \sqrt{2} \, \sqrt{2B_{IF} \eta_x/2} = \sqrt{2B_{IF}\eta_x} \qquad ...(4.9)$$

After substitution of eqn. (4.9) into eqn. (4.7), the resulting expression can be applied to eqn. (4.8). After the integration, the normalised output power of the phase detector around DC is equal to:

$$S_{out} = \frac{A_1^2 A_2^2}{4} \cos^2(\phi_1 - \phi_2) + \frac{A_1^2 \eta_2 B_{IF}}{4} + \frac{A_2^2 \eta_1 B_{IF}}{4} + \frac{\eta_1 \eta_2 B_{IF}^2}{2} \qquad ...(4.10)$$

The first term of the above equation represents the signal x signal contribution, the two middle terms the signal x noise cross products and the last term the noise x noise contribution. The output signal-to-noise ratio is therefore:

$$\frac{S}{N}_{out} = \frac{\dfrac{A_1^2 A_2^2}{4} \cos^2(\phi_1 - \phi_2)}{\dfrac{A_1^2 \eta_2 B_{IF}}{4} + \dfrac{A_2^2 \eta_1 B_{IF}}{4} + \dfrac{\eta_1 \eta_2 B_{IF}^2}{2}} = \frac{2 \dfrac{S}{N_{in1}} \dfrac{S}{N_{in2}} \cos^2(\phi_1 - \phi_2)}{\dfrac{S}{N_{in1}} + \dfrac{S}{N_{in2}} + 1} \qquad ...(4.11)$$

with:

$$S_{inx} = A_x^2/2$$
$$N_{inx} = B_{IF} \eta_x$$

The expression for the signal-to-noise ratio at the output of a phase detector is equivalent to the result obtained by Andrews, who derived it using statistical arguments [And 73]. Eqn. (4.11) is very similar to the signal-to-noise ratio at the output of a mixer, the two equations only differ by a factor of $2 \cos^2(\phi_1-\phi_2)$. Since the input signals to the phase detector have the same frequency, the lower sideband after the mixing process is DC, and the contributions from both sides of the double sided spectrum add at zero frequency, which accounts for the factor of 2. Furthermore, the phase difference between the two input signals cannot be neglected because the lower sideband is DC, which yields the factor $\cos^2(\phi_1-\phi_2)$ in eqn. (4.11). There is no loss in signal-to-noise ratio for strong input signals at the phase detector, provided the two signals are in phase or anti-phase; a mixer would have 3 dB loss under the same circumstances.

The relationships between the input and output signals of a phase detector in the spectral domain are given in Fig. 4-6. The output spectral densities are obtained by convolving the relevant input terms. For the sake of completeness, the unwanted upper sideband is also included in the diagram. As can be seen in Fig. 4-6, idealised bandpass filters with a rectangular pass band have been assumed as IF filters.

Knowing the shape of the noise power spectral density at the phase detector output, it is now possible to assess the effect of narrow low-pass filters on the output signal-to-noise ratio. As shown in figure 4-7, it is assumed that the low-pass filter has an idealised rectangular pass band.

Inputs:

signal$_{1/2}$:

$\dfrac{A_{1/2}}{2}$

noise$_{1/2}$:

$\dfrac{\eta_{1/2}}{2}$

B_{IF}

Outputs:

signal$_1$ x signal$_2$:

$\dfrac{A_1^2 A_2^2}{4}\cos^2(\phi_1-\phi_2)$

$\dfrac{A_1^2 A_2^2}{16}$

signal$_1$ x noise$_2$:

$\dfrac{A_1^2 \eta_2}{4}$

$\dfrac{A_1^2 \eta_2}{8}$

noise$_1$ x noise$_2$:

$\dfrac{\eta_1\eta_2}{2} B_{IF}$

$\dfrac{\eta_1\eta_2}{4} B_{IF}$

Total noise spectrum:

$1 = \dfrac{\eta_1\eta_2}{4} B_{IF}$

$2 = \dfrac{\eta_1\eta_2}{4} B_{IF} + \dfrac{A_1^2 \eta_2}{4}$

$3 = \dfrac{\eta_1\eta_2}{4} B_{IF} + \dfrac{A_1^2 \eta_2}{4} + \dfrac{A_2^2 \eta_1}{4}$

$4 = \dfrac{\eta_1\eta_2}{2} B_{IF} + \dfrac{A_1^2 \eta_2}{4} + \dfrac{A_2^2 \eta_1}{4}$

Spectrum the same
as at $-2f_0$
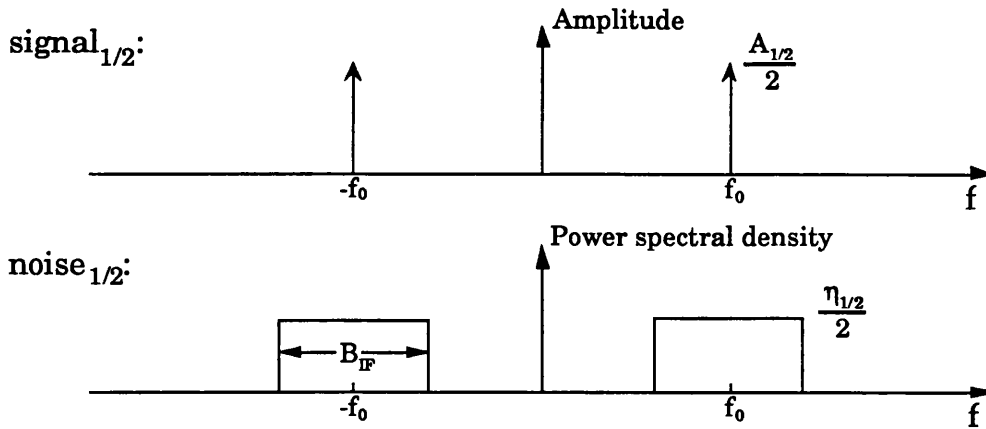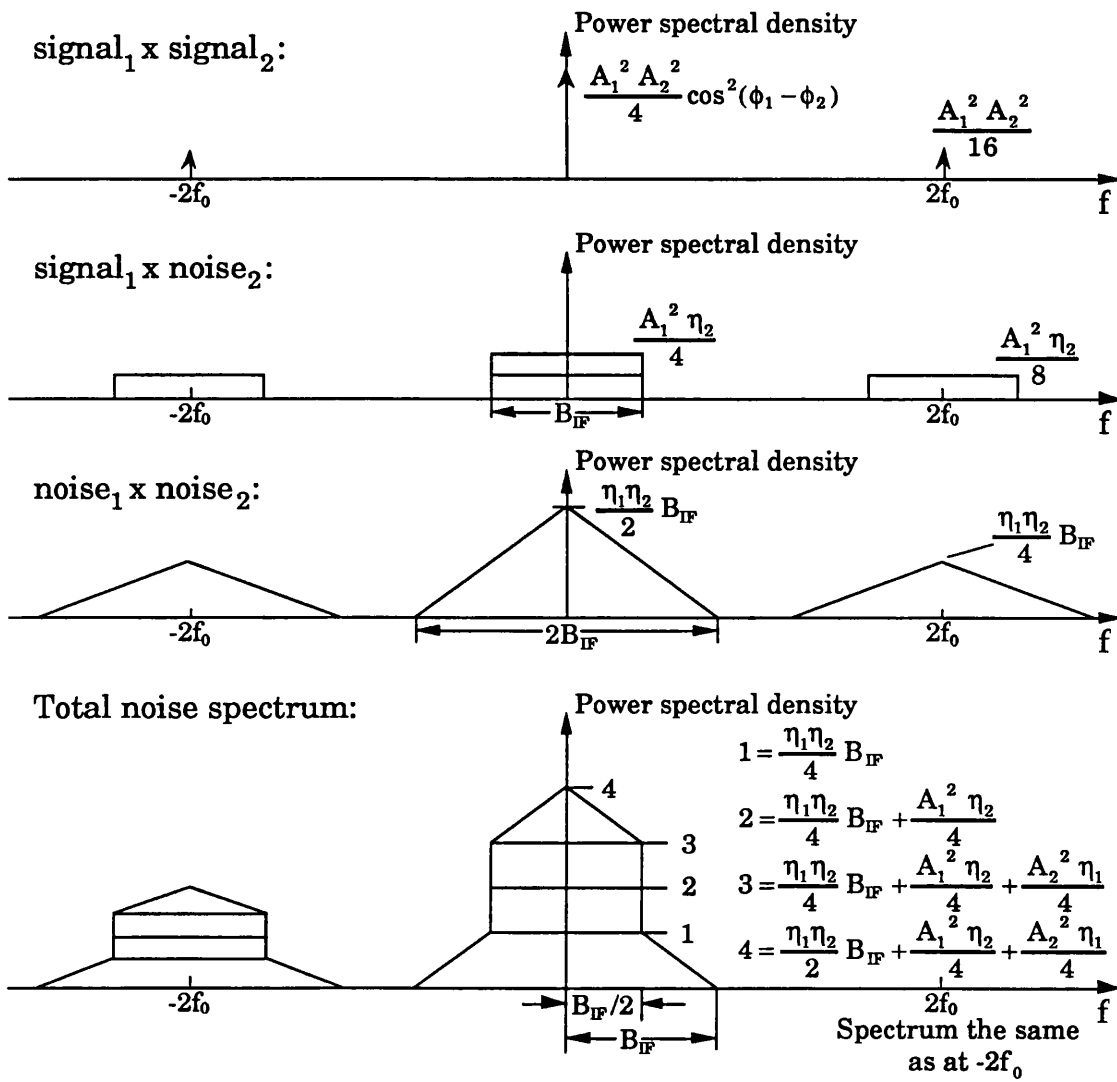
Figure 4-6: Representation of the phase detector signals in the spectral domain

with:

$$y_1 = \frac{\eta_1 \eta_2}{2} B_{IF} + \frac{A_1^2 \eta_2}{4} + \frac{A_2^2 \eta_1}{4}$$

$$y_2 = \frac{\eta_1 \eta_2}{2} B_{IF}$$
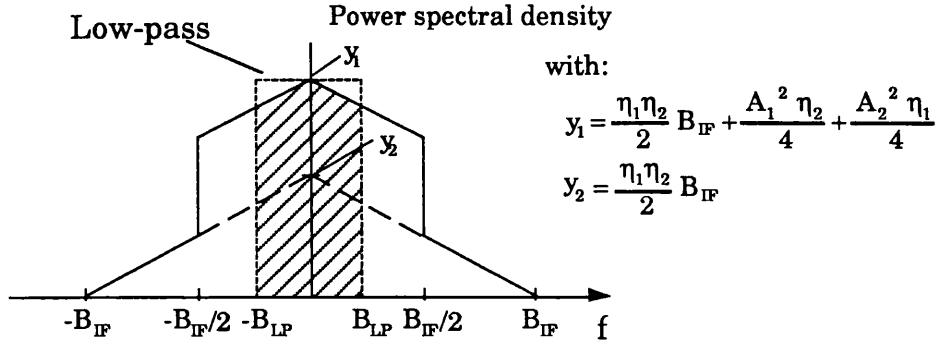
Figure 4-7: Effect of the low-pass filter on the noise power

The noise power at the output of the low-pass filter for $0 \le B_{LP} \le B_{IF}/2$ is given by:

$$N_{LP} = 2 \int_0^{B_{LP}} \left(y_1 - \frac{\eta_1 \eta_2}{2} f \right) df = 2 y_1 B_{LP} - \frac{\eta_1 \eta_2}{2} B_{LP}^2 \qquad ...(4.12)$$

The constants y_1 and y_2 are given in Fig. 4-7. For $B_{IF}/2 < B_{LP} \le B_{IF}$, the noise power at the low-pass filter output can be obtained using the following formula:

$$N_{LP} = 2 \int_0^{B_{IF}/2} \left(y_1 - \frac{\eta_1 \eta_2}{2} f \right) df + 2 \int_{B_{IF}/2}^{B_{LP}} \left(y_2 - \frac{\eta_1 \eta_2}{2} f \right) df$$

$$\qquad ...(4.13)$$

$$= y_1 B_{IF} + y_2 (2 B_{LP} - B_{IF}) - \frac{\eta_1 \eta_2}{2} B_{LP}^2$$

For very narrow low-pass filters, i.e. $B_{LP} \ll B_{IF}/2$, N_{LP} can be approximated by (see Fig. 4-7):

$$N_{LP} \approx y_1 2 B_{LP} \qquad ...(4.14)$$

The gain in signal-to-noise ratio due to low-pass filtering is now given as the ratio of the noise powers at the phase detector and low-pass output:

$$G_{S/N} = 10 \log \left(\frac{N_{Pd}}{N_{LP}} \right) \text{ [dB]} \qquad ...(4.15)$$

The phase detector noise power N_{Pd} can be calculated using eqn. (4.13) with $B_{LP} = B_{IF}$:

$$N_{Pd} = y_1 \, B_{IF} \qquad\qquad ...(4.16)$$

For high signal-to-noise ratios, the low-pass filter must have a bandwidth of less than $B_{IF}/2$ to yield noticeable improvement. This is due to the fact that between $B_{IF}/2$ and B_{IF}, all the noise power is originated by the noise x noise term of the phase detector output, and its contribution is negligible for high signal-to-noise ratios (see Fig. 4-6). Fig. 4-8 shows the signal-to-noise ratio of a phase detector that is followed by a low-pass filter. It can clearly be seen that a low-pass with a bandwidth between B_{IF} and $B_{IF}/2$ can yield an improvement for low input signal-to-noise ratios, whereas for high signal-to-noise ratios, its bandwidth must be less than $B_{IF}/2$.
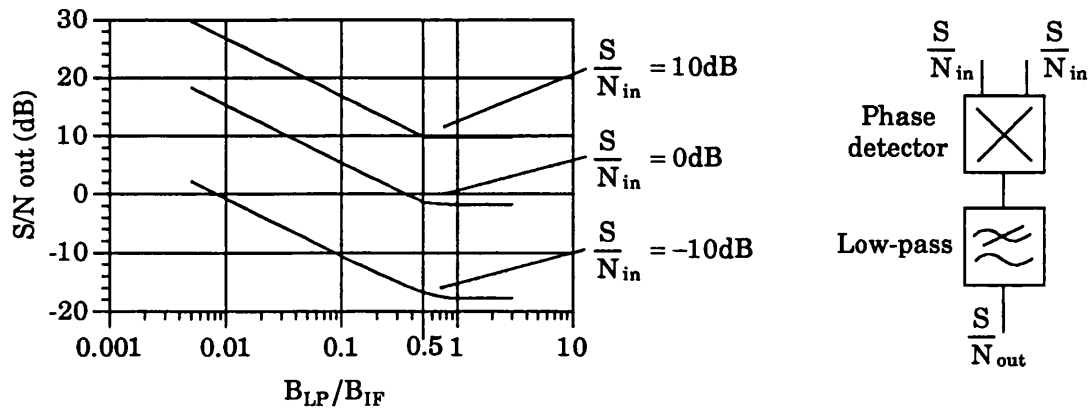
Figure 4-8: S/N improvement due to low-pass filtering

If the low-pass is very narrow, so that eqn. (4.14) holds, the improvement in signal-to-noise ratio becomes:

$$G_{S/N} \approx 10\log\left(\frac{B_{IF}}{2B_{LP}}\right) \ [dB]$$

$$...(4.17)$$

4.5.2. Baseband sampling

Another way of improving the baseband signal-to-noise ratio is to sample and average the phase detector output. Since the signal components

of the samples are fully correlated, the signal amplitudes are added, and the total signal power increases with n^2 for n samples. If the noise contained in successive samples is uncorrelated, the noise powers of the individual samples will be added and the total noise power increases by a factor of n for n samples. The gain in signal-to-noise ratio for n uncorrelated samples is therefore n. Throughout the rest of the text, samples will be referred to as being uncorrelated if the only source of correlation between them is the DC signal component. This definition is useful for the following analysis, which examines the correlation of the noise in successive samples.

The next aim is to determine the sampling rate needed to obtain uncorrelated samples. If the noise power spectral density at the phase detector output was uniform over the entire frequency range of interest, all samples would be independent and hence uncorrelated. However, this is not the case, as it is illustrated in Fig. 4-6. Since sampling is essentially looking at a waveform at different instances in time, the correlation between different samples is described by the autocorrelation function (acf) of the waveform, which is the inverse Fourier transform of the corresponding power spectral density:

$$R(\tau) = \mathcal{F}^{-1}[G(f)] = \int_{-\infty}^{\infty} G(f)\, e^{j2\pi f \tau}\, df$$

...(4.18)

Because the noise power spectral density at the phase detector output is an even function (see Fig. 4-7), it is possible to use the cosine Fourier transform to calculate the corresponding autocorrelation function. Excluding the DC spike at zero frequency caused by the signal component, the acf of the baseband noise becomes:

$$R_N(\tau) = 2\int_{0}^{\infty} G(f)\cos(2\pi f\tau)df$$

$$= 2\int_{0}^{B_{IF}/2} \left(y_1 - \frac{\eta_1 \eta_2}{2} f \right)\cos(2\pi f\tau)df + 2\int_{B_{IF}/2}^{B_{IF}} \left(y_2 - \frac{\eta_1 \eta_2}{2} f \right)\cos(2\pi f\tau)df$$

...(4.19)

where y_1 and y_2 are defined in Fig. 4-7, and the index n denotes that $R_N(\tau)$ is the autocorrelation function of the output noise. After solving the integrals in eqn. (4.19), the following expression for the acf can be obtained:

$$R_N(\tau) = \frac{\sin(\pi B_{IF}\tau)}{\pi B_{IF}\tau}(y_1 - y_2)B_{IF} + \left(\frac{\sin(\pi B_{IF}\tau)}{\pi B_{IF}\tau}\right)^2 y_2 B_{IF}$$

...(4.20)

The autocorrelation function of the phase detector output noise normalised with $R_N(0)$ is shown in Fig. 4-9 for different input signal-to-noise ratios. The abscissa is normalised to $1/B_{IF}$ to make the curves independent of the IF bandwidth.
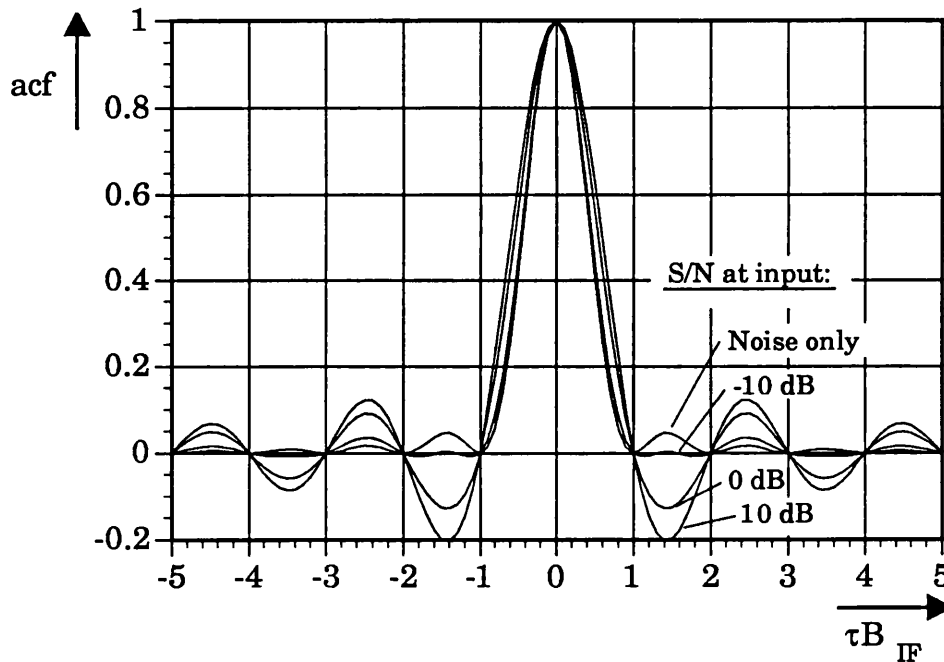
Figure 4-9: Autocorrelationfunction of the noise at the phase detector output

It can be seen that the acf has zeros for integer τB_{IF}, which means that the samples are uncorrelated for a sampling rate of B_{IF} or an integer fraction of it. Sampling at B_{IF} clearly gives the best results, since all other possible sampling frequencies are lower, so that less samples are obtained within a given period of time. Fig. 4-9 also shows that the autocorrelation function has sinx/x shape for high signal-to-noise ratios. This is expected, because the power spectral density has a rectangular form in the case of dominant noise x signal cross products (see Fig. 4-6). Analogous observations can be made for the noise only situation. Here, the triangular

shape of the power spectral density corresponds to the $(\sin x/x)^2$ pattern of the autocorrelation function. The only reason that the zeros of the $\sin x/x$ and $(\sin x/x)^2$ type autocorrelation functions coincide is that the power spectral density function for the triangular noise x noise term is twice as wide as the power spectral density function of the rectangular signal x noise term (see Fig. 4-6).

Having shown that a sampling frequency of B_{IF} yields maximum gain in signal-to-noise ratio for sampling over a limited period of time, it will now be examined how the gain is affected by a change of sampling rate. In order to do this, the output of the phase detector is considered as a stationary random process $z(t)$ with mean m and variance σ^2. That is:

$$E\big[z(t)\big] = m \qquad\qquad\qquad \text{...(4.21a)}$$

$$E\big[(z(t)-m)^2\big] = E\big[z^2(t)\big] - m^2 = \sigma^2 \qquad\qquad \text{...(4.21b)}$$

Costas derived an expression for the second moment, i.e. the total power of the sampled process [Cos 50]:

$$E\big[z^2(t)\big] = \frac{\sigma^2 + m^2}{n} + \frac{1}{n^2}\sum_{k=1}^{n-1} 2(n-k)\, R_z(kt_0) \qquad \text{...(4.22)}$$

where $R_z(t)$ is the autocorrelation function of the random process, n is the number of samples and the time interval between samples is t_0. With eqn. (4.21b), the variance of n correlated samples is equal to:

$$\sigma_n^2 = \frac{\sigma^2}{n} + \frac{1}{n^2}\sum_{k=1}^{n-1} 2(n-k)\, R_z(kt_0) + \frac{m^2(1-n)}{n} \qquad \text{...(4.23)}$$

The mean of the phase detector output is the value of the DC signal, since the probability density function of the noise is symmetrical about the signal term, which will be shown in chapter 5. The signal contribution is a delta function with amplitude m^2 at zero frequency in the power spectral density domain and translates into a constant of value m^2 in the autocorrelation function. The acf of the phase detector output can therefore

be rewritten in terms of m^2 and the autocorrelation function of the noise $R_N(\tau)$:

$$R(\tau) = R_N(\tau) + m^2 \qquad \text{...(4.24)}$$

By exploiting this fact and using the following relation,

$$\sum_{k=1}^{n-1} k = \frac{n(n-1)}{2}$$

the variance for n samples can be expressed in terms of the acf of the noise at the phase detector output:

$$\sigma_n^2 = \frac{\sigma^2}{n} + \frac{1}{n^2} \sum_{k=1}^{n-1} 2(n-k)R_N(kt_0) \qquad \text{...(4.25)}$$

$R_N(\tau)$ is given in eqn. (4.20). An interesting special case of eqn. (4.25) occurs when t_0 is such that each sample is statistically independent of all other samples. $R_N(kt_0)$ is then zero for all k and eqn. (4.25) has the following simplified form:

$$\sigma_n^2 = \frac{\sigma^2}{n} \qquad \text{...(4.26)}$$

This is the well-known expression for the variance of n independent samples.

The variance calculated in eqn. (4.25) is effectively the noise power after summation and averaging of n samples. The change of signal-to-noise ratio gain as a function of the sampling rate is best observed when the sampling takes place over a fixed period of time. The improvement in signal-to-noise ratio, which can also be called the effective number of uncorrelated samples, is obtained by normalising eqn. (4.25) to σ^2, the variance of the process z(t). Since σ^2 is the noise power at the phase detector output, it can be calculated using eqn.(4.13) with $B_{LP} = B_{IF}$:

$$\sigma^2 = y_1 B_{IF} + y_2 B_{IF} - \frac{\eta_1 \eta_2}{2} B_{IF}^2 = y_1 B_{IF} \qquad \text{...(4.27)}$$

Figs. 4-10a-c show the results for different sampling periods and signal-to-noise ratios at the input of the phase detector. The sampling frequency f_s is normalised to B_{IF}, the highest sampling rate that yields uncorrelated samples. It can be seen that the signal-to-noise ratio improvement for sampling rates that do not produce uncorrelated samples depends on signal strength at the phase detector input. For low signal-to-noise ratios, the improvement is proportional to the increase in sampling rate. However, strong signals at the input of the phase detector produce a curve for the signal-to-noise ratio gain that approaches a staircase function for sampling over long time intervals (see Fig. 4-10c). This is because the autocorrelation function of the noise at the phase detector output has a sinx/x shape for strong signals and therefore sometimes becomes negative. Consequently, the sum term in eqn. (4.25), which indicates the magnitude and sign of the overall correlation between the different samples, can also be negative. In the case of the sinx/x autocorrelation function, the sum term and the sampling rate interact such that this staircase function curve is generated.

For sampling rates that are integer fractions of B_{IF}, the samples are uncorrelated and the signal-to-noise gain is, as expected, independent of the signal strength at the phase detector input (see Fig. 4-10). To achieve the best improvement in signal-to-noise ratio, sampling at a rate of B_{IF} is sufficient. Higher sampling frequencies will not give any further improvement, whereas lower rates can immediately yield quite considerable losses in signal-to-noise ratio, especially for weak signals and long sampling times, as illustrated in Fig. 4-10c. With the help of eqn. (4.25) and eqn. (4.27), the gain in signal-to-noise ratio for sampling n samples at an arbitrary rate is given as:

$$G_{S/N} = 10 \log\left(\frac{\sigma^2}{\sigma_n^2}\right) \text{ [dB]}$$

...(4.28)

For the special case where the sampling rate is B_{IF} or an integer fraction thereof, which yields uncorrelated samples, the above equation can be simplified to:

$$G_{S/N} = 10 \log(n) \text{ [dB]}$$

...(4.29)

Figure 4-10a   Total sampling time: $10/B_{IF}$ f_s/B_{IF}

Figure 4-10b   Total sampling time: $100/B_{IF}$ f_s/B_{IF}

Figure 4-10c   Total sampling time: $1000/B_{IF}$ f_s/B_{IF}
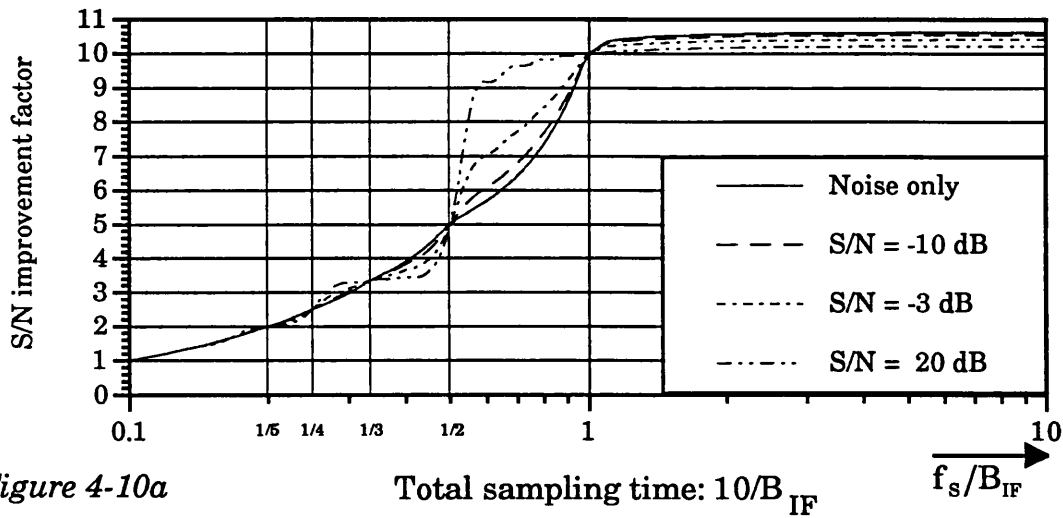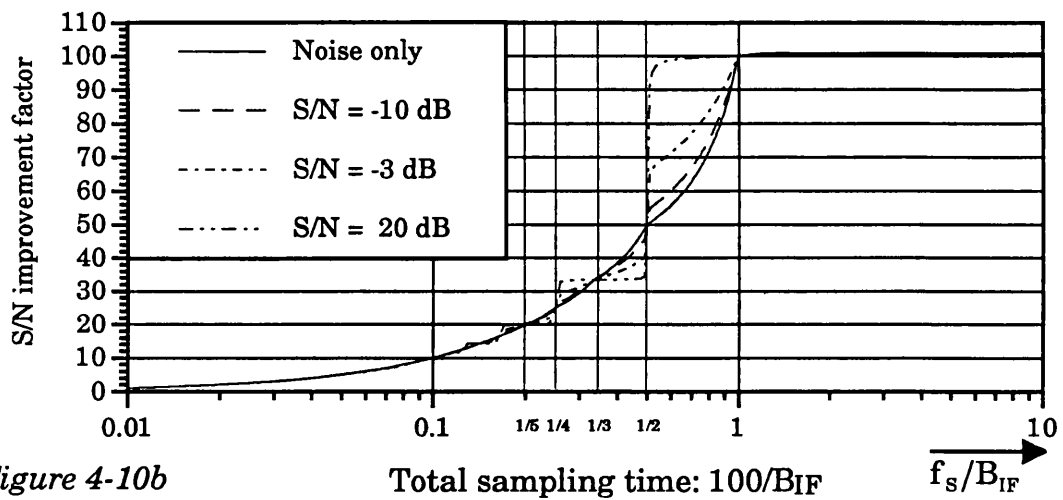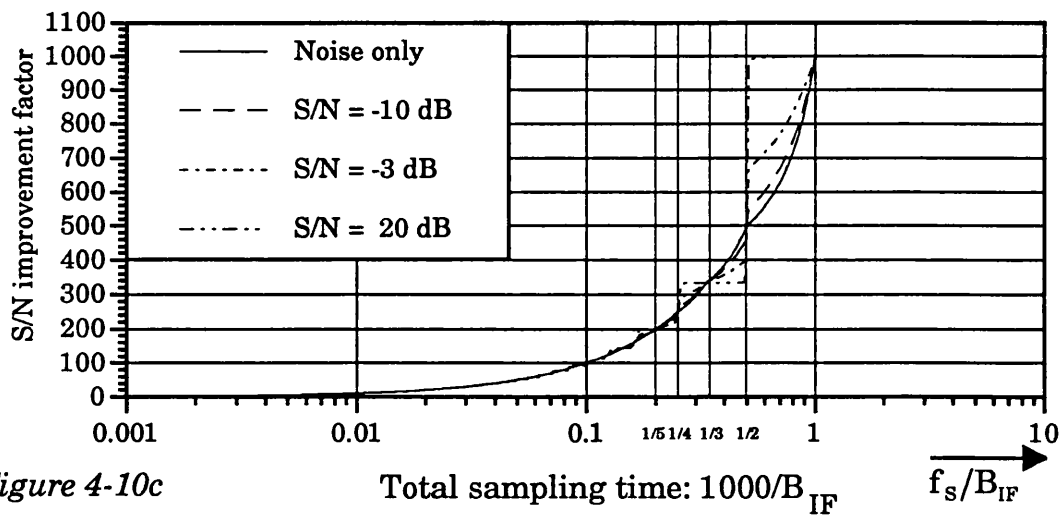
Figure 4-10:   S/N improvement as a function of the sampling rate f_s

4.5.3. Baseband integration

The third option to enhance the system sensitivity at baseband is to perform continuous integration over a limited period of time. This technique is closely related to the previous method, since integration can be seen as an increase of the number of samples beyond limit within a given time interval, i.e. sampling at extremely high sampling rates. The equation obtained by Costas for the variance after integration over a fixed period T is [Cos 50]:

$$\sigma_\infty^2 = \frac{2}{T} \int_{\tau=0}^{\tau=T} \left(1 - \frac{\tau}{T}\right) R_N(\tau) \, d\tau$$

...(4.30)

With $R_N(\tau)$ from eqn. (4.20) and after some manipulations, which can be found in appendix 5, the following expression for σ_∞^2 can be obtained:

$$\sigma_\infty^2 = \frac{2}{T}\left[u(c_1 + 2c_3) + \sum_{n=1}^{\infty} (-1)^n \left[\frac{c_1(u)^{2n+1} + c_3(2u)^{2n+1}}{(2n+1)\,(2n+1)!} + \frac{c_4(2u)^{2n}}{4n\,(2n)!} \right] + \right.$$

$$\left. c_2\big(\cos(u)-1\big) - c_3\left(\frac{\sin^2(u)}{u}\right) \right]$$

...(4.31)

with:

$$u = \pi \, B_{IF} T$$

$$c_1 = \frac{y_1 - y_2}{\pi}; \quad c_2 = \frac{y_1 - y_2}{\pi^2 B_{IF} T}; \quad c_3 = \frac{y_2}{\pi}; \quad c_4 = \frac{y_2}{\pi^2 B_{IF} T}$$

and as defined in Fig. 4-7:

$$y_1 = \frac{\eta_1 \eta_2}{2} B_{IF} + \frac{A_1^2 \eta_2}{4} + \frac{A_2^2 \eta_1}{4}; \quad y_2 = \frac{\eta_1 \eta_2}{2} B_{IF}$$

The results of the calculations of the noise power σ_∞^2 after integration over a period T have been compared with those obtained for the noise power after sampling the output of the phase detector over the same period of time. Sampling at 10 B_{IF} and continuous integration gave the same improvement in signal-to-noise ratio. This confirms the trend shown in Figs. 4-10a-c that no significant increase in signal-to-noise ratio gain can be obtained with sampling frequencies higher than B_{IF}.

4.5.4. Discussion of the three schemes

The method which is probably easiest to implement is low-pass filtering of the baseband signal. The problem is, however, that the filters have a certain settling time and furthermore, an analogue low-pass is an IIR filter (Infinite Impulse Response filter) with theoretically infinite memory. This causes problems when sequential processing is used, where the same IF/baseband unit processes the signals of all pairs in turn. To be certain that the signals of the previous pair do not affect the result any more, long settling times must be allowed between successive measurements. This unnecessarily increases the time required for a complete bearing.

Sampling the baseband signal reduces the analogue hardware of the direction finder. After analogue-to-digital conversion of the output of the sample and hold unit, the processing of the samples can be done digitally. The number of channels required can be reduced to two, since the sampled
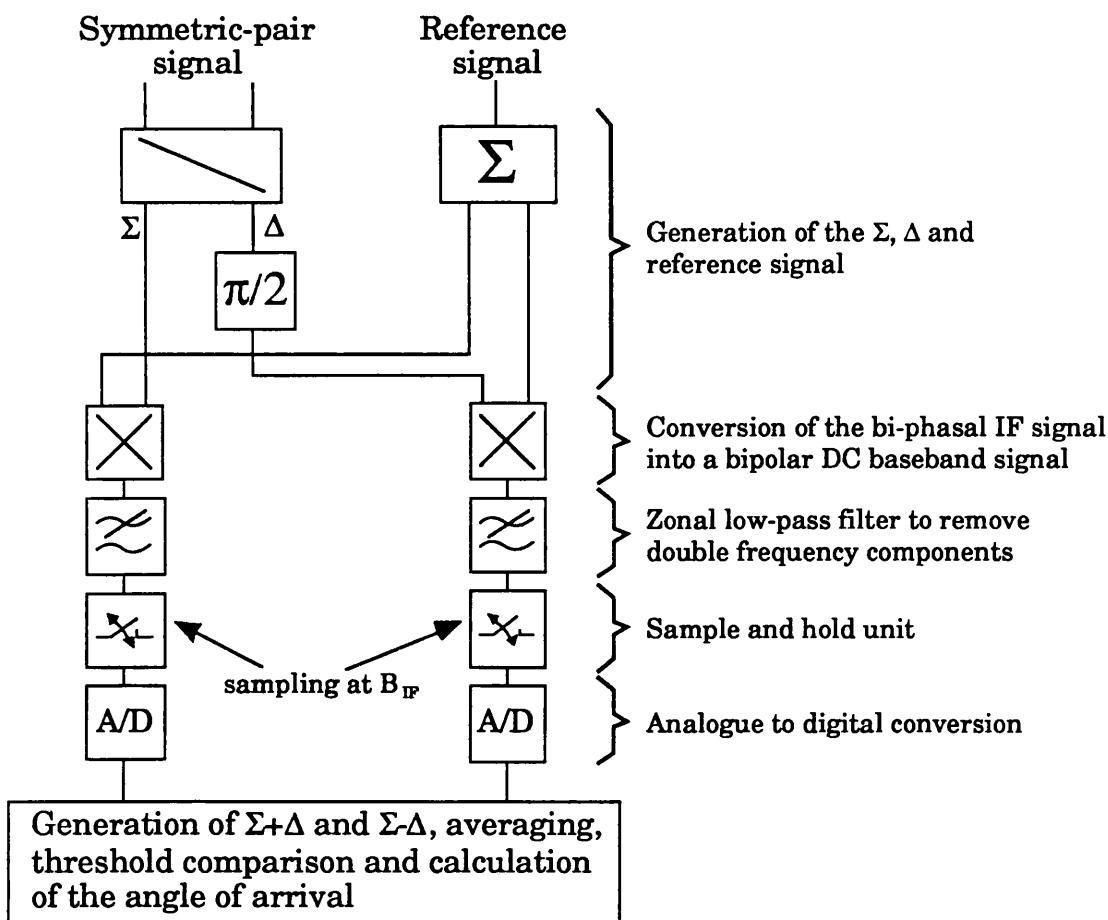
Figure 4-11: IF / baseband block for sampling

and digitised Σ and Δ signals allow us to generate the required $\Sigma{+}\Delta$ and $\Sigma{-}\Delta$ signals without any problem. Fig. 4-11 shows the modified IF/baseband signal processing unit. An advantage of this arrangement is that it is easily possible to change threshold settings and it generally allows more flexible signal processing. Quantisation errors due to the analogue-to-digital conversion should not cause serious problems, as long the quantisation is fine enough and a sufficiently large number of samples is taken for each measurement to average out the errors.

The use of analogue integrators to enhance the signal-to-noise ratio at baseband has the advantage that it is still possible, as in the case of low-pass filters, to have a separate unit that calculates the angle of arrival from the tri-state outputs of the threshold detector bank. By using integrators that can be reset quickly, i.e. 'integrate and dump' principle, the results of successive integrations are independent of each other, and switching between pairs does not cause any problems. The only constraint of this configuration is that the four integrators must show identical performance.

The effectiveness of the different techniques can also be assessed by comparing the time it takes to get a certain signal-to-noise ratio improvement. To obtain for example 10 dB gain, it would be necessary to add 100 samples at B_{IF} (eqn.(4.29)), or to integrate for $100/B_{IF}$ seconds. There is therefore no difference in time between sampling and integration. To achieve the same improvement with a low-pass filter, its corresponding bandwidth had to be $B_{IF}/200$ (eqn.(4.17.)). Certain assumptions concerning the filter and the system must be made, to specify a duration for the low-pass measurement. For the case of a first order low-pass filter, and allowing $4/B_{LP}$ seconds to avoid the problems caused by the memory of the filter (after $4/B_{LP}$, the step response of a first order low-pass filter is more than 98% of its designated value), measurements could be made every $4/B_{LP}$. The time between successive measurements would then be $800/B_{IF}$, which is eight times as much compared to sampling or integration. Assuming a total settling time of only $1/B_{IF}$, it would still take twice the time to complete a measurement with the low-pass approach compared to the other alternatives.

In view of the above discussion, it is best to use the integrate and dump approach, since it allows fast measurements combined with the best possible

signal-to-noise ratio improvement. It also uses only analogue hardware to generate the tri-state signals, which contain the information of the angle of arrival.

4.6. EFFECTS OF AMPLITUDE AND PHASE ERRORS ON THE BASEBAND SIGNAL

Investigations of phase and amplitude errors and their effects on the signals of a symmetric-pair have already been carried out in chapter 2. However, this analysis was done at RF level, to keep the results independent of any particular subsequent processing. The aim was to focus on the location of the phase transition angles as a function of these errors. In this subsection, it is investigated whether the chosen processing technique produces additional side effects, like a decrease of the signal-to-noise ratio, that might affect the performance of the direction finding system.

To assess the impact of the downconversion to baseband it is convenient to derive the signal-to-noise ratio at the output of the phase detector. In the analysis, a single symmetric-pair with Σ processing is considered. The 3 dB loss in signal-to-noise ratio caused by the RF to IF mixing process is ignored, since only signal-to-noise ratio changes due to phase and amplitude errors as a function of the angle of arrival are of interest. The insertion loss of the zonal low-pass filter which removes the double frequency components is also neglected. Fig. 4-12 shows the examined arrangement.
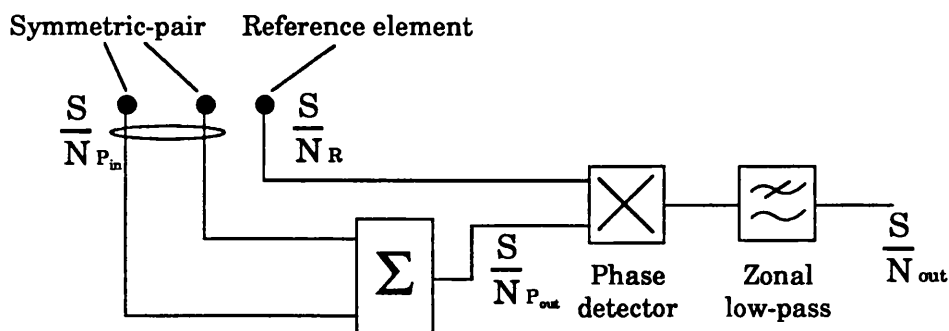
Figure 4-12: Arrangement for phase and amplitude error investigations

The signal-to-noise ratio at the output of a phase detector is given by eqn. (4.11).

$$\frac{S}{N}_{out} = \frac{2\frac{S}{N_{in1}}\frac{S}{N_{in2}}\cos^2(\phi_1 - \phi_2)}{\frac{S}{N_{in1}} + \frac{S}{N_{in2}} + 1} \qquad ...(4.32)$$

The output signal-to-noise ratio of the symmetric-pair S/N_{Pout} can be obtained using eqn. (2.16), which gives the peak output amplitude of a pair with amplitude and symmetrical phase errors (non symmetrical phase errors are not considered here). Hence with Fig. 4-12:

$$\frac{S}{N_{in1}} = \frac{S}{N_R} \qquad ...(4.33a)$$

and

$$\frac{S}{N_{in2}} = \frac{S}{N_{P_{out}}} = \frac{(2V + \Delta V)^2 \cos^2(\phi) + \Delta V^2 \sin^2(\phi)}{4} \qquad ...(4.33b)$$

where:

$$V = \sqrt{2}\ 10^{((S/N)_{Pin}/20)} \qquad \text{and} \qquad \Delta V = V\left(10^{(\Delta(S/N)/20)} - 1\right)$$

To simplify the above calculation of the signal-to-noise ratio at the pair output, a noise power of 0 dBm (=1 mW) at the output has been assumed. $\Delta S/N$ is the difference in signal-to-noise ratio between the elements of the pair in dB. With eqn. (A1.2) and eqn. (2.13) the phase of the error affected pair becomes:

$$\phi_1 = \phi + \frac{2\pi}{\lambda}(\Delta_{al}\cos\theta + \Delta_{or}\sin\theta) - \arctan\left(\frac{V\sin(2\phi)}{\Delta V + 2V\cos^2\phi}\right) \qquad ...(4.33c)$$

where Δ_{al} and Δ_{or} are the offsets along and orthogonal to the element axis as defined in section 2.5.1, θ the angle of incidence and ϕ is given by eqn. (2.1) as:

$$\phi = \frac{\pi d}{\lambda}\cos\theta$$

If equal and opposite phase shifts are used, ϕ in eqn. (4.33c) must be replaced by ϕ':

$$\phi' = \phi + \Delta\phi = \frac{\pi d}{\lambda}\cos\theta + \Delta\phi$$

where $\pm\Delta\phi$ is the phase shift applied at the elements of the symmetric-pair. ϕ_2 in eqn. (4.32) accounts for phase errors in the reference path. The output signal-to-noise ratio of the phase detector for various error constellations can now be calculated with eqn. (4.32) and eqns. (4.33a-c).

Fig. 4-13 illustrates the effects of different errors for a symmetric-pair with 4λ spacing and 0 dB signal-to-noise ratio at all elements. The error free case is shown in Fig. 4-13a, and all nulls in the signal-to-noise ratio pattern are due to phase transitions of the pair output signal. For phase errors in the reference path, the output signal-to-noise ratio is reduced by a factor of $\cos^2(\phi_1 - \phi_2)$ (see eqn. 4.32), as illustrated in Fig. 4-13b. Amplitude imbalances between the elements of the pair as the only source of error also only change the signal-to noise ratio by a constant (Fig. 4-13c), and the same happens for amplitude errors in the reference path or a combination of both types of amplitude errors. However, the location of the phase transition angles is changed for a combination of amplitude errors at the pair and phase errors in the reference channel, and the output signal-to-noise ratio is modulated as a function of the angle of arrival, as shown in Fig. 4-13d.

If the only source of error is an offset of the reference element from the phase centre of the pair, the resulting symmetrical phase errors can be described with Δ_{al} and Δ_{or} (see section 2.5.1), and ϕ_2 is zero. This type of error only modulates the output signal-to-noise ratio as a function of the angle of incidence and does not change the location of the phase transition angles (Figs. 4-13e/f). However, if the maximum phase error is 90° (or an odd multiple of it), the term $\cos^2(\phi_1 - \phi_2)$ in eqn. (4.32) is zero and two extra nulls appear. This is demonstrated in Fig. 4-13f for the case of $\Delta_{al} = 0.25\lambda$, where additional nulls occur at 0° and 180°. Comparisons with the results of section 2.5 show that the phase detector translates the modulation of the output phase of the pair at RF level into a modulation of signal-to-noise ratio at baseband.

(a) $\Delta_{al} = 0\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = 0dB;\ \phi_2 = 0°$

(b) $\Delta_{al} = 0\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = 0dB;\ \phi_2 = 60°$

(c) $\Delta_{al} = 0\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = -10dB;\ \phi_2 = 0°$

(d) $\Delta_{al} = 0\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = -10dB;\ \phi_2 = 60°$

(e) $\Delta_{al} = 0\lambda;\ \Delta_{ar} = 0.2\lambda;\ \Delta S / N = 0dB;\ \phi_2 = 0°$

(f) $\Delta_{al} = 0.25\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = 0dB;\ \phi_2 = 0°$

(g) $\Delta_{al} = 0\lambda;\ \Delta_{ar} = 0.1\lambda;\ \Delta S / N = 0dB;\ \phi_2 = 130°$

(h) $\Delta_{al} = 0.2\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = 0dB;\ \phi_2 = 20°$

(i) $\Delta_{al} = 0.2\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = -10dB;\ \phi_2 = 0°$

(j) $\Delta_{al} = 0.2\lambda;\ \Delta_{ar} = 0\lambda;\ \Delta S / N = -10dB;\ \phi_2 = 60°$
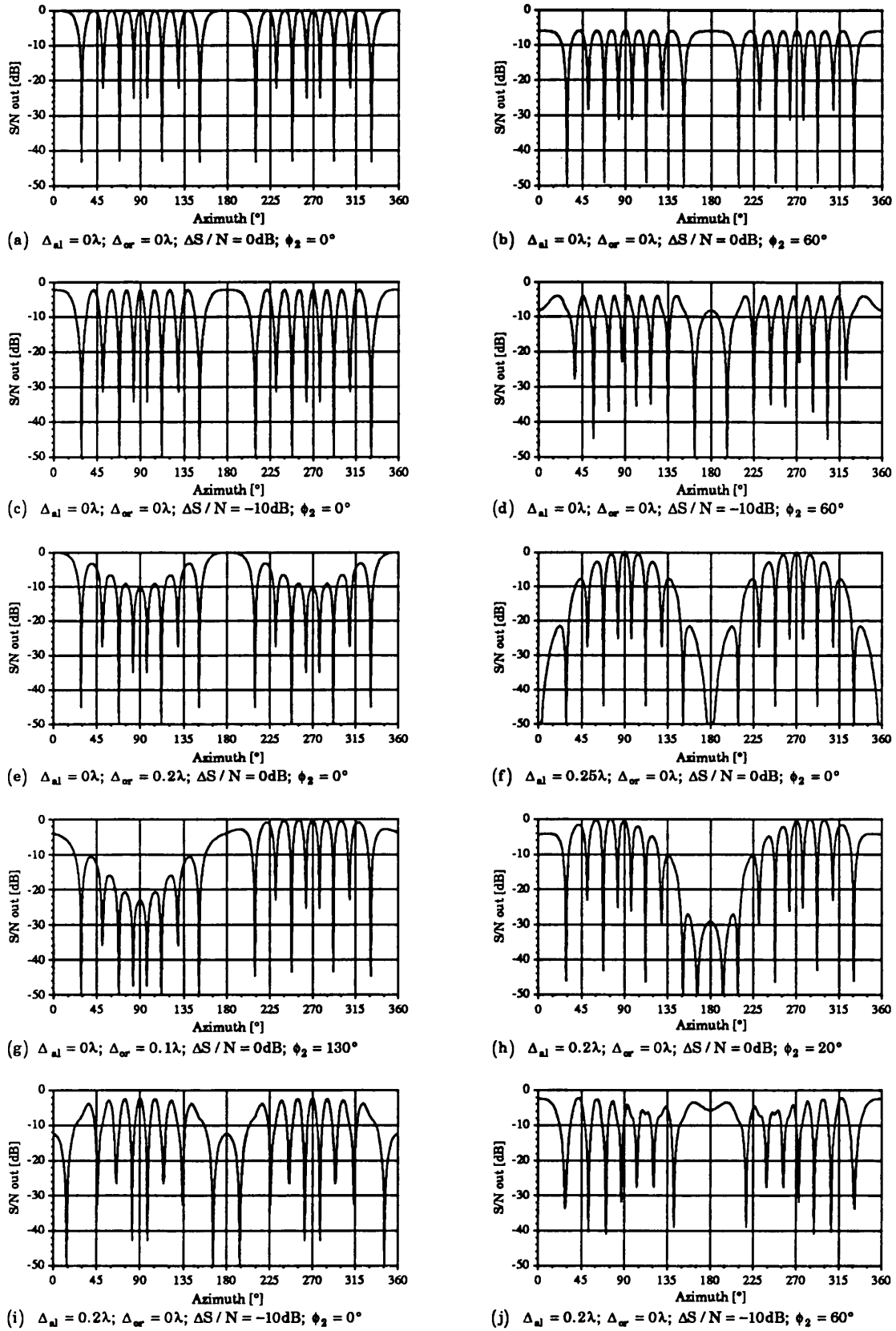
Figure 4-13:   Output signal-to-noise ratio of the phase detector for different errors and 0dB S / N at all elements

Phase errors at the symmetric-pair combined with phase error in the reference channel also introduce extra nulls as soon as the term $\cos^2(\phi_1-\phi_2)$ becomes zero. To avoid this condition, $\phi_1-\phi_2$ must be smaller than $\pm90°$ at all times. Fig. 4-13g shows a somewhat unusual case, where no extra nulls occur despite large errors, because $\phi_1-\phi_2$ is always larger than $\pm90°$. The condition for extra nulls is met by the error constellation of Fig. 4-13h, and two additional nulls can be found around 180°.

Phase *and* amplitude errors at the pair clearly change the location of the phase transition angles (Fig. 4-13h), and additional phase errors in the reference path have the same effect, as shown in Fig. 4-13j.

The above discussion can be summarised by stating that:

(1) The location of phase transition angles does not change as long as only one type of error occurs, i.e. either phase or amplitude error in the reference channel and/or at the symmetric-pair (see Fig. 4-13b, Fig. 4-13c and Fig. 4-13e).

(2) Under certain circumstances ($\cos^2(\phi_1-\phi_2) = 0$), each of the two phase error types, or a combination of them, can introduce additional nulls, but the location of the other nulls does not change (see Fig. 4-13f, Fig. 4-13g and Fig. 4-13h).

(3) A combination of amplitude errors at the pair and any of the two investigated types of phase error changes the phase transition angles (see Fig. 4-13d, Fig. 4-13i and Fig. 4-13j).

(4) All errors change the signal-to-noise ratio of the output, and dependent on the magnitude and type of error, the signal-to-noise ratio may be modulated as a function of the angle of arrival with dips of 10 dB or more. It should be noted here that rather large error parameters have been used for the simulations to highlight the effects of the different types of error.

As long as the errors only affect the signal-to-noise ratio at the phase detector output, the accuracy of the direction finding system will not be affected, only its sensitivity. If the errors are such that the location of the

phase transition angles changes, the direction finder may give incorrect bearings dependent on the angle of incidence.

The envelope of the signal-to-noise ratio at the phase detector output as a function of the angle of arrival remains unchanged for different equal and opposite phase shifts $\Delta\phi$ at the elements of a symmetric-pair, as demonstrated in Fig. 4-14.
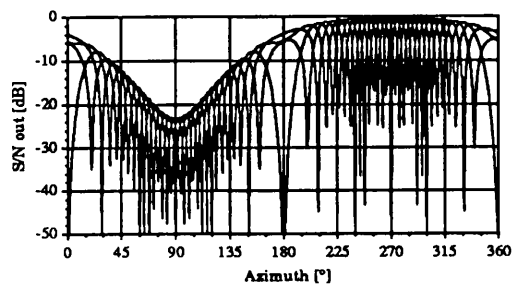
Error parameters:

$\Delta_{al} = 0\lambda$

$\Delta_{or} = 0.1\lambda$

$\Delta S/N = 0$ dB

$\phi_2 = 130°$

Traces for:

$\Delta\phi = 0°$, $\Delta\phi = 90°$, $\Delta\phi = 45°$, $\Delta\phi = -45°$

Figure 4-14: Superimposed outputs of the phase detector for different $\Delta\phi$

5 PREDICTED PERFORMANCE OF THE DF SYSTEM IN LOW SIGNAL-TO-NOISE ENVIRONMENTS

Several aspects of the direction finding system must be studied in order to assess its overall performance. Up to now, the operation of the system under ideal as well as realistic conditions has been analysed, and the impact of phase and amplitude errors on its accuracy and sensitivity has also been evaluated. Furthermore, it has been shown that different techniques can be employed to improve the system sensitivity and the pros and cons of each method have been discussed. The aspect to be scrutinised in this chapter is the performance of the direction finding system in low signal-to-noise environments.

So far, only the long term mean power contribution of the noise has been of interest, which allowed us to represent noise in a non-statistical way. However, to evaluate the system performance for low signal-to-noise ratios, noise has to be treated as a random or stochastic process. The statistics of the noise have to be considered and the results can only be described in probabilistic terms, such as probability of false-alarm, or probability for the detection of a signal.

The probability of detection will be calculated for cases where the phase detector baseband outputs are integrated over a length of time. The relevant analysis can be simplified if successive integrated samples are independent. To find the maximum sampling rate that still gives independent samples, a theorem from detection theory can be used. It states that samples which are more than one correlation period apart may be considered to be independent [Mar 60]. Together with the previously obtained acf of the phase detector output (see section 4.5.2, Fig. 4-9), the above theorem yields a (maximum) sampling rate for independent samples of $1/B_{IF}$.

5.1. CONSIDERATIONS FOR A SINGLE CHANNEL

To be able to analyse the performance of the four channel IF/baseband unit of the symmetric-pair direction finder in low signal-to-noise conditions, it is helpful to investigate the properties of a single isolated channel first. The results obtained can then be used to tackle the more complex four channel case.

A single channel system comprises a symmetric-pair plus the reference element, where only one signal of the pair, for example the Σ signal, is processed using the methods suggested in chapter 4. The phase detector followed by the zonal low-pass is effectively a cross-correlator, and its output is a measure of the correlation between the reference signal and the signal from the symmetric-pair. Any subsequent circuitry, be it a low-pass, integrator or sampling unit, just averages the output signal to reduce the noise contribution and to allow the detection of the signal, which was buried in noise before.

The use of correlation analysis to detect signals in noise was suggested by Lee, Cheatham and Wiesner ([Lee 50a], [Lee 50b]). Fano, Davenport, Horowitz and Green subsequently investigated the output signal-to-noise ratio of different types of correlators ([Fan 51], [Dav 51], [Hor 55], [Gre 57]). Different authors considered then the problem of determining the probability density function (pdf) for the output of an analogue cross-correlator. Lampard, who used low-pass *filtered* inputs to the correlator, showed in his analysis that for a narrow low-pass filter at the correlator output, the pdf tends to become Gaussian, as would be expected from the central limit theorem [Lam 56]. Cooper introduced narrow band noise inputs to the correlator by adding bandpass filters [Coo 65]. Several other authors derived expressions for the probability density function and cumulative distribution function (cdf) of the output of an analogue cross-correlator with various degrees of generality for both low-pass and bandpass inputs ([Kai 66], [Bro 67], [Str 70], [Mur 71], [And 73], [Mil 74], [Bri 82]). Wilson, Andrews and Brice, as well as Milne, give expressions for the sum of n sampled baseband outputs of an analogue cross-correlator ([Wil 79], [And 83], [Mil 91]).

Other types of cross-correlators have also been considered. Roe and White and Andrews discussed the output probability density function of cross-correlators that use sampled input data ([Roe 61], [And 74a]). To complete the overview on correlator literature, it should be mentioned that correlators that clip the incoming waveforms before multiplying them have also been analysed ([Ekr 63], [Che 68], [And 74b], [And 80]). Compared to the other types of correlators, this kind of cross-correlator shows a loss in output signal-to-noise ratio for low input signal-to-noise ratios.

To obtain the false-alarm rate and probability of detection for different threshold settings, it is necessary to know the probability density function and the corresponding cumulative distribution function of the correlator output. Since sampling or integration can be used to enhance the output signal-to-noise ratio (see section 4.5), the pdf and cdf for the summation of n independent samples must be known. As mentioned above, several authors analysed this case. Wilson and Andrews approach the problem by deriving the characteristic function and exploiting the fact that the characteristic function of a sum of n independent samples is that for one sample raised to the power n [Wil 79], [And 83]. Wilson derives expressions under the assumption that the noises at the input of the cross-correlator are uncorrelated, whereas Andrews takes correlation into account. Both of them consider a sinusoid plus narrow band noise as input signals. Milne on the other hand only allows correlated noise at the inputs [Mil 91]. He obtains the pdf of the sum of n independent samples by repeatedly convolving the pdf of a single sample with itself. Since Milne's analysis has not been published, it is given in appendix 6. It is also shown in this appendix that Andrews's results, with a correction factor for the pdf, and Milne's expressions can be transformed into one another.

5.1.1. Probability density function for the phase detector output

For zero mean narrow band Gaussian noise at the inputs, the pdf of the output of an analogue cross-correlator after the summation of n independent samples is given by Milne as follows (see appendix 6, eqn. (A6.10) and eqn. (A6.11)):

$$p_n(z) = \frac{(\beta \alpha)^n}{(\beta + \alpha)^{2n-1}} \exp(-\alpha z) \sum_{k=0}^{n-1} \frac{(2n-2-k)! \; (\beta+\alpha)^k z^k}{k! \; (n-1-k)! \; (n-1)!} \qquad \text{for } z > 0$$

$$...(5.1a)$$

and

$$p_n(z) = \frac{(\beta \alpha)^n}{(\beta + \alpha)^{2n-1}} \exp(\beta z) \sum_{k=0}^{n-1} \frac{(2n-2-k)! \; (\beta+\alpha)^k (-z)^k}{k! \; (n-1-k)! \; (n-1)!} \qquad \text{for } z < 0$$

$$...(5.1b)$$

with

$$\alpha = \frac{2}{\sigma_1 \sigma_2 \left(\sqrt{1 - \lambda_0^2} + \rho_0 \right)}$$

$$...(5.2a)$$

and

$$\beta = \frac{2}{\sigma_1 \sigma_2 \left(\sqrt{1 - \lambda_0^2} - \rho_0 \right)}$$

$$...(5.2b)$$

where σ_1 and σ_2 are the standard deviations of the noise at the cross-correlator inputs, ρ_0 is the in-phase and λ_0 is the quadrature correlation coefficient. The overall correlation coefficient ρ between the inputs of the cross-correlator is defined as:

$$\rho^2 = \rho_0^2 + \lambda_0^2$$

$$...(5.3)$$

To compare the pdf for the output of a cross-correlator with well known probability density functions, it is necessary to compute its mean and variance as a function of the number of added samples.

Since the mean of a sum is always equal to the sum of the means, the expectation of the correlator output after n additions is given by:

$$m_n = E_n[z] = n \, E[z] = n \int_{-\infty}^{\infty} z \, p_1(z) \, dz$$

$$...(5.4)$$

With eqns. (5.1a/b) inserted in eqn. (5.4), $E_n[z]$ results in:

$$E_n[z] = n \left(\int_{-\infty}^{0} z \, \frac{\beta \alpha}{\beta + \alpha} e^{\beta z} \, dz + \int_{0}^{\infty} z \, \frac{\beta \alpha}{\beta + \alpha} e^{-\alpha z} \, dz \right) = n \frac{\beta - \alpha}{\alpha \beta}$$

$$...(5.5)$$

The variance of the cross-correlator output after n summations is defined as:

$$\sigma_n^2 = E_n\left[\left(z - E_n[z]\right)^2\right] = E_n\left[z^2\right] - E_n[z]^2 \qquad \text{...(5.6)}$$

The above equation shows that the second moment of the output of the cross-correlator must be calculated in order to obtain the variance. Substituting eqns. (5.1a/b) into the definition of the second moment, and after integration (which involves incomplete Gamma function type integrals), $E_n[z^2]$ has the following form:

$$E_n\left[z^2\right] = \int_{-\infty}^{\infty} z^2 \, p_n(z) \, dz$$

$$= \frac{(\beta \, \alpha)^n}{(\beta+\alpha)^{2n-1}} \sum_{k=0}^{n-1} \frac{(2n-2-k)!}{(n-1-k)!} \frac{(\beta+\alpha)^k}{(n-1)!} (k+2)(k+1)\left(\frac{1}{\beta^{k+3}} + \frac{1}{\alpha^{k+3}}\right) \qquad \text{...(5.7)}$$

Using the above formulae, it is possible to plot the pdf of the output of an analogue cross-correlator for different correlation coefficients and numbers of independent samples, as shown in Figs. 5-1a-i. Unit variances have been assumed at the input of the correlator for all calculations. The corresponding Gaussian distribution with the same mean and variance as the correlator pdf is also plotted in Fig. 5-1. This illustrates how the pdfs become more and more Gaussian as the number of samples increases, which is expected from central limit theorem arguments, The corresponding Gaussian pdf is defined as:

$$p(z) = \frac{1}{\sqrt{2 \, \pi \, \sigma_n^2}} \exp\left[\frac{-(z - m_n)^2}{2 \, \sigma_n^2}\right] \qquad \text{...(5.8)}$$

The pdfs for 1, 2 and 10 uncorrelated noise samples are given in Figs. 5-1a-c. As can be seen, the correlator and Gaussian pdf are virtually identical in the central region after summation of 10 independent samples. The pdfs for quadrature correlation λ_0 between the inputs are slimmer compared to the pdfs for uncorrelated noise, but they are still symmetrical, have zero mean and converge similarly fast to the Gaussian shape for an increasing number of samples (Figs. 5-1d-f). It should be noted that the quadrature coefficient λ_0 is zero when the bandpass filters at the cross-correlator inputs are perfectly centred on the carrier
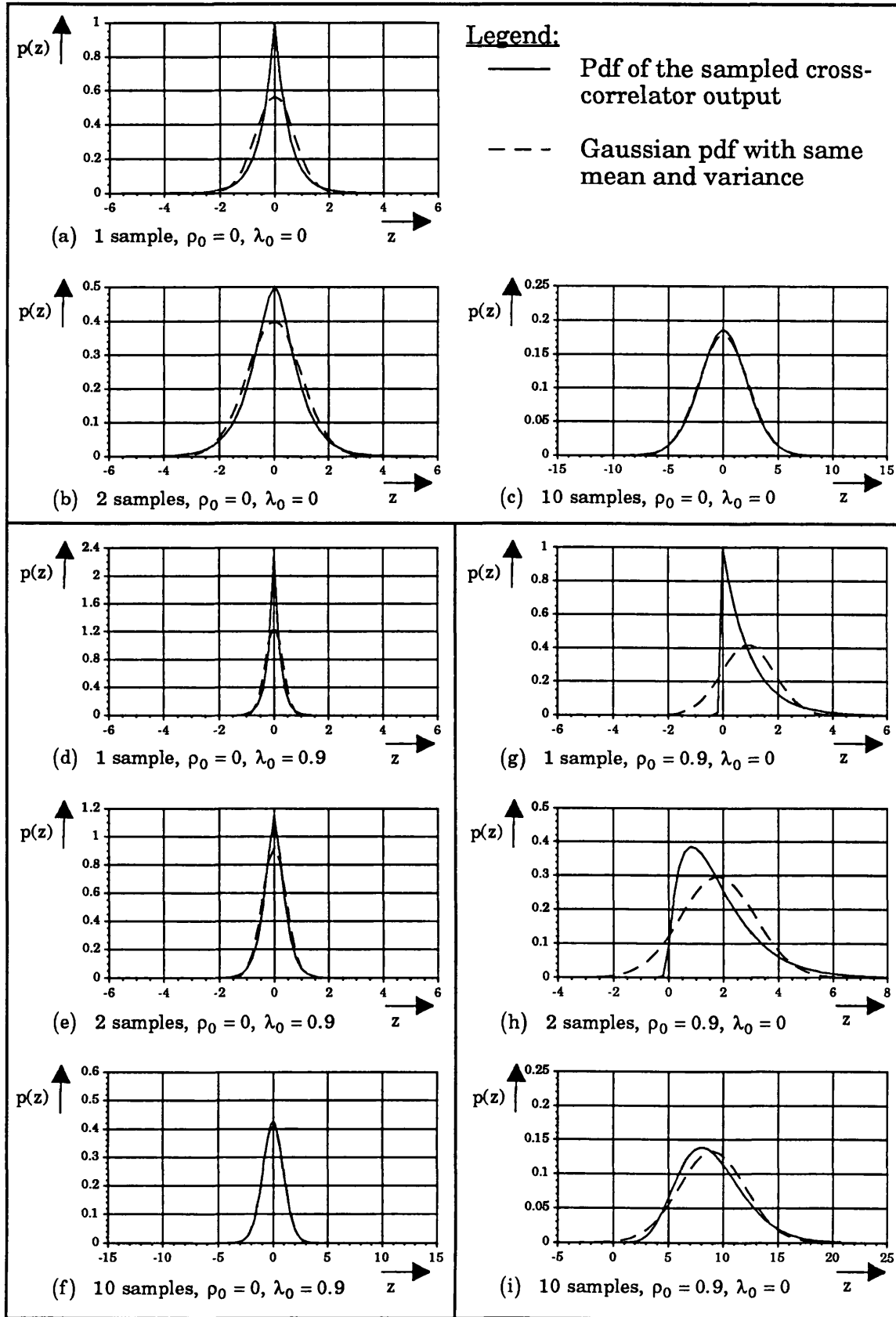
Figure 5-1:   Pdf of the cross-correlator output for different numbers of
samples and correlation coefficients

frequency of the signal ([Ric 45], section 3.7). In-phase correlation destroys the symmetry of the pdf, and for $\rho_0 = 1$ and $\lambda_0 = 0$, the noise at the input would be fully correlated and the cross-correlator would give the same performance as a signal squarer. Figs. 5-1g-i show that despite the highly non-symmetrical pdf for one sample, a Gaussian shape of the distribution will still be obtained for a large number of samples. For high in-phase correlation coefficients, the pdf of the correlator output tends to become single sided positive (Fig. 5-1g). In such cases, the two input signals have virtually always the same sign (and magnitude), which produces the positive output after the correlation inherent multiplication process.

As a summary it can be stated that in-phase correlation makes the pdf non-symmetrical and introduces a mean, whereas quadrature correlation only narrows the pdf compared to the uncorrelated case. If both types of correlation occur, their effects are simply superimposed.

5.1.2. False-alarm rate

It is reasonable to assume for the following calculations that the noise at the cross-correlator inputs is uncorrelated ($\rho = \rho_0 = \lambda_0 = 0$). This implies that the IF bandpass filters of the DF system are centred around the IF frequency to avoid any quadrature correlation, and that no strong external noise sources or jammers are present which might introduce in-phase correlation between the correlator inputs.

Bipolar threshold detectors are employed in the symmetric-pair DF system to decide whether a signal is present or absent in a background of noise. Their performance can be described in terms of two probabilities, the false-alarm probability P_{fa} and the probability of detection P_d. The presence of a signal is assumed if the input to the detector is strong enough to cross the threshold. However, there is a finite probability that the threshold will be exceeded in the absence of a signal. This is because the statistics of random noise are such that there is always a small but non zero probability that the noise voltage can attain a value larger than the threshold level. The probability that the threshold is exceeded when there is no signal present is called the false-alarm probability, which is defined as [Sko 70]:

$$P_{fa} = P(z > T) = \int_T^\infty p_N(z) \, dz$$

$$...(5.9)$$

where T is the threshold level and $p_N(z)$ is the pdf of the noise fed into the detector. Since the correlation between the input noises of the correlator is assumed to be zero, which means that the pdf for the integrated output of the correlator is always symmetrical with zero mean, as shown in the previous section, the false-alarm for the bipolar threshold becomes:

$$P_{fa_n} = P(|z| > T) = 2\int_T^\infty p_n(z) \, dz$$

$$...(5.10)$$

The index n stands for the number of integrated independent samples. Milne gives the probability that the integrated output of a cross-correlator which is fed with Gaussian noise exceeds a certain positive threshold (see appendix 6). After insertion of the series representation of the incomplete Gamma function into eqn. (A6.14b), the false-alarm rate of the bipolar threshold detector for n integrated samples has the following form:

$$P_{fa_n} = 2\left(\frac{\beta\,\alpha}{\beta+\alpha}\right)^n \exp\left(\frac{-2T}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}+\rho_0\right)}\right) \sum_{p=0}^{n-1}\sum_{k=0}^{n-p-1}\left(\frac{\alpha}{\beta+\alpha}\right)^p \frac{(n-1+p)!\ \alpha^k\,T^k}{p!\ (n-1)!\ k!}$$

$$...(5.11)$$

where α and β are given in eqns. (5.2a/b). For uncorrelated noise at the inputs, i.e. $\rho_0 = \lambda_0 = 0$, the above equation can be modified to:

$$P_{fa_n} = 2\left(\frac{1}{\sigma_1\sigma_2}\right)^n \exp\left(\frac{-2T}{\sigma_1\sigma_2}\right) \sum_{p=0}^{n-1}\sum_{k=0}^{n-p-1}\frac{(n-1+p)!\ T^k}{2^{p-k}\ p!\ (n-1)!\ k!\ (\sigma_1\sigma_2)^k}$$

$$...(5.12)$$

The false-alarm probability as a function of the threshold for different numbers of independent samples is given in Fig. 5-2. The mean m_n and the variance σ_n^2, used to normalise the threshold, are defined in eqn. (5.4) and eqn. (5.6) ($m_n = 0$ for uncorrelated noise). Calculations showed that there is no significant difference between the curve for 1000 samples and
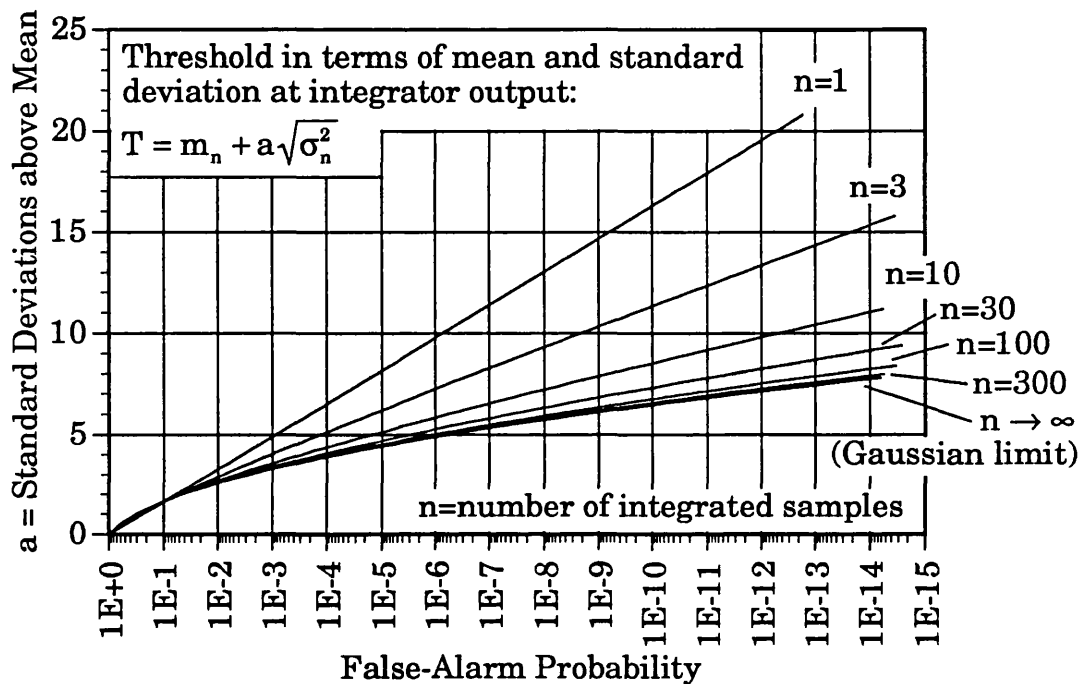
Figure 5-2: False-alarm rate for uncorrelated noise at the cross-correlator inputs
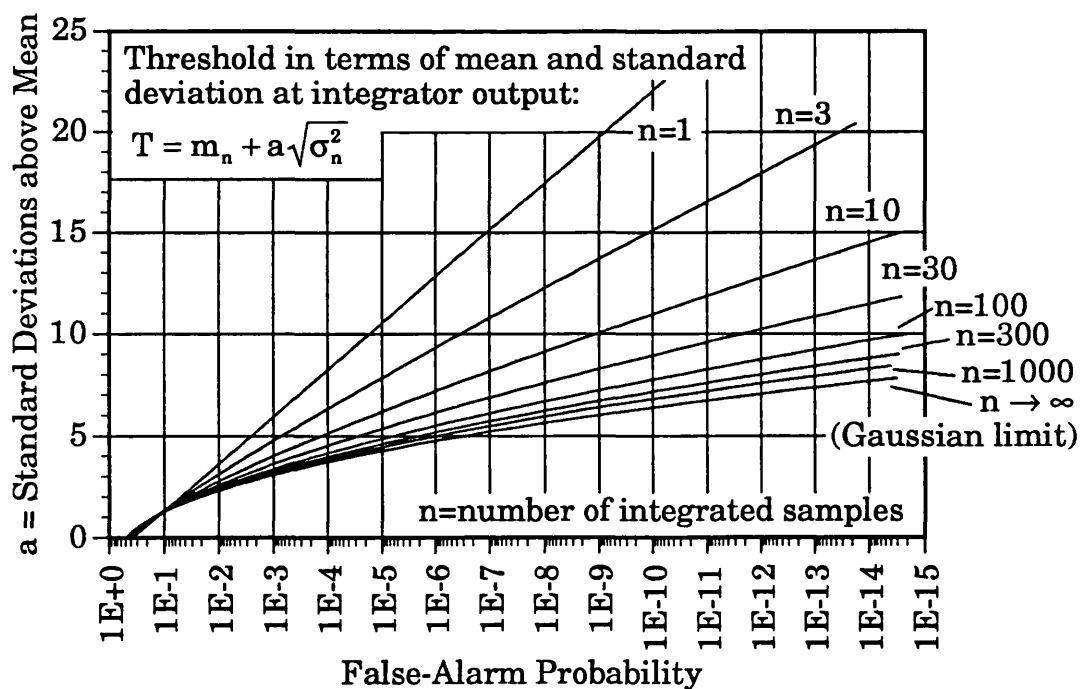
Figure 5-3: False-alarm rate for fully in-phase correlated noise at the cross-correlator inputs (squarer law detector)

the Gaussian limit. Inspection of Fig. 5-2 shows that 100 samples are a reasonable approximation to the Gaussian limit for false-alarm rates up to the order of 10^{-6} or even 10^{-8}. To illustrate the impact of correlation between the input signals of the cross-correlator, the false-alarm probabilities for fully in-phase correlated noise have been calculated. With $\sigma_1 = \sigma_2 = 1$ (unit variances), $\rho_0 = 1$ and $\lambda_0 = 0$ (conditions for square law detection), and bearing in mind that the pdf is now single sided, which accounts for a factor of two, eqn. (5.11) reduces to:

$$P_{fa_n} = \exp(-T) \sum_{k=0}^{n-1} \frac{T^k}{k!}$$

...(5.13)

The false-alarm probability as a function of the threshold setting for fully in-phase correlated noise at the cross-correlator input is shown in Fig. 5-3. In can clearly be seen that due to the correlation, higher threshold settings would be required to achieve the same false-alarm rate for a given number of independent samples.

5.1.3 Probability of detection

The second term to describe the performance of the bipolar threshold detector is its probability of detection P_d. The probability of detection is given by an equation similar to the one for the false-alarm rate. The probability density function to be integrated is now that of signal plus noise present at the detector [Sko 70]:

$$P_d = P(z > T) = \int_T^\infty p_{SN}(z) \, dz$$

...(5.14)

In the case of the symmetric-pair DF system, p_{SN} is the probability density function for the cross-correlator output after integration. The probability of detection is therefore a function of the following parameters:

(1) The signal-to-noise ratio at the antenna elements and hence at the cross-correlator inputs, whereby the signal strength at one of the inputs is dependent on the angle of incidence of the received signal.

(2) The threshold of the detector, which is given by the desired false-alarm rate.

(3) The number of integrated independent samples.

The false-alarm rate, which is determined by the tails of the pdf, has been calculated for different numbers of independent samples in the previous section. No approximations have been used for these evaluations.

To calculate the probability of detection after integration and for a chosen threshold, the cross-correlator output in the presence of a signal must be known. For the calculations, it is assumed that the signals after downconversion to the intermediate frequency can be represented as a sinusoid with narrow band Gaussian noise. (Since the IF bandwidth of the DF system will be much smaller than the RF bandwidth, it is valid to infer that the noise at IF has again Gaussian distribution.) The signal component of the symmetric-pair and the reference element at IF can be expressed in the following form:

$$s_x(t) = A_x \cos(\omega_{IF} + \phi_x)$$

...(5.15)

where A_x is the signal peak amplitude and ϕ_x is the signal phase with:

$\phi_1 = \phi_0 + \phi$ (for the first element of the symmetric-pair)

$\phi_2 = \phi_0 - \phi$ (for the second element of the symmetric-pair)

$\phi_3 = \phi_0$ (for the reference element)

The narrow band Gaussian noise at IF is given by its quadrature representation [Wha 70]:

$$n_x(t) = n_{cx}(t)\cos(\omega_{IF}t) - n_{sx}(t)\sin(\omega_{IF}t)$$

...(5.16)

The index x refers either to the elements of the symmetric-pair (x = 1, 2) or to the reference element (x = 3). $n_{cx}(t)$ and $n_{sx}(t)$ are the quadrature components of the narrow band process which have also zero mean and Gaussian distribution. There is no correlation between $n_{cx}(t)$ and $n_{sx}(t)$. Since the processes under consideration have zero mean, their power and

variance is identical. The power of the narrow band Gaussian noise and its quadrature components is equal to:

$$E\left[n_x^2(t)\right] = E\left[n_{cx}^2(t)\right] = E\left[n_{sx}^2(t)\right] = \sigma_x^2$$

...(5.17)

It is now straightforward to express the signals at the input of the cross-correlators of the 4 channel DF system with the help of the above definitions (for the block diagram of the system see Fig. 5-4.). The reference signal at the correlator inputs is:

$$y_r(t) = \frac{1}{2}\left(s_3(t) + n_3(t)\right)$$

...(5.18)
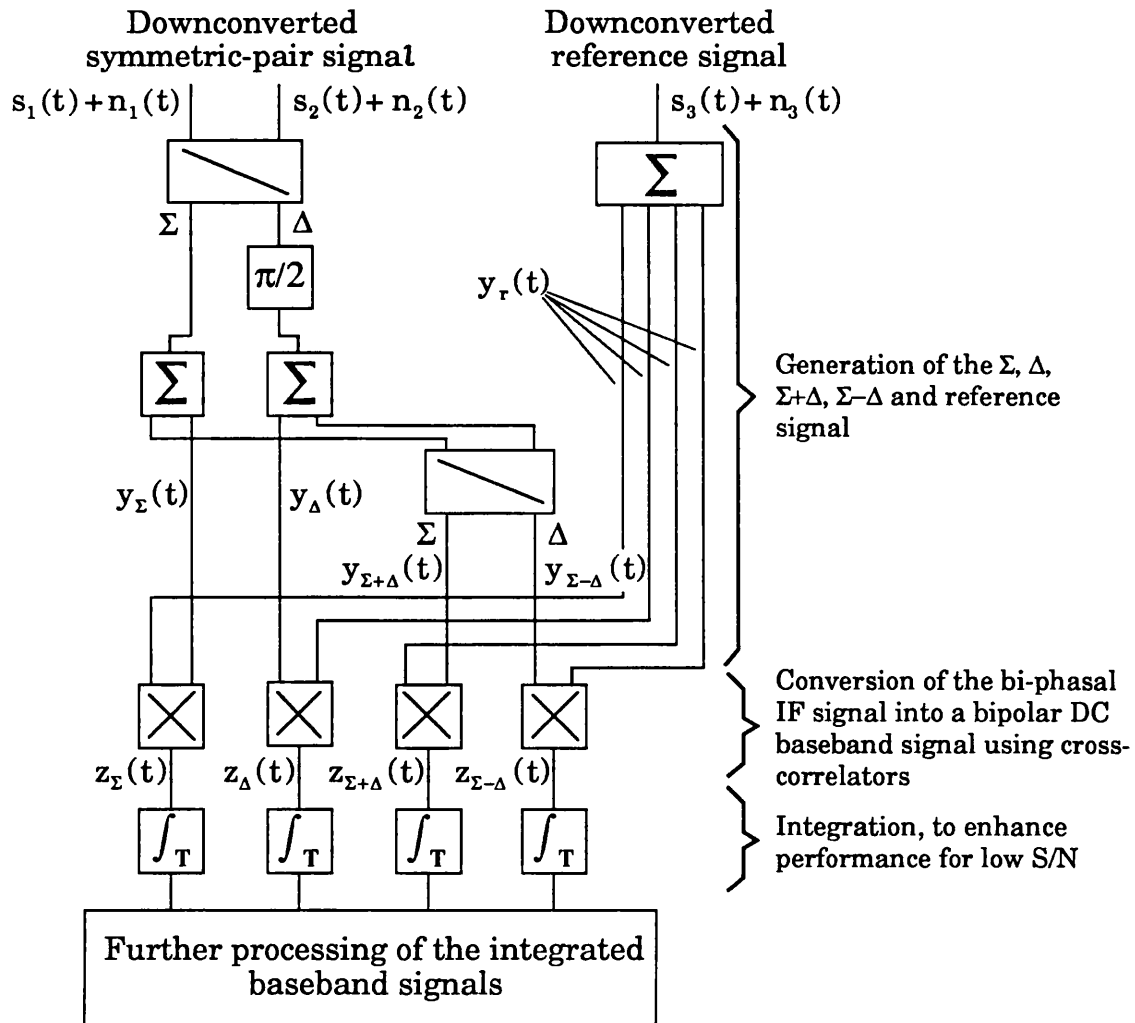
Figure 5-4: Block diagram of the correlator section of the DF system

The signal of the Σ channel at the correlator input is equal to:

$$y_\Sigma(t) = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}\left(\left(s_1(t) + n_1(t)\right) + \left(s_2(t) + n_2(t)\right)\right)\right) \quad \text{...(5.19)}$$

The output of the cross-correlator is then:

$$z_\Sigma(t) = y_\Sigma(t)\, y_r(t) \quad \text{...(5.20)}$$

For further calculations it has been assumed that the amplitudes at the pair and reference antenna elements are equal ($A_1 = A_2 = A_3 = A$) and that $\phi_0 = 0$. Substituting eqn. (5.15) to eqn. (5.19) into eqn. (5.20), and after removal of the double frequency terms, the output of the Σ channel cross-correlator becomes:

$$z_\Sigma(t) = \frac{1}{8}\left[2A^2\cos(\phi) + 2A\cos(\phi)n_{c3}(t) + A\left(n_{c1}(t) + n_{c2}(t)\right)\right.$$
$$\left. + \left(n_{c1}(t) + n_{c2}(t)\right)n_{c3}(t) + \left(n_{s1}(t) + n_{s2}(t)\right)n_{s3}(t)\right] \quad \text{...(5.21)}$$

Applying the same methods and taking the additional 90° phase shifter into account, it is possible to derive the output of the Δ channel cross-correlator:

$$z_\Delta(t) = \frac{1}{8}\left[2A^2\sin(\phi) + 2A\sin(\phi)n_{c3}(t) + A\left(n_{s1}(t) - n_{s2}(t)\right)\right.$$
$$\left. + \left(n_{s1}(t) - n_{s2}(t)\right)n_{c3}(t) + \left(n_{c2}(t) - n_{c1}(t)\right)n_{s3}(t)\right] \quad \text{...(5.22)}$$

The output signals of the $\Sigma+\Delta$ and $\Sigma-\Delta$ channel cross-correlators are then equal to:

$$z_{\Sigma+\Delta}(t) = \frac{1}{\sqrt{2}}\left(z_\Sigma(t) + z_\Delta(t)\right) \quad \text{...(5.23)}$$

$$z_{\Sigma-\Delta}(t) = \frac{1}{\sqrt{2}}\left(z_\Sigma(t) - z_\Delta(t)\right) \quad \text{...(5.24)}$$

Knowing the output of the cross-correlator for signal plus Gaussian noise at the inputs of the direction finder, it would now be necessary to obtain the corresponding pdf to be able to calculate the desired probability

of detection. However, instead of deriving the exact probability density function, it will be shown that it is possible to use the Gaussian pdf as a reasonable approximation instead. This will just be demonstrated for the Σ channel, since the same arguments apply for the other three channels.

The output signal of the Σ cross-correlator consists of three components which are signal x signal, noise x noise and the signal x noise cross products (see eqn. (5.21). The signal x signal term of the correlator output is DC, which yields the mean of the pdf. The pdf of the noise x noise component has been investigated in the false-alarm rate calculations, and it has been shown, that the pdf after the summation of a large number of independent samples converges in the limit to a Gaussian distribution. Looking at the signal x noise cross products, it can be seen that they have Gaussian distribution with the signal term acting merely as an amplitude factor. Furthermore, since the noises at the antenna elements are assumed to be independent, there is no cross correlation between the quadrature components of any of the three narrow band noises. The pdf of the correlator output will therefore be symmetrical and centred on the mean.

The pdf for the probability of detection has, compared to the false-alarm rate pdf, the additional signal x noise components. Because these terms have a Gaussian distribution, the overall pdf of the correlator output will converge even faster to the Gaussian limit as the number of integrated samples increases. Additionally, since this distribution will be used to calculate the probability of detection and *not* the false-alarm rate, only the central region of the pdf is of interest, and differences between the tails of the exact and the Gaussian distribution are of no importance. It has previously been shown (see section 5.1.1) that the shapes of the central regions are virtually identical after the integration of only 10 noise samples. The Gaussian approximation for the calculation of the probability of detection of a single channel is therefore valid in situations, where much less than 100 samples are integrated. (100 samples was the required minimum, to use the Gaussian pdf for false-alarm rate calculations; see previous section.) However, the appropriate curves in Fig. 5-2 must be used to obtain the correct threshold for a wanted false-alarm rate, when less than 100 samples are integrated.

The approximation to the exact pdf of the cross-correlator output is therefore a Gaussian probability density function with mean and variance of the correlator output:

$$p_{SN}(z_\Sigma) = \frac{1}{\sqrt{2\pi\sigma_{nz_\Sigma}^2}} \exp\left[\frac{-\left(z_\Sigma - m_{nz_\Sigma}\right)^2}{2\sigma_{nz_\Sigma}^2}\right]$$

...(5.25)

The derivation and a cross check of the expressions for the mean m_{nz_Σ} and variance $\sigma_{nz_\Sigma}^2$ of the Σ cross-correlator output after summation of n independent samples and for signal plus noise at the inputs is contained in appendix 7. The mean is given by eqn. (A7.9a):

$$m_{nz_\Sigma} = \frac{n}{4}A^2\cos(\phi)$$

...(5.26)

and the formula for the variance has the following form (eqn. (A7.9c):

$$\sigma_{nz_\Sigma}^2 = \frac{n}{16}\left[A^2\,\sigma^2\left(\cos^2(\phi)+0.5\right)+\left(\sigma^2\right)^2\right]$$

...(5.27)

Since $p_{SN}(z_\Sigma)$ is Gaussian, the corresponding probability of detection for the bipolar threshold detector is equal to:

$$P_{d_\Sigma} = P(|z|>T) = \int_{-\infty}^{-T}p_{SN}(z_\Sigma)dz + \int_{T}^{\infty}p_{SN}(z_\Sigma)dz = 1 - \int_{-T}^{T}p_{SN}(z_\Sigma)dz$$

$$= 1 - 0.5\left(\text{erf}\left(\frac{T - m_{nz_\Sigma}}{\sqrt{2\sigma_{nz_\Sigma}^2}}\right) - \text{erf}\left(\frac{-T - m_{nz_\Sigma}}{\sqrt{2\sigma_{nz_\Sigma}^2}}\right)\right)$$

...(5.28)

where the error function is defined as:

$$\text{erf}(z) = \frac{2}{\sqrt{\pi}}\int_{0}^{z}e^{-t^2}dt$$

([Abr 70], eqn. (7.1.1))

and:

erf(-z)=-erf(z)

([Abr 70], eqn. (7.1.9))

Rational approximations of the error function are given by Abramowitz and Stegun (equations 7.1.25 to 7.1.28) [Abr 70]. High level programming languages like *Mathematica* ® have built in routines for the error function [Wol 91].

The threshold T in eqn. (5.28) is a function of the desired false-alarm rate and can be expressed in terms of the mean and standard deviation of the integrated cross-correlator output when no input signal is present. Using eqn. (5.26) and eqn. (5.27) with A=0 (absence of a signal), T is defined as:

$$T = m_{nz_I} + a\sqrt{\sigma^2_{nz_I}}\Big|_{A=0} = a\sqrt{\frac{n}{16}(\sigma^2)^2} = a\frac{\sqrt{n}}{4}\sigma^2 \qquad \text{...(5.29)}$$

The constant a determines the false-alarm rate and can be extracted from Fig. 5-2, using the trace appropriate trace.

The probability of detection as a function of the signal-to-noise ratio at the antenna elements for different numbers of integrated samples is shown in Fig. 5-5. It can be seen that the curves become steeper as the number of integrated independent samples increases. As expected, lower false-alarm rates require higher signal-to-noise ratios to achieve a certain probability of detection.

The signal-to-noise ratio at the antenna elements as a function of the false-alarm rate and the number of integrated samples for a given probability of detection is presented in Fig. 5-6. These results have been calculated by changing iteratively the value of the signal-to-noise ratio at the antenna elements, using eqn. (5.26) to eqn. (5.28), until the required probability of detection had been obtained. As illustrated in Fig. 5-6, the signal-to-noise ratio required to obtain a certain probability of detection decreases by about 5 to 6 dB for an increase of the number of integrated samples by a decade. This is similar to results obtained by Marcum for the case of a square law detector. He reports changes of the same order of magnitude [Mar 60].

Fig. 5-7 shows the probability of detection as a function of the phase difference ϕ between one of the symmetric-pair elements and the phase
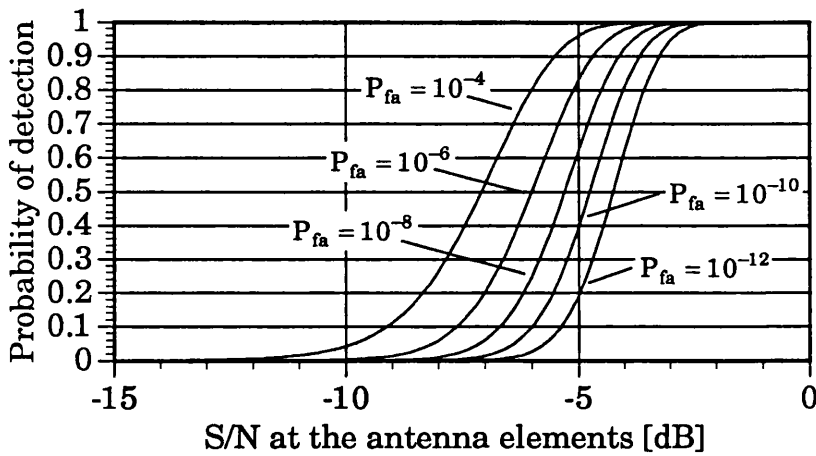
Number of integrated samples:

n=100

Phase difference:

$\phi = 0°$

Figure 5-5a
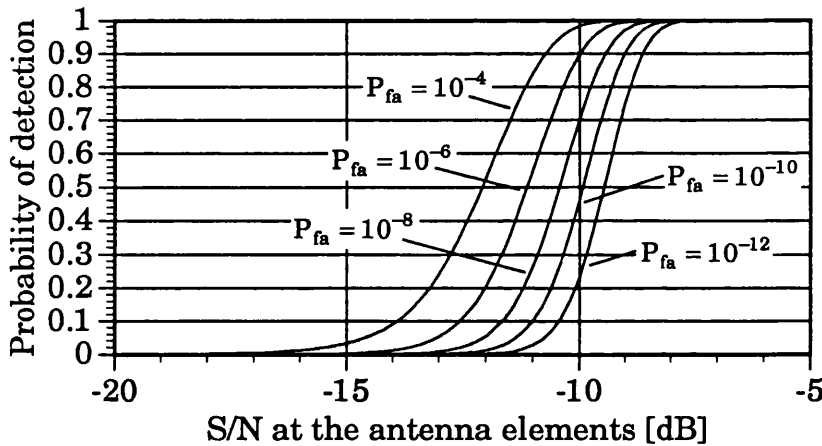
Number of integrated samples:

n=1000

Phase difference:

$\phi = 0°$

Figure 5-5b
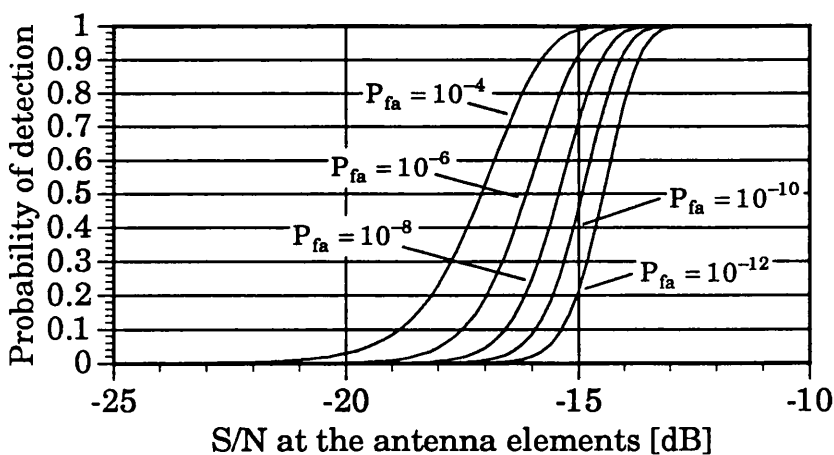
Number of integrated samples:

n=10000

Phase difference:

$\phi = 0°$

Figure 5-5c

Figure 5-5: *Probability of detection as a function of the signal-to-noise ratio at the antenna elements for different false-alarm probabilities and different numbers of integrated samples.*

Probability of
detection:

$$P_{d_r} = 0.5$$

Phase difference:

$$\phi = 0°$$

Figure 5-6a

n = number of integrated independent samples

Probability of
detection:

$$P_{d_r} = 0.9$$

Phase difference:

$$\phi = 0°$$

Figure 5-6b

n = number of integrated independent samples

Probability of
detection:

$$P_{d_r} = 0.99$$

Phase difference:

$$\phi = 0°$$

Figure 5-6c

n = number of integrated independent samples
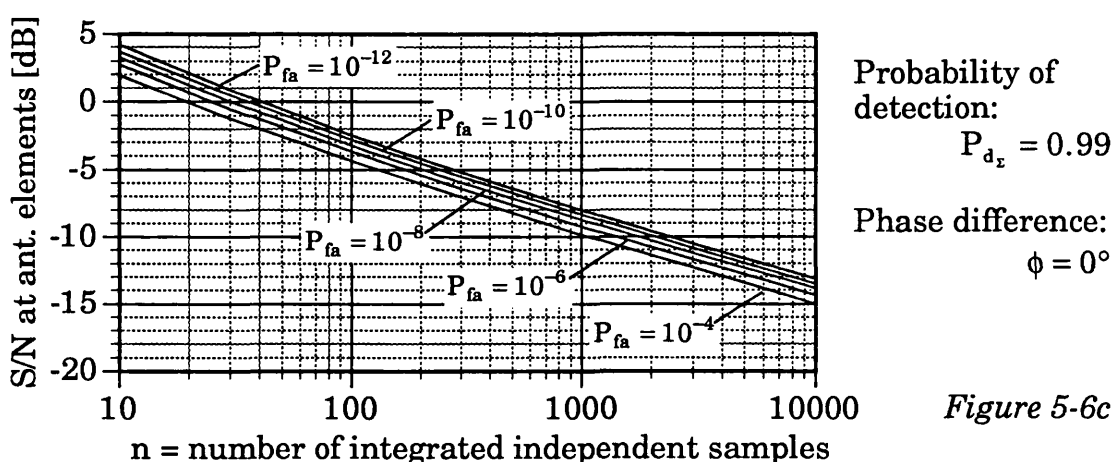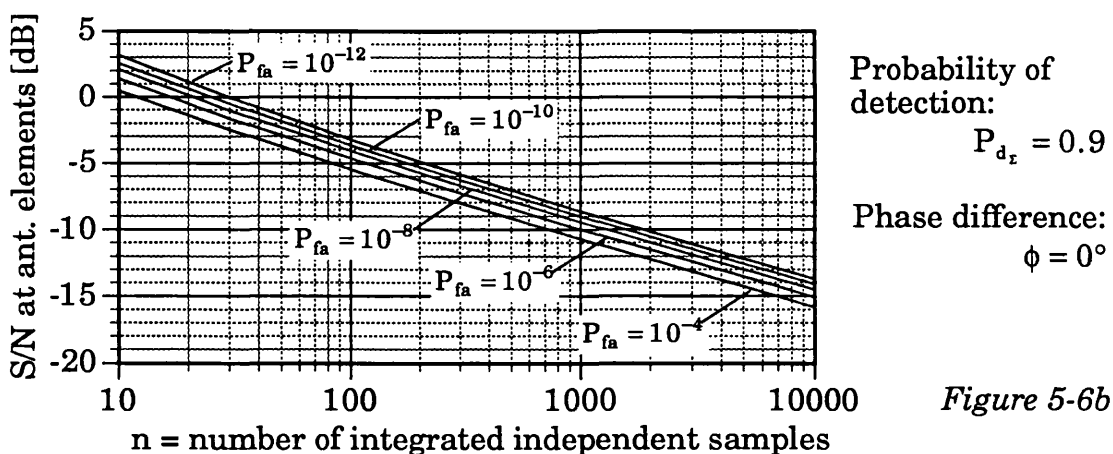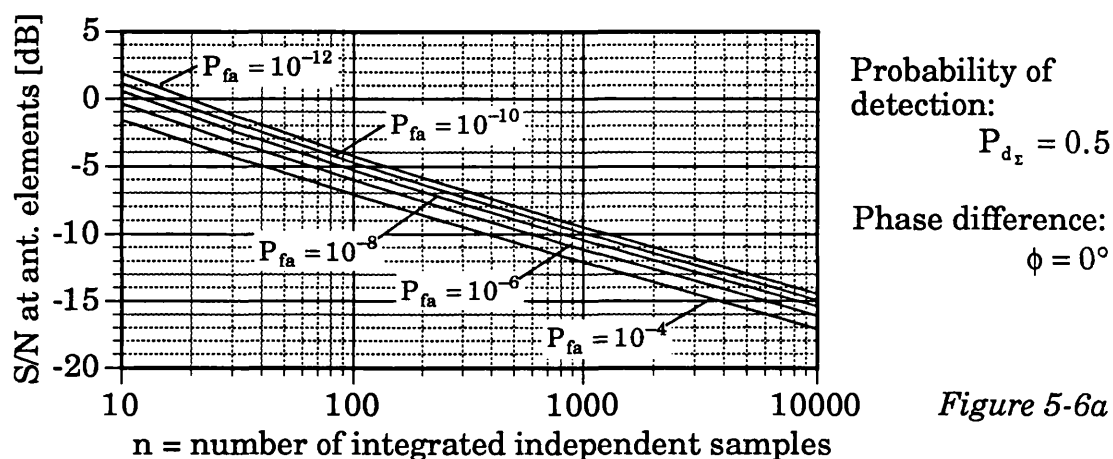
Figure 5-6: Required signal-to-noise ratio at the antenna elements as a function of the number of integrated samples at the cross-correlator output for different false-alarm probabilities, constant propabilitiy of detection and nonfluctuating signal.

centre of the pair. The advantage of this representation is that ϕ depends on the angle of arrival as well as the spacing of the symmetric-pair, which means that it covers all possible values of the two variables.
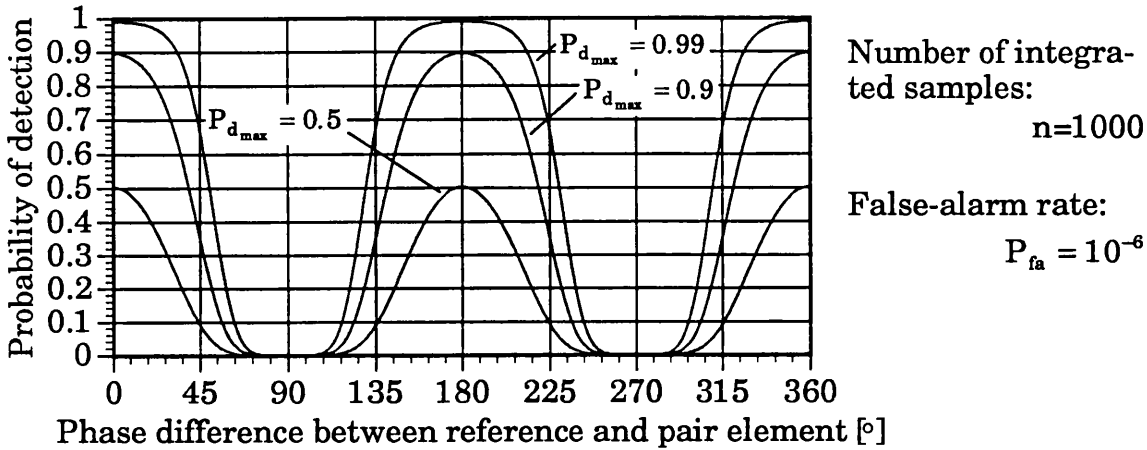
Figure 5-7: Probability of detection of the Σ channel detector as a function of the phase difference between the reference and symmetric-pair element

As can be seen, the probability of detection as a function of ϕ is periodic. For angles of arrival where ϕ is an odd multiple of 90°, the output signal of the Σ channel is always zero, which entails zero probability of detection. Since the Σ signal on its own extracts 1 bit of DF information, i.e. two different subsectors can be distinguished, two periods of the probability of detection function are contained in a full cycle of ϕ (see above).

The above calculations are only for a single channel, and it is therefore impossible to incorporate the information provided by the 'don't know' state of the bipolar threshold detector. As outlined in chapter 4.2, at least two channels are required to be able to use this additional data. The next section will demonstrate that the extra information provided by the 'don't know' state allows us to achieve virtually constant probability of detection irrespective of the received signal's angle of arrival.

5.2. PERFORMANCE PREDICTIONS FOR THE FOUR CHANNEL SYSTEM

5.2.1. Assumptions made for the calculations

The outputs of the four channel system are the Σ, Δ, $\Sigma+\Delta$ and $\Sigma-\Delta$ signals of the symmetric-pair under consideration. However, two of the four signals are linear combinations of the Σ and Δ outputs, which reduces the relevant mathematics to bivariate calculations. To keep the bivariate analysis tractable it is assumed that 100 or more samples are always integrated, which allows us to use Gaussian distribution theory, as shown in the previous section. Since the Σ and Δ outputs are therefore Gaussian variates, their joint pdf will be jointly Gaussian. Good treatments of bivariate Gaussian distributions can be found in Johnson and Kotz and Leon-Garcia ([Joh 72], [Leo 89]).

It is furthermore still assumed that the noise at all antenna elements is uncorrelated. The only correlation between the output signals of the four channels is introduced by the presence of a signal.

5.2.2. The bivariate Gaussian probability density function

The bivariate Gaussian (normal) distribution of the two random variables Z_Σ and Z_Δ has the following standard form:

$$p(z_\Sigma, z_\Delta) = \frac{1}{2\pi\sigma_{nz_\Sigma}\sigma_{nz_\Delta}\sqrt{1-\rho_{\Sigma\Delta}{}^2}} \exp\left[\frac{-\sigma_{nz_\Delta}^2 \left(z_\Sigma - m_{nz_\Sigma}\right)^2 + }{} \right.$$

$$\left. \frac{2\rho_{\Sigma\Delta}\sigma_{nz_\Sigma}\sigma_{nz_\Delta}\left(z_\Sigma - m_{nz_\Sigma}\right)\left(z_\Delta - m_{nz_\Delta}\right) - \sigma_{nz_\Sigma}^2\left(z_\Delta - m_{nz_\Delta}\right)^2}{2\sigma_{nz_\Sigma}^2\sigma_{nz_\Delta}^2\left(1-\rho_{\Sigma\Delta}{}^2\right)} \right]$$

...(5.30)

Where:

z_Σ and z_Δ are possible values of the two random variables

m_{nz_Σ} and m_{nz_Δ} their respective means after summation of n samples

$\sigma_{nz_\Sigma}^2$ and $\sigma_{nz_\Delta}^2$ their respective variances after summation of n samples

$\rho_{\Sigma\Delta}$ the correlation coefficient of the two random variables

The expressions for m_{nz_Σ}, m_{nz_Δ}, $\sigma^2_{nz_\Sigma}$, $\sigma^2_{nz_\Delta}$ and $\rho_{\Sigma\Delta}$ are derived in appendix 7 (eqn. (A7.9a), eqn. (A7.10a), eqn. (A7.9c), eqn. (A7.10c) and eqn. (A7.14)). It should be noted that this appendix also contains an analysis that cross checks these expressions.

To simplify further calculations, the following substitutions have been made:

$$a = \sigma_{nz_\Sigma}\left(z_\Delta - m_{nz_\Delta}\right)$$

$$b = \sigma_{nz_\Delta}\left(z_\Sigma - m_{nz_\Sigma}\right)$$

$$v = \sigma^2_{nz_\Sigma}\sigma^2_{nz_\Delta}\left(1 - \rho_{\Sigma\Delta}^{\ 2}\right)$$

Using these substitutions and by adding and subtracting the factor $\rho_{\Sigma\Delta}^{\ 2}\, b^2$ in the numerator, eqn. (5.30) can be modified to:

$$p(z_\Sigma, z_\Delta) = \frac{1}{\sqrt{2\pi}}\exp\left[\frac{-\left(1-\rho_{\Sigma\Delta}^{\ 2}\right)b^2}{2\,v}\right]\frac{1}{\sqrt{2\pi v}}\exp\left[\frac{-\left(a-\rho_{\Sigma\Delta}b\right)^2}{2\,v}\right] \qquad ...(5.31)$$

The second part of the formula has the form of a Gaussian pdf with variate a, mean $\rho_{\Sigma\Delta}b$ and variance v. This form of the bivariate Gaussian pdf will later be very useful for the calculation of the probability of detection where the two variates Z_Σ and Z_Δ are correlated due to the presence of a signal.

The purpose of the rest of this section is to investigate the properties of the bivariate Gaussian distribution with respect to the signals of the symmetric-pair direction finder. The bivariate Gaussian pdf is centred at the point $(m_{nz_\Sigma}, m_{nz_\Delta})$ and has a bell shape that depends on the values of $\sigma^2_{nz_\Sigma}$, $\sigma^2_{nz_\Delta}$ and $\rho_{\Sigma\Delta}$. If the two variates are uncorrelated and have equal variance (i.e. $\rho_{\Sigma\Delta}=0$ and $\sigma^2_{nz_\Sigma} = \sigma^2_{nz_\Delta}$), the contours of constant pdf are circles. In all other cases, the contours are ellipses (see Fig. 5-8). The angle of the major axis of the ellipse with the z_Σ axis is given by [Joh 72]:

$$\theta = \frac{1}{2}\arctan\left[\frac{2\rho_{\Sigma\Delta}\sigma_{nz_\Sigma}\sigma_{nz_\Delta}}{\sigma^2_{nz_\Sigma} - \sigma^2_{nz_\Delta}}\right] \qquad ...(5.32)$$

After insertion of the expressions for $\sigma_{nz_\Sigma}^2$, $\sigma_{nz_\Delta}^2$ and $\rho_{\Sigma\Delta}$ from appendix 7 in the above equation, the angle between the major axis of the ellipse and the z_Σ axis becomes:

$$\theta = \frac{1}{2}\arctan\left[\frac{\sin(2\phi)}{\cos^2(\phi)-\sin^2(\phi)}\right] = \phi$$

...(5.33)

This means that in the case of the symmetric-pair DF system, the bivariate pdf rotates around the origin as a function of the angle of arrival and the pair spacing, because ϕ is the phase difference between one of the symmetric-pair elements and the phase centre of the pair.

The next problem to be addressed is whether the major and minor axis of the ellipse (i.e. a and b in Fig. 5-8) are a function of ϕ, or whether the shape of the rotating ellipse is independent of the angle of arrival and the pair spacing. The pdf is constant for values of z_Σ and z_Δ for which the argument of the exponent in the equation for the bivariate Gaussian pdf is constant. Johnson and Kotz give the following formula for an ellipse that contains $100\alpha\%$ of the distribution [Joh 72] $(0 \leq \alpha < 1)$:

$$\left(\frac{z_\Sigma - m_{nz_\Sigma}}{\sigma_{nz_\Sigma}}\right)^2 - 2\rho_{\Sigma\Delta}\left(\frac{z_\Sigma - m_{nz_\Sigma}}{\sigma_{nz_\Sigma}}\right)\left(\frac{z_\Delta - m_{nz_\Delta}}{\sigma_{nz_\Delta}}\right) + \left(\frac{z_\Delta - m_{nz_\Delta}}{\sigma_{nz_\Delta}}\right)^2$$

$$= -2\left(1 - \rho_{\Sigma\Delta}^2\right)\log(1 - \alpha)$$

...(5.34)

Since the major and minor axes of the rotating ellipse are of interest, a new coordinate system whose axes are aligned with those of the ellipse must be defined (see Fig. 5-8). The necessary coordinate transformation into the new x, y coordinate system can be accomplished by using the following rotation matrix:

$$\begin{bmatrix} z_\Sigma \\ z_\Delta \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

...(5.35)

Substituting the expressions for m_{nz_Σ}, m_{nz_Δ}, $\sigma_{nz_\Sigma}^2$, $\sigma_{nz_\Delta}^2$ and $\rho_{\Sigma\Delta}$ from appendix 7 in eqn. (5.34) and after the coordinate transformation and

some reasonably straightforward but lengthy mathematics, eqn. (5.34) can be written in the form of an equation for an ellipse:

$$\frac{(x-m)^2}{a^2} + \frac{y^2}{b^2} = 1$$

...(5.36)

with:

$$a^2 = \frac{-n\,\sigma^2\left(3A^4 + 8A^2\sigma^2 + 4(\sigma^2)^2\right)\log(1-\alpha)}{16(A^2 + 2\sigma^2)}$$

$$b^2 = \frac{-n\,\sigma^2\left(3A^4 + 8A^2\sigma^2 + 4(\sigma^2)^2\right)\log(1-\alpha)}{16(3A^2 + 2\sigma^2)}$$

$$m = \frac{n\,A^2}{4}$$

Inspection of eqn. (5.36) shows that the major and minor axis of the ellipse are independent of ϕ and hence of the angle of arrival as well as the spacing of the symmetric-pair. The two axes of the ellipse are only a function of the signal amplitude A, the noise power σ^2 and the number of
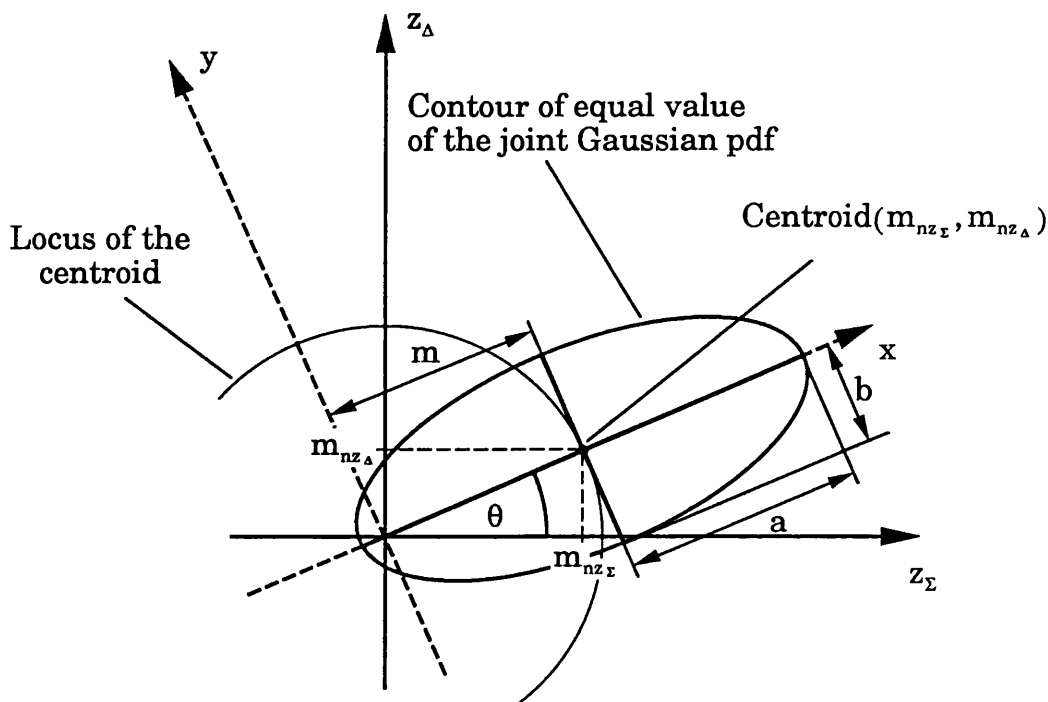
Figure 5-8: Contour of equal value of the bivariate Gaussian pdf

integrated samples n. The offset m of the major axis can be cross checked by performing the following calculation with the help of Fig. 5-8:

$$m = \sqrt{m_{n z_\Sigma}{}^2 + m_{n z_\Delta}{}^2} = \sqrt{\left(\frac{nA^2}{4}\cos(\phi)\right)^2 + \left(\frac{nA^2}{4}\sin(\phi)\right)^2} = \frac{n A^2}{4}$$

In summary, it can be said that the shape of the bivariate Gaussian density is independent of the angle of arrival, and the centroid of the pdf rotates on a circle around the origin (see Fig. 5-8). The radius m of this circle is dependent on the signal strength and is zero in the absence of a signal. The angle between the major axis of the ellipse and the z_Σ axis is ϕ, and therefore a function of the angle of arrival and the pair spacing.

5.2.3. False-alarm rate

The four channel system allows us to utilise the additional information provided by the 'don't know' state of the bipolar threshold detectors. As a consequence, the definition of a false-alarm is slightly more complicated, since signals which are too weak to cross the threshold can also provide useful information. This fact must be taken into account in the definition of a false-alarm. As shown in chapter 4.2, where the 'don't know' state was introduced, at least one out of the four channels of an octant selecting DF system must be allowed to flag the 'don't know' state. This is required in order to cover also angles of arrival around the grating lobe nulls of the symmetric-pair. On the other hand, up to three out of the four channels may produce a 'don't know' output, and it is still possible to define unambiguously the required subsector. In the following false-alarm rate calculations, all three possible cases, i.e. up to one, two or three 'don't knows' per measurement, will be considered. To understand why it is sensible to investigate all three options, it is necessary to look at the respective definitions of a false-alarm and the associated implications for the threshold settings.

Superficially, it could be argued that up to three 'don't know' events would be by far the most favourable case, since this yields the widest 'don't know' zone which allows signals with lower signal-to-noise ratios to be detected, as illustrated in Fig. 5-9. This argument is only valid as long as

the threshold settings remain unchanged. However, in order to maintain the same false-alarm rate for all cases, the threshold has to be changed according to the maximum number of permitted 'don't knows'.
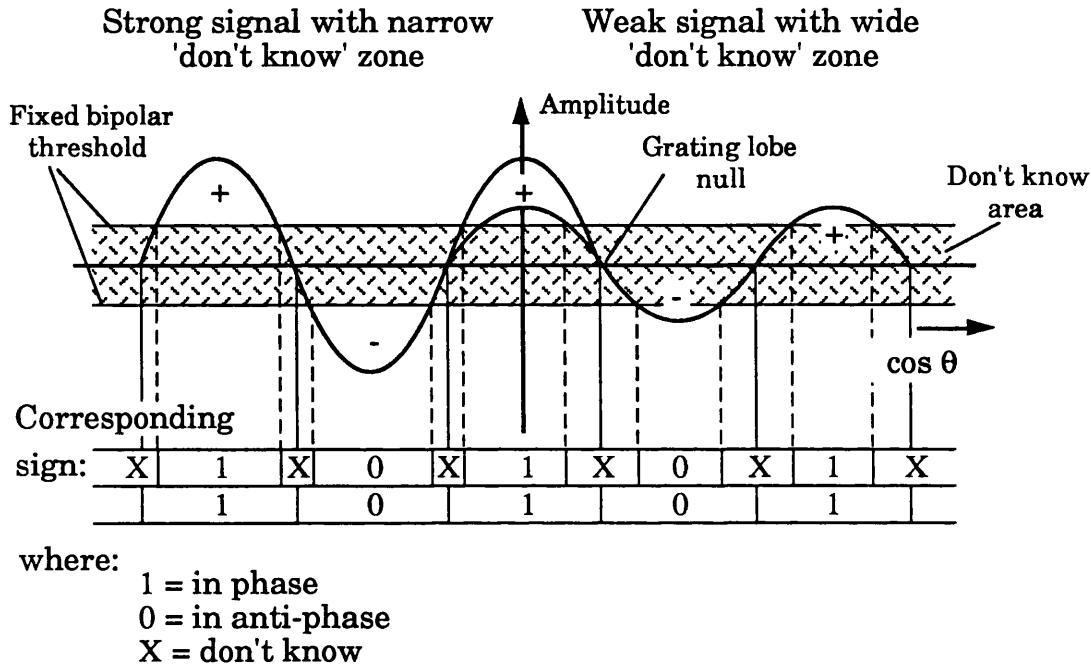
Figure 5-9 'Don't know' zone width for different signal strengths

When up to three 'don't know' states are allowed, a false-alarm occurs as soon as at least one of the four detectors is triggered by noise at the input of the DF system. For just up to one permitted 'don't know', the event that three out of the four or all four detectors are triggered is called a false-alarm. A comparison of these two cases shows that for a given false-alarm rate, higher threshold settings are required in the former example due to the different definition of the event 'false-alarm'.

Two contrary mechanisms can therefore be identified:

(1) The wider the permitted 'don't know' zone, the lower is the required signal-to-noise ratio to detect signals from angles of arrival around grating lobe nulls. (Fixed threshold assumed.)

(2) To maintain a constant false-alarm rate, the threshold setting must be increased for wider 'don't know' zones. Higher signal-to-noise ratios are then necessary in order to detect a signal.

The question is now, which one of the two effects dominates and whether the DF system should consequently be designed for the minimum or maximum permitted 'don't know' zone width.

In order to answer this question, the different false-alarm probabilities as a function of the threshold setting must be evaluated. The definition of a false-alarm in the three investigated cases is:

(1) **Up to one 'don't know' allowed:**

A false-alarm occurs, when all four or any three of the four detectors are triggered by noise at the input of the DF system.

$$P_{fa(1)} = \underbrace{\left(P_\Sigma \cap P_\Delta \cap P_{\Sigma+\Delta} \cap \overline{P}_{\Sigma-\Delta}\right)}_{P_I} \cup \underbrace{\left(P_\Sigma \cap P_\Delta \cap \overline{P}_{\Sigma+\Delta} \cap P_{\Sigma-\Delta}\right)}_{P_{II}} \cup$$

$$\underbrace{\left(P_\Sigma \cap \overline{P}_\Delta \cap P_{\Sigma+\Delta} \cap P_{\Sigma-\Delta}\right)}_{P_{III}} \cup \underbrace{\left(\overline{P}_\Sigma \cap P_\Delta \cap P_{\Sigma+\Delta} \cap P_{\Sigma-\Delta}\right)}_{P_{IV}} \cup$$

$$\underbrace{\left(P_\Sigma \cap P_\Delta \cap P_{\Sigma+\Delta} \cap P_{\Sigma-\Delta}\right)}_{P_V}$$

...(5.37)

where:

$$P_{\Sigma/\Delta} = P\left(|z_{\Sigma/\Delta}| > T\right) \overset{\substack{\text{since mutually} \\ \text{exclusive events}}}{=} P\left(z_{\Sigma/\Delta} > T\right) + P\left(z_{\Sigma/\Delta} < -T\right)$$

$$\overline{P}_{\Sigma/\Delta} = P\left(|z_{\Sigma/\Delta}| < T\right) = P\left(-T < z_{\Sigma/\Delta} < T\right)$$

$$P_{\Sigma\pm\Delta} = P\left(\left|(z_\Sigma \pm z_\Delta)/\sqrt{2}\right| > T\right) = P\left((z_\Sigma \pm z_\Delta)/\sqrt{2} > T\right) + P\left((z_\Sigma \pm z_\Delta)/\sqrt{2} < -T\right)$$

$$\overline{P}_{\Sigma/\Delta} = P\left(\left|(z_\Sigma \pm z_\Delta)/\sqrt{2}\right| < T\right) = P\left(-T < (z_\Sigma \pm z_\Delta)/\sqrt{2} < T\right)$$

Since P_I, P_{II}, P_{III}, P_{IV} and P_V are mutually exclusive events, $P_{fa(1)}$ becomes:

$$P_{fa(1)} = P_I + P_{II} + P_{III} + P_{IV} + P_V$$

(2) **Up to two 'don't knows' allowed:**

A false-alarm occurs, when all four or any three or any two of the four detectors are triggered by noise at the input of the DF system.

$$P_{fa(2)} = 1 - [\,\underbrace{\left(P_\Sigma \cap \overline{P}_\Delta \cap \overline{P}_{\Sigma+\Delta} \cap \overline{P}_{\Sigma-\Delta}\right)}_{P_I} \cup \underbrace{\left(\overline{P}_\Sigma \cap P_\Delta \cap \overline{P}_{\Sigma+\Delta} \cap \overline{P}_{\Sigma-\Delta}\right)}_{P_{II}} \cup$$

$$\underbrace{\left(\overline{P}_\Sigma \cap \overline{P}_\Delta \cap P_{\Sigma+\Delta} \cap \overline{P}_{\Sigma-\Delta}\right)}_{P_{III}} \cup \underbrace{\left(\overline{P}_\Sigma \cap \overline{P}_\Delta \cap \overline{P}_{\Sigma+\Delta} \cap P_{\Sigma-\Delta}\right)}_{P_{IV}} \cup$$

$$\underbrace{\left(\overline{P}_\Sigma \cap \overline{P}_\Delta \cap \overline{P}_{\Sigma+\Delta} \cap \overline{P}_{\Sigma-\Delta}\right)}_{P_V}\,]$$

...(5.38)

Since P_I, P_{II}, P_{III}, P_{IV} and P_V are again mutually exclusive events, $P_{fa(2)}$ becomes:

$$P_{fa(2)} = 1 - \left(P_I + P_{II} + P_{III} + P_{IV} + P_V\right)$$

(3) **Up to three 'don't knows' allowed:**

A false-alarm occurs, when at least one of the four detectors is triggered by noise at the input of the DF system.

$$P_{fa(3)} = 1 - \left(\overline{P}_\Sigma \cap \overline{P}_\Delta \cap \overline{P}_{\Sigma+\Delta} \cap \overline{P}_{\Sigma-\Delta}\right)$$

...(5.39)

To obtain the desired false-alarm probabilities, the bivariate Gaussian pdf must be integrated over certain regions in the z_Σ z_Δ plane. It is convenient to use graphical methods to determine the relevant area for each of the three different false-alarm definitions given in eqn. (5.37) to eqn. (5.39). Fig. 5-10 illustrates this technique for the example of $z_\Sigma > T$, $z_\Delta < -T$, $(z_\Sigma - z_\Delta)/\sqrt{2} > T$ and $\left|(z_\Sigma + z_\Delta)/\sqrt{2}\right| < T$, which is one of the possible combinations of P_I in eqn. (5.37). All different permutations of threshold crossings that can yield a false-alarm must be examined in this way, and the total area over which the pdf has to be integrated is obtained by piecing together all individual regions. Fig. 5-11 shows the respective zones for up to one, two or three permitted 'don't knows'.

Inspection of the correlation coefficient $\rho_{\Sigma\Delta}$ (eqn. (A7.14)) shows that z_Σ and z_Δ are uncorrelated when no signal is present (i.e. A=0). Since the two variates are Gaussian, they are also independent when they are uncorrelated, and the joint pdf simplifies to:

$$p(z_\Sigma, z_\Delta) = p(z_\Sigma)\, p(z_\Delta)$$

...(5.40)

z_Δ

$z_\Sigma > T$

$(1+\sqrt{2})\,T$

$\left|\dfrac{z_\Sigma + z_\Delta}{\sqrt{2}}\right| < T$

$\sqrt{2}\,T$

T

$\dfrac{z_\Sigma - z_\Delta}{\sqrt{2}} > T$

$-(1+\sqrt{2})\,T$ $-\sqrt{2}\,T$ $-T$ T $\sqrt{2}\,T$ $(1+\sqrt{2})\,T$ z_Σ

$-T$

$-\sqrt{2}\,T$

$z_\Delta < -T$

$-(1+\sqrt{2})\,T$

Region defined by the intersection of the four different areas
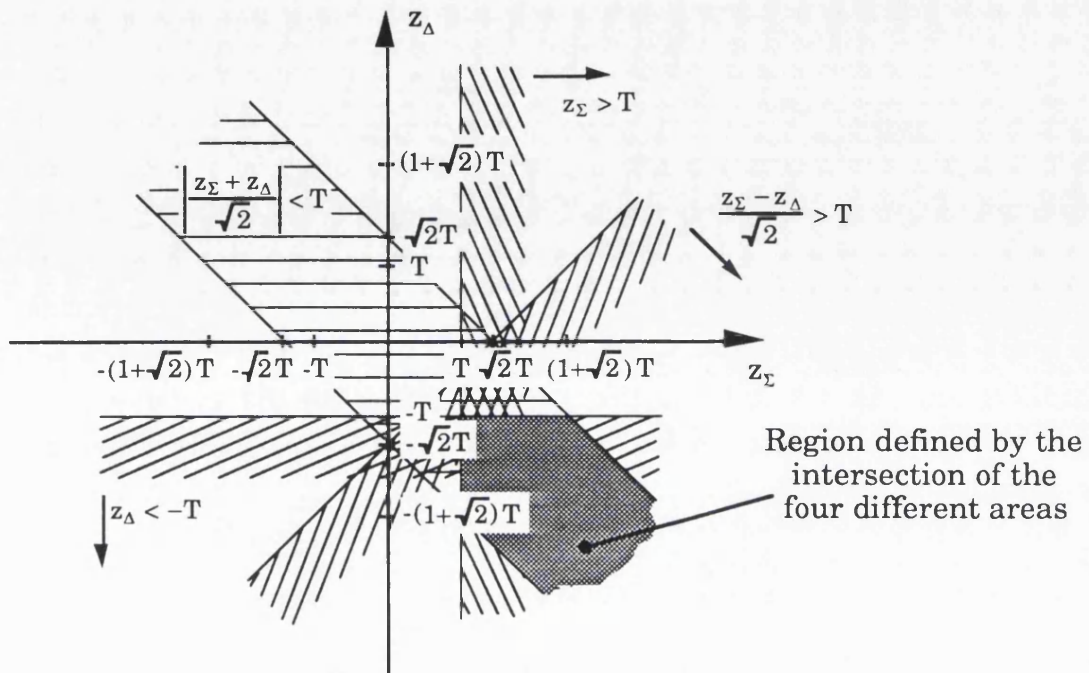
Figure 5-10: Region in the $z_\Sigma z_\Delta$ plane defined by the intersection of $z_\Sigma > T$, $z_\Delta < -T$, $\left(z_\Sigma - z_\Delta\right)/\sqrt{2} > T$ and $\left|\left(z_\Sigma + z_\Delta\right)/\sqrt{2}\right| < T$

z_Δ

$\sqrt{2}\,T$ $(1+\sqrt{2})\,T$

T

z_Σ
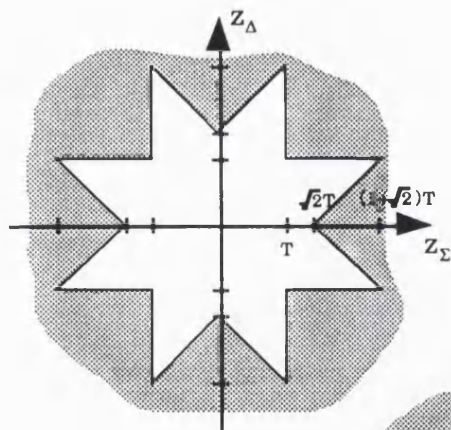
Figure 5-11a
(up to 1 'don't know')

z_Δ

$(\sqrt{2}-1)\,T$

$\sqrt{2}\,T$

T

z_Σ
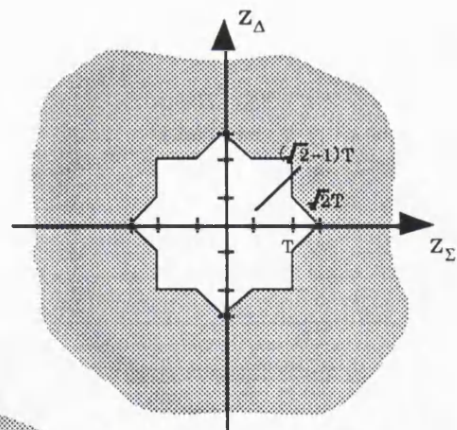
Figure 5-11b
(up to 2 'don't knows')

z_Δ

$(\sqrt{2}-1)\,T$

$\sqrt{2}\,T$

T

z_Σ

Complement

P_{fa} = integration over this region
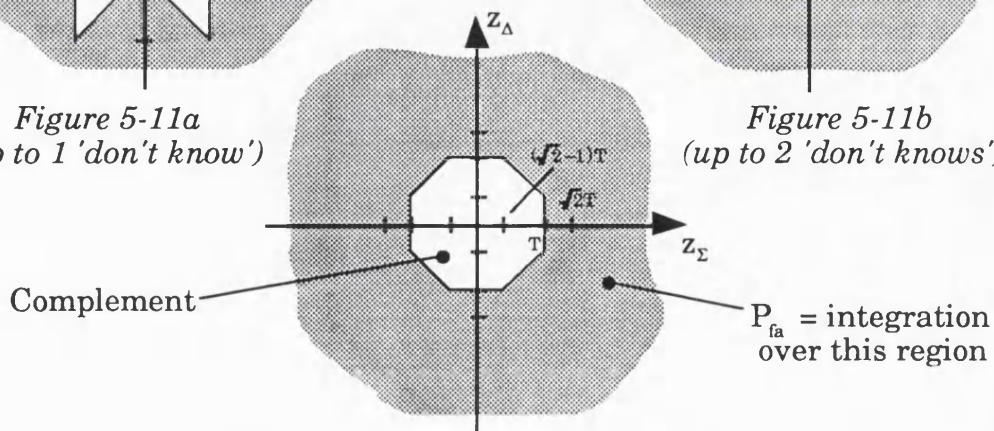
Figure 5-11c (up to 3 'don't knows')

Figure 5-11: Regions, over which to integrate for false-alarm rate calculations for up to 1, 2 or 3 'don't knows' per measurement

$p(z_\Sigma)$ and $p(z_\Delta)$ are Gaussian distributions with variances $\sigma^2_{nz_\Sigma}$ and $\sigma^2_{nz_\Delta}$ and have zero mean, since m_{nz_Σ} and m_{nz_Δ} are zero when no signal is present. For the calculation of the false-alarm rate, it is convenient to evaluate the complement of the regions defined in Fig. 5-11 and subtract it from one to obtain P_{fa}. Since the pdf has zero mean and $\sigma^2_{nz_\Sigma} = \sigma^2_{nz_\Delta}$ for A = 0, the bivariate Gaussian distribution is centred at the origin of the coordinate system and the contours of equal probability are circles, as discussed in section 5.2.2. For the integration, the relevant areas can be broken down into smaller sectors, and by utilising the symmetry of the pdf, it is not necessary to determine all sectors. The false-alarm probability can now be obtained by evaluating:

$$P_{fa} = 1 - \iint\limits_{\substack{\text{complement} \\ \text{region}}} p(z_\Sigma)\, p(z_\Delta)\, dz_\Sigma dz_\Delta$$

$$...(5.41)$$

The contributions of the different sectors can be calculated by using one of the following three different types of integral:

Type I:

$$\int_{-x}^{x} p(z_\Sigma) dz_\Sigma \int_{-y}^{y} p(z_\Delta) dz_\Delta = \text{erf}\left(\frac{x}{\sqrt{2\sigma^2_{nz_\Sigma}}}\right) \text{erf}\left(\frac{y}{\sqrt{2\sigma^2_{nz_\Delta}}}\right)$$

$$...(5.42a)$$

Type II:

$$\int_{x_1}^{x_2} p(z_\Sigma) dz_\Sigma \int_{-y}^{y} p(z_\Delta) dz_\Delta = 0.5\left[\text{erf}\left(\frac{x_2}{\sqrt{2\sigma^2_{nz_\Sigma}}}\right) - \text{erf}\left(\frac{x_1}{\sqrt{2\sigma^2_{nz_\Sigma}}}\right)\right]\text{erf}\left(\frac{y}{\sqrt{2\sigma^2_{nz_\Delta}}}\right)$$

$$...(5.42b)$$

Type III:

$$\int_{x_1}^{x_2} p(z_\Sigma)\left[\int_{z_\Sigma+y_1}^{y_2} p(z_\Delta)dz_\Delta\right]dz_\Sigma = 0.5\int_{x_1}^{x_2} p(z_\Sigma)\left[\text{erf}\left(\frac{y_2}{\sqrt{2\sigma^2_{nz_\Sigma}}}\right) - \text{erf}\left(\frac{z_\Sigma+y_1}{\sqrt{2\sigma^2_{nz_\Sigma}}}\right)\right]dz_\Sigma$$

$$...(5.42c)$$

The remaining integration in integrals of type III must be done numerically. The formulas for the Gaussian distribution, the variances and the threshold are given in eqn. (5.25), eqn. (A7.9c), eqn. (A7.10c) and eqn. (5.29) respectively. Fig. 5-12 shows the false-alarm probabilities as a

function of the threshold setting for up to one, two or three 'don't knows' per measurement. It can be seen that the highest threshold settings are required for up to three 'don't knows'. This confirms the argument of the previous quantitative discussion that the threshold must be increased for wider 'don't know' zones to maintain a constant false-alarm rate.
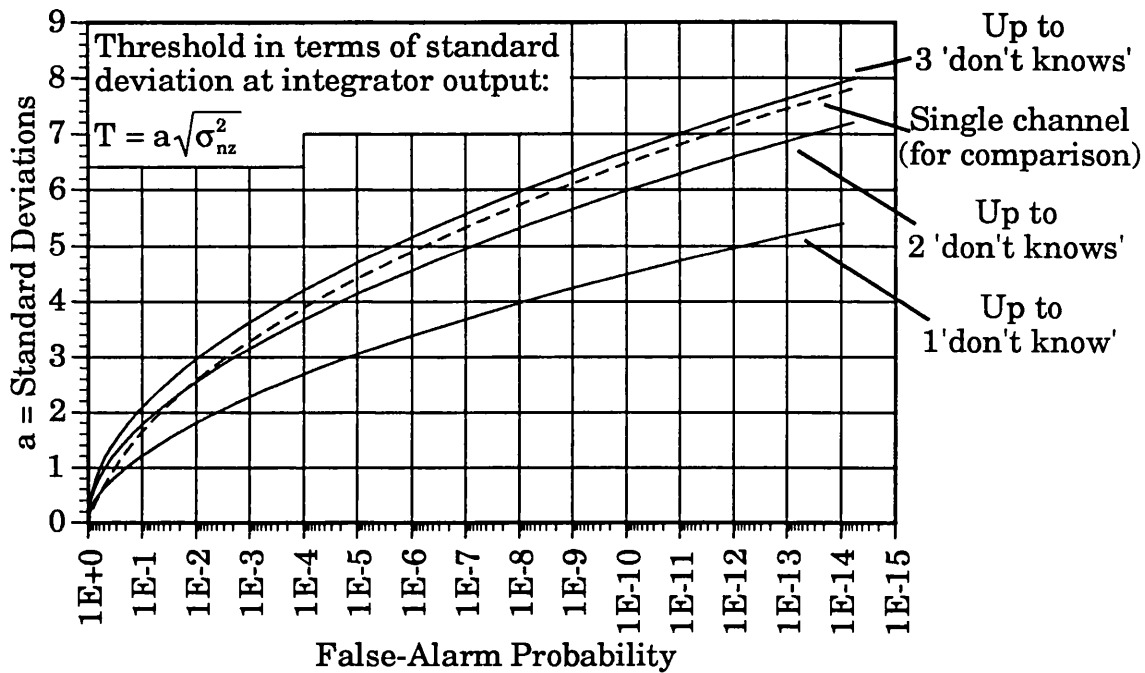
Figure 5-12: *False-alarm probability for up to 1, 2 or three 'don't knows' per measurement in a 4 channel system*

5.2.4. Probability of detection

The probability of detection for a given false-alarm rate can now be calculated employing similar methods. The areas over which the bivariate Gaussian pdf has to be integrated are the same as given in Figs. 5-11a-c. However, the form of the density function is now different due to the presence of a signal, which causes correlation between the Σ and the Δ signals. The contours of equal probability of the bivariate pdf are now ellipses instead of circles. The probability of detection is therefore defined as:

$$P_d = 1 - \iint\limits_{\substack{\text{complement} \\ \text{region}}} p(z_\Sigma, z_\Delta) \, dz_\Delta dz_\Sigma$$

...(5.43)

where $p(z_\Sigma, z_\Delta)$ is the bivariate Gaussian pdf given in eqn. (5.31). One of the two integrations can be carried out by changing the integration variable z_Δ to a, to enable the integration of the Gaussian pdf with variate a, mean ρb and variance v using the error function. The complement region can again be subdivided into separately integrable smaller sectors, and the contribution of one of these sectors is equal to:

$$P\left(z_{\Sigma_1} \leq z_\Sigma \leq z_{\Sigma_2}, z_{\Delta_1} \leq z_\Delta \leq z_{\Delta_2}\right) = \frac{1}{\sqrt{2\pi}} \int\limits_{z_{\Sigma_1}}^{z_{\Sigma_2}} \exp\left[\frac{-\left(1 - \rho_{\Sigma\Delta}^2\right)b^2}{2v}\right]$$

$$\frac{0.5}{\sigma_{nz_\Sigma}^2}\left[\text{erf}\left(\frac{\sigma_{nz_\Sigma}^2\left(z_{\Delta_2} - m_{nz_\Delta}\right) - \rho_{\Sigma\Delta}b}{\sqrt{2}v}\right) - \text{erf}\left(\frac{\sigma_{nz_\Sigma}^2\left(z_{\Delta_1} - m_{nz_\Delta}\right) - \rho_{\Sigma\Delta}b}{\sqrt{2}v}\right)\right] dz_\Sigma$$

...(5.44)

with:

$$b = \sigma_{nz_\Delta}\left(z_\Sigma - m_{nz_\Sigma}\right)$$

$$v = \sigma_{nz_\Sigma}^2 \sigma_{nz_\Delta}^2 \left(1 - \rho_{\Sigma\Delta}^2\right)$$

The expressions for m_{nz_Σ}, m_{nz_Δ}, $\sigma_{nz_\Sigma}^2$, $\sigma_{nz_\Delta}^2$ and $\rho_{\Sigma\Delta}$ are given in appendix 7 and the bounds for z_Δ will sometimes be a function of z_Σ. The remaining integration in eqn. (5.44) must be done using numerical algorithms.

The results for up to one 'don't know' event per measurement will now be discussed in some detail, since they exhibit all the typical features of the four channel system. After that, the discussion will be extended to the case of up to two and up to three 'don't knows'.

The probability of detection as a function of ϕ, which is the phase difference between the reference and pair element, is shown in Figs. 5-13a-c. As |in the case of a single channel, P_d is again a periodic function. This time, there are eight periods in a full cycle of ϕ, since the four channel system extracts three bit of DF information per measurement. In contrast to the single channel (see Fig. 5-7), there is a finite probability of

$P_{d_{max}} = 0.99$

Number of integrated samples:

$n = 1000$

False-alarm rate:

$P_{fa} = 10^{-6}$

Up to 1 'don't know'

$P_{d_{max}} = 0.25$

Figure 5-13a

Number of integrated samples:

$n = 1000$

Maximum probability of detection:

$P_{d_{max}} = 0.9$

Up to 1 'don't know'

Figure 5-13b

False-alarm rate:

$P_{fa} = 10^{-6}$

Maximum probability of detection:

$P_{d_{max}} = 0.9$

Up to 1 'don't know'

Figure 5-13c

Figure 5-13: *Probability of detection of the 4 channel DF system as a function of the phase difference between the reference and symmetric-pair element*

detection for all values of ϕ, even for $P_{d_{max}}$ as low as 0.25 (see Fig. 5-13a). This is due to the use of the 'don't know' state, which allows us to deal with angles of arrival around the grating lobe nulls of the pair. The change in probability of detection is larger for medium values of $P_{d_{max}}$ as illustrated in Fig. 5-13a. This phenomenon can be explained by looking at the graphs for the probability of detection as a function of the signal-to-noise ratio at the antenna elements, as given in Figs. 5-14a/b. Changing ϕ has an effect similar to changing the signal-to-noise ratio at all elements, with the difference that in the former case, the signal strength at only one instead of both phase detector inputs changes. As can be seen in Figs. 5-14a/b, a reduction in signal-to-noise ratio entails a decrease of the probability of detection, and for a given change in signal-to-noise ratio, the probability of detection will change more for medium values of P_d, since there the slope of the curve is steeper.

The same argument can be used to explain the differences in the plots of Fig. 5-13b and Fig. 5-13c where the probability of detection is given for different false-alarm rates and different numbers of integrated samples. Fig. 5-14a and Fig. 5-14b show the corresponding curves of the probability of detection as a function of the signal-to-noise ratio at the antenna elements. The lower ¦the desired false-alarm rate or the larger the number of integrated samples, the steeper is the curve of the probability of detection, which yields larger excursions of the periodic P_d functions in Fig. 5-13b and Fig. 5-13c.

Figs. 5-14a/b show also the traces for $\phi = 22.5°$. This value of ϕ, or odd multiples of it, yield the lowest probability of detection for a given false-alarm rate and number of integrated samples (see Figs. 5-13a-c). The difference between the trace for $\phi = 0°$ and $\phi = 22.5°$ gives the amount, by which the minimum signal-to-noise ratio at the antenna elements would have to be increased to compensate for the 'dips' in Figs. 5-13a-c. The resulting minimum probability of detection would be irrespective of the angle of arrival at least $P_{d_{max}}$. As can be seen in Fig. 5-14b, the difference between the two curves increases only slightly with the number of integrated samples. For a probability of detection of for example 0.9 and a false-alarm rate of 10^{-6}, an increase of about 1 dB in minimum signal-to-noise ratio would be sufficient. However, a change of the false-alarm rate clearly changes the amount of increase in signal-to-noise ratio needed to

Number of integra-
ted samples:

$$n=1000$$

Up to 1 'don't know'

——— $\phi = 0°$

- - - - $\phi = 22.5°$

Figure 5-14a

False-alarm rate:

$$P_{fa} = 10^{-6}$$

Up to 1 'don't know'

——— $\phi = 0°$

- - - - $\phi = 22.5°$
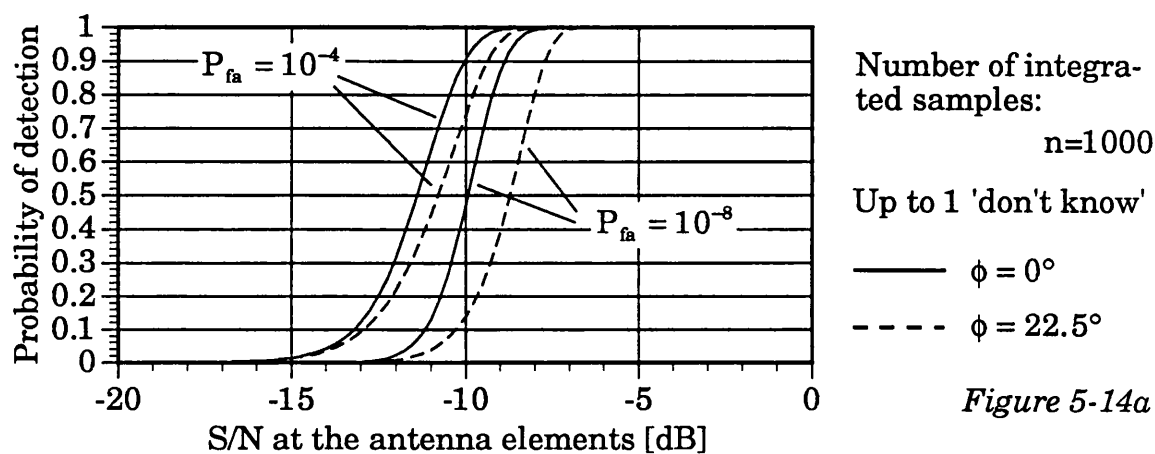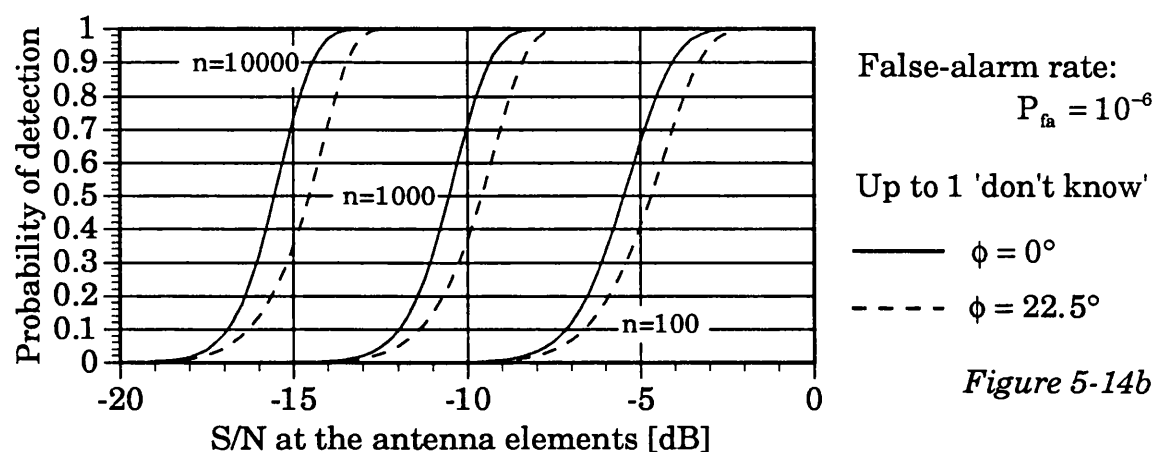
Figure 5-14b

Figure 5-14: Probability of detection as a function of the signal-to-noise ratio at the antenna elements

Probability of
detection:

$$P_d = 0.9$$

Phase difference:

$$\phi = 0°$$

Up to 1 'don't know'
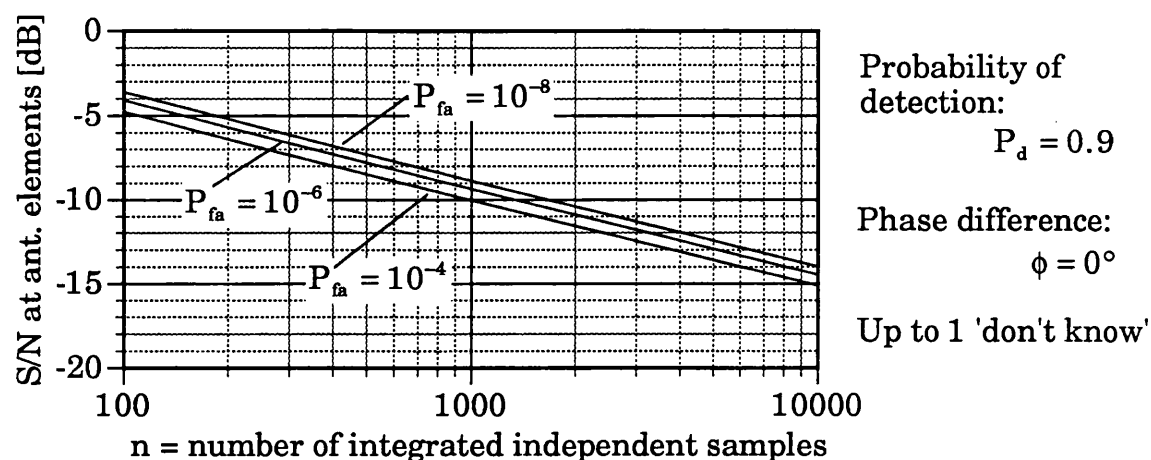
Figure 5-15: Required signal-to-noise ratio at the antenna elements as a function of the number of integrated samples for different false-alarm probabilities and constant probability of detection of the 4 channel DF system

compensate the 'dips' (see Fig. 5-14a). The lower P_{fa}, the larger the required increase.

The signal-to-noise ratio at the antenna elements as a function of the number of integrated samples for a constant probability of detection (P_d = 0.9) and for different false-alarm rates is given in Fig. 5-15. The minimum number of integrated samples in the graph is 100, since the Gaussian assumption does not hold for smaller values. A comparison with single channel results from Fig. 5-6b shows that the four channel system requires a signal-to-noise ratio, which is about 0.5 dB higher, to achieve the same false-alarm probability. The difference is of the same order of magnitude for P_d = 0.99 and P_d = 0.5; the respective graphs for the four channel system are not given here.

The corresponding graphs for up to two or three 'don't knows' have generally the same tendencies as the ones for up to one 'don't know'. However, inspection of Figs. 5-16a-c shows that the larger the 'don't know' zones (i.e. the more 'don't knows' allowed), the smaller is the difference in signal-to-noise ratio between the traces for $\phi = 0°$ and $\phi = 22.5°$. As a consequence, the probability of detection of the four channel system as a function of the phase difference between the reference and symmetric-pair element is less dependent on the ϕ, or the angle of arrival respectively (see Fig. 5-17). It should also be noted that the trace for up to two 'don't knows' is offset by half a period relative to the other two plots. This just reflects the fact that the boundaries of the subsections defined by two 'don't knows' are also offset by half a subsector width relative to the subsector pattern defined for one or three 'don't knows' (see chapter 4.2).

The required minimum signal-to-noise ratio as a function of the number of integrated samples for up to one, two or three 'don't knows' is given in Fig. 5-18. As can be seen, all three traces lie within one dB of each other. This means that none of the two contrary mechanisms identified in section 5.2.3. clearly dominates. Only a gain in signal-to-noise ratio of less than one dB can be achieved by allowing wider 'don't know' zones. (The average gain [averaged over ϕ] in signal-to-noise ratio relative to up to 1 'don't know' is 0.69 dB for up to 2 'don't knows' and 0.93 dB for up to 3 'don't knows'.) However, as demonstrated in Fig. 5-17, wider 'don't know' zones reduce the dependence of the probability of detection on the angle of
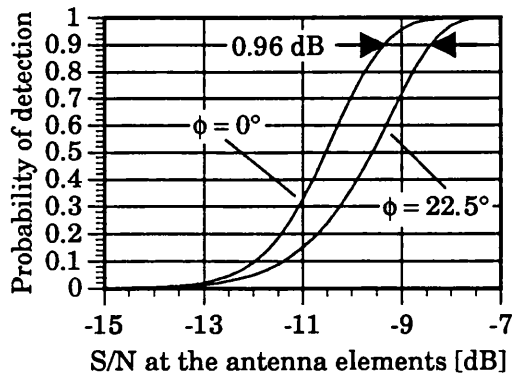
Figure 5-16a (up to 1'don't know')
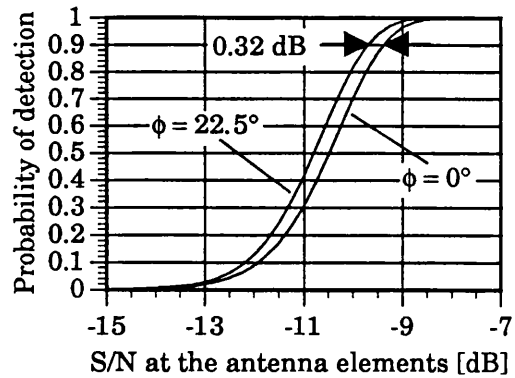
Figure 5-16b (up to 2'don't knows')
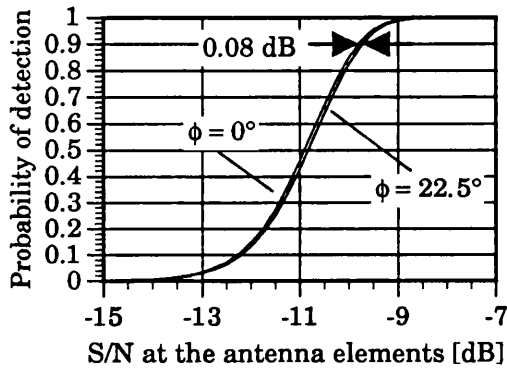
Figure 5-16c (up to 3'don't knows')

Number of integra-
ted samples:

n=1000

False-alarm rate:

$P_{fa} = 10^{-6}$

Figure 5-16: *Probability of detection as a function of the signal-to-noise ratio at the antenna elements for different numbers of permitted 'don't knows' (traces for $\phi = 0$ ° and $\phi = 22.5$ °)*

Phase difference between reference and pair element [°]
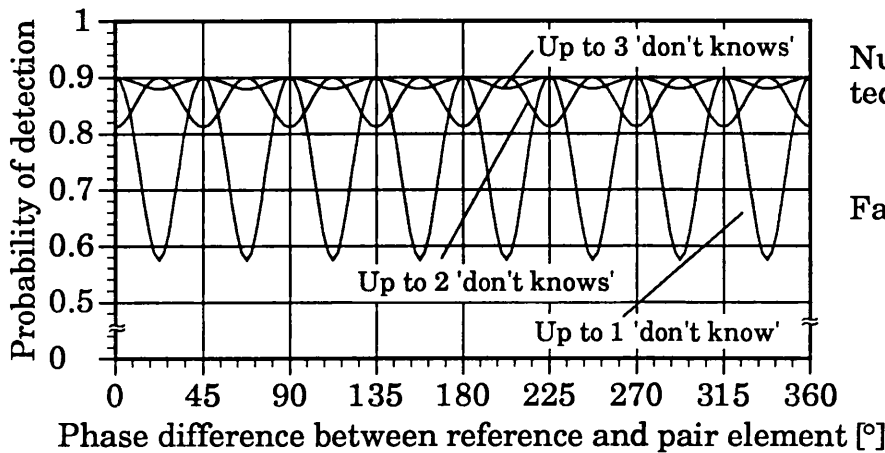
Number of integra-
ted samples:

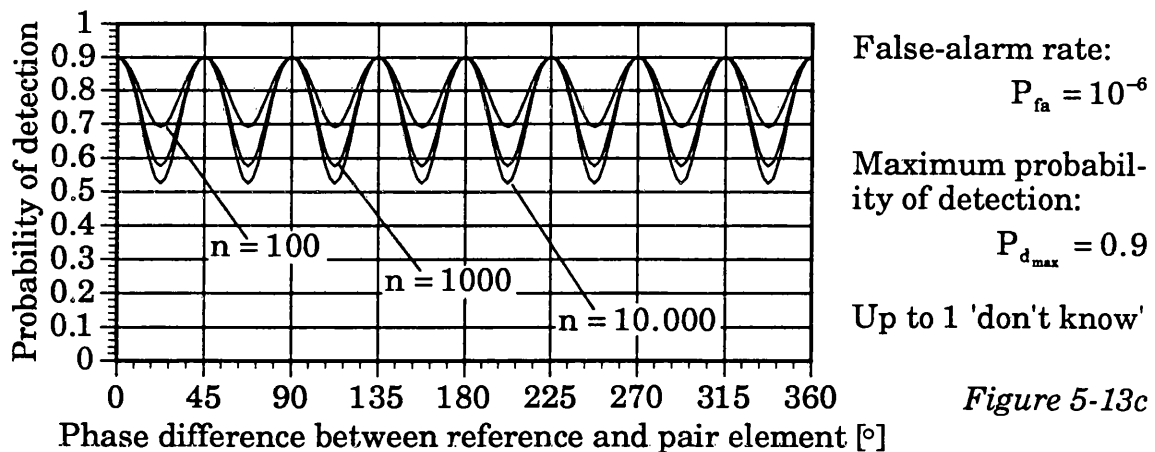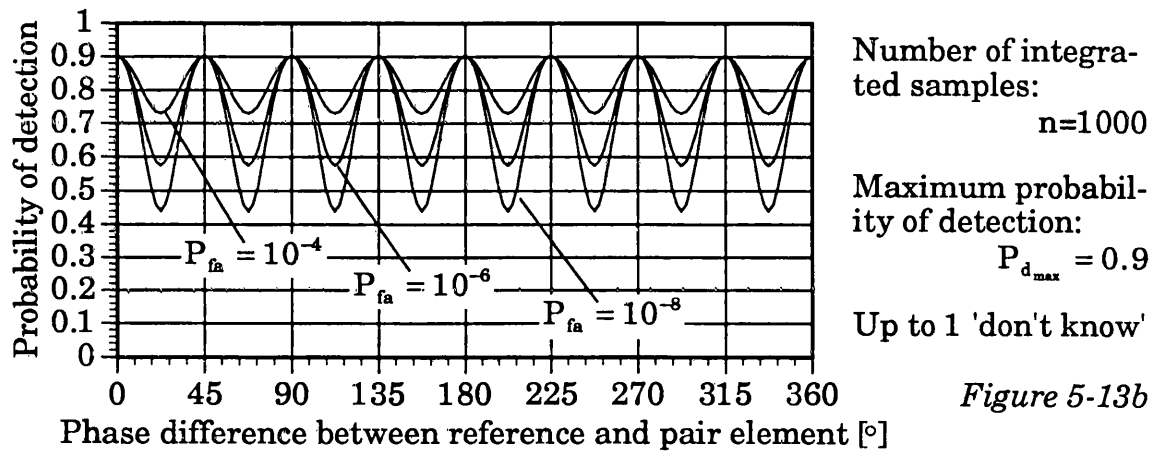n=1000

False-alarm rate:
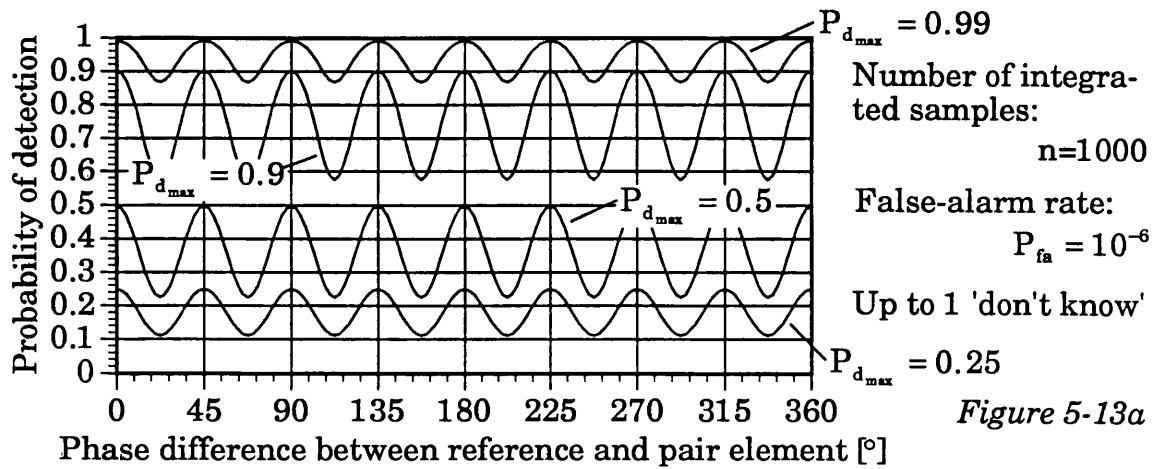
$P_{fa} = 10^{-6}$

Figure 5-17: *Probability of detection of the 4 channel DF system as a function of the phase difference between the reference and symmetric-pair element*

arrival for low signal-to-noise ratios at the antenna elements, which is highly desirable. The thresholds and the data processing of the symmetric-pair direction finding system should therefore be designed for up to three 'don't knows' events per measurement.
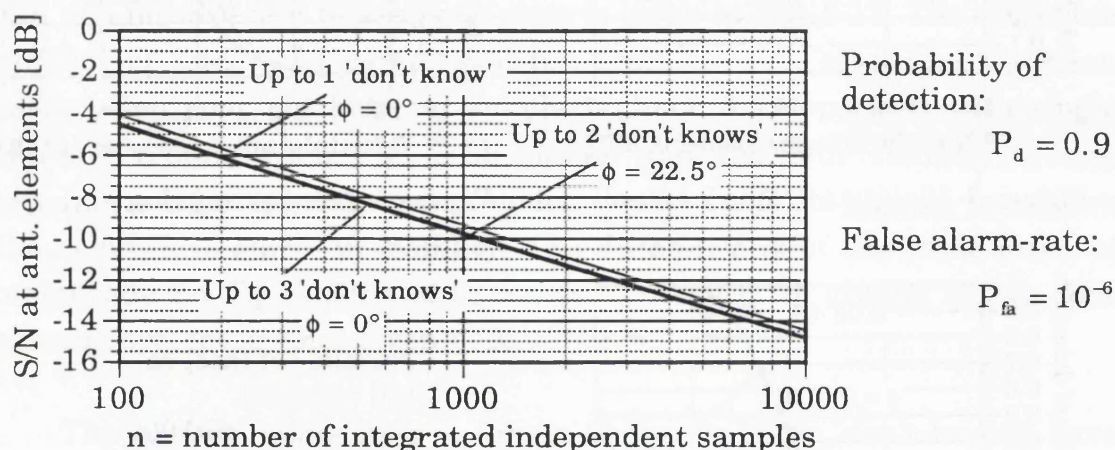
Figure 5-18: Required signal-to-noise ratio at the antenna elements as a function of the number of integrated samples for up to one, two or three 'don't know' events per measurement.

5.3. THRESHOLD CONSIDERATIONS

All previous calculations have been done on the premises that the noise power at the input of the system does not fluctuate and the gain of the IF amplifier stages is fixed. These assumptions slightly simplified the relevant mathematics and allowed us to concentrate on the investigation of the relations between threshold setting and the corresponding false-alarm rates and probabilities of detection. However, the operating conditions for a direction finding system will not be as ideal, and measures have to be considered to cope with varying background noise and changing IF amplification. (The AGC at IF is used to increase the dynamic range over which the system can operate.) Two different approaches will be considered and assessed in the following sections. The first one uses a fixed threshold of the bipolar detector and assumes that the output signal plus noise power of the AGC stage is always constant. The second option uses a variable threshold, whose setting is derived from the estimated actual noise power at the system input.

5.3.1. Threshold set for constant signal plus noise power at the output of the AGC stages

This method of setting the threshold is based on the realistic assumption that the aggregate signal plus noise power at the output of the IF AGC stage is always constant. Hence, the dynamic range of the AGC circuitry must be such that this output power level is also reached in the absence of a signal, i.e. by amplification of the input noise. This situation yields the highest possible noise level at the bipolar threshold detectors. Setting the threshold levels for the desired false-alarm rate for this worst case, the actual false-alarm rate will be even lower in the presence of a signal. Fig. 5-19 shows the design of the AGC stage as suggested in chapter 4.4. As can be seen, the reference channel has an active AGC loop and acts as a 'master' to derive the AGC voltages for all three channels. The two symmetric-pair channel amplifiers act as 'slaves' that follow the 'master', to avoid gain imbalances between the three channels. Using the reference channel as 'master' has the advantage that the AGC signal is not affected when the system switches between different symmetric-pairs in order to take a complete bearing.
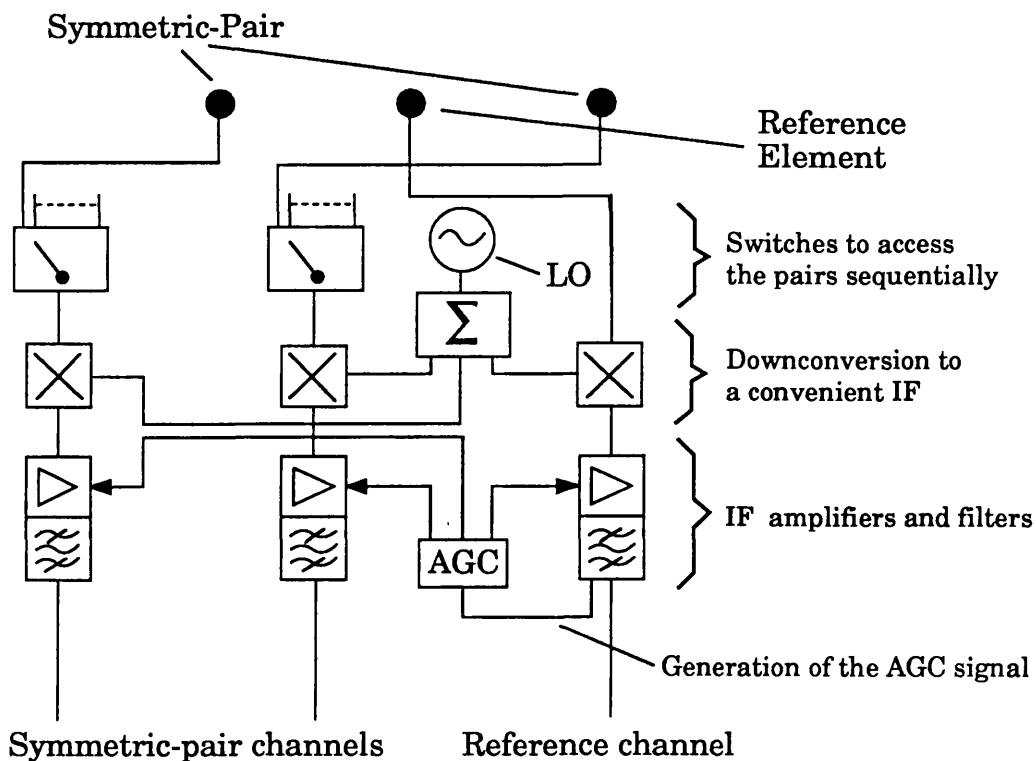
Figure 5-19: Downconversion and IF amplification of the signals

The condition that the signal plus noise power at the AGC output is constant can be expressed as follows:

$$P_{AGC_{out}} = G_{AGC}{}^2 \left(\frac{A^2}{2} + N \right) = const. \tag{5.45}$$

where A is the signal peak amplitude and N the noise power in the symmetric-pair and reference channels at IF. G_{AGC} represents the voltage gain of the IF amplifier. Neglecting the loss of the RF components and the conversion loss of the mixer, the gain of the AGC amplifier for a given noise power and signal-to-noise ratio at the antenna elements is then defined as:

$$G_{AGC} = \sqrt{\frac{P_{AGC_{out}}}{\frac{A^2}{2} + N}} = \sqrt{\frac{P_{AGC_{out}}[mW]}{10^{\left(\frac{S/N[dB]+N[dBm]}{10}\right)} + 10^{\left(\frac{N[dBm]}{10}\right)}}} \tag{5.46}$$

To obtain the relevant formulae for the calculation of the probability of detection, the expressions for the signal and noise components at IF (eqn. (5.15) and eqn. (5.16)) must be multiplied with the factor G_{AGC}. After consideration of this factor in the rest of the analysis, it turns out that eqn. (5.43) and eqn. (5.44) can be used to evaluate the probability of detection subject to the following modifications of the expressions for the mean, variance and threshold:

$$m_{nz_{\Sigma/\Delta}}{}^* = G_{AGC}{}^2\, m_{nz_{\Sigma/\Delta}} \tag{5.47a}$$

$$\sigma^2_{nz_{\Sigma/\Delta}}{}^* = G_{AGC}{}^4\, \sigma^2_{nz_{\Sigma/\Delta}} \tag{5.47b}$$

$$T^* = T\, G_{AGC}{}^2\big|_{A=0} \tag{5.47c}$$

where $G_{AGC}\big|_{A=0}$ is the gain of the IF amplifiers in the absence of a signal, which can be obtained from eqn. (5.46) with S/N=-∞dB. It should be noted that the correlation coefficient $\rho_{\Sigma\Delta}$ is independent of the AGC amplifier gain, since the amplifiers of the three IF channels have the same amplification. Inserting the starred parameters from above in eqn. (5.44) and using eqn. (5.46) to calculate G_{AGC} for different signal-to-noise ratios, it
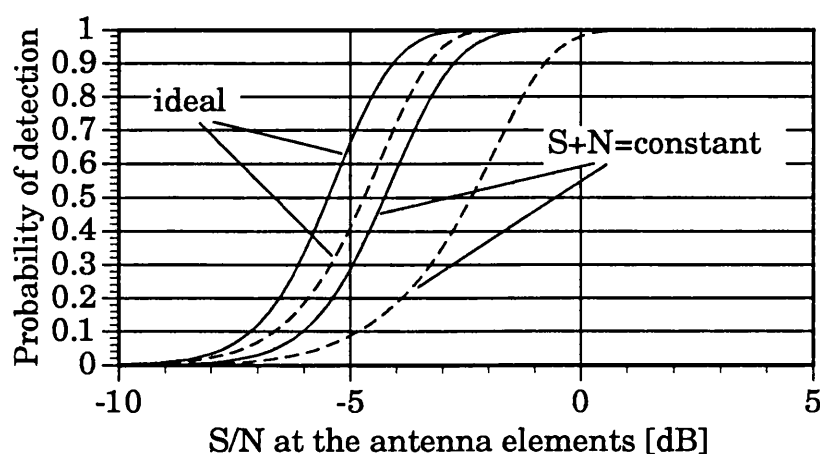
Number of integra-
ted samples:

n=100

False-alarm rate:

$P_{fa} = 10^{-6}$

Up to 1 'don't know'

———— $\phi = 0°$

‒ ‒ ‒ ‒· $\phi = 22.5°$

Figure 5-20a
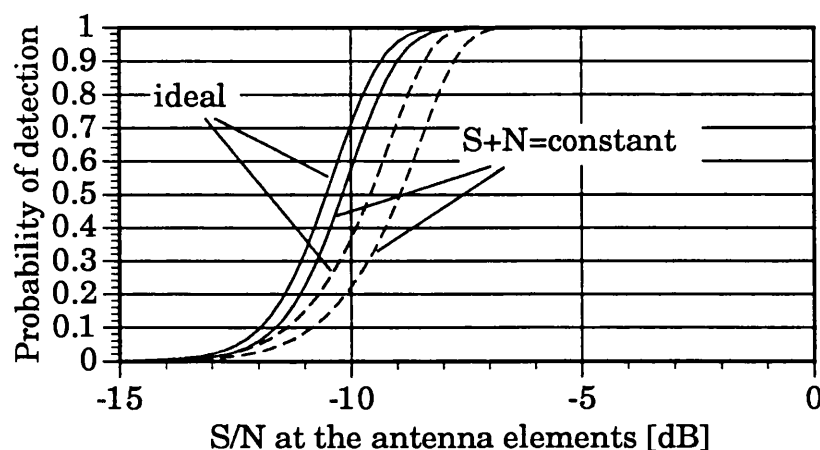
Number of integra-
ted samples:

n=1000

False-alarm rate:

$P_{fa} = 10^{-6}$

Up to 1 'don't know'

———— $\phi = 0°$

‒ ‒ ‒ ‒· $\phi = 22.5°$

Figure 5-20b
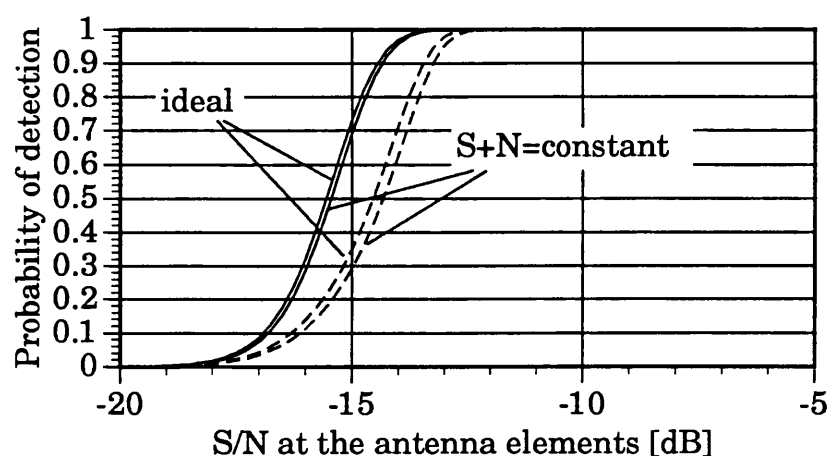
Number of integra-
ted samples:

n=10000

False-alarm rate:

$P_{fa} = 10^{-6}$

Up to 1 'don't know'

———— $\phi = 0°$

‒ ‒ ‒ ‒· $\phi = 22.5°$

Figure 5-20c

*Figure 5-20: Probability of detection as a function of the signal-to-noise ratio
at the antenna elements with ideal threshold and threshold set
for constant signal plus noise power at the output of the AGC
stages for different numbers of integrated samples*

is possible to evaluate the probability of detection as a function of the signal-to-noise ratio at the antenna elements. Fig. 5-20 and Fig. 5-21 show the relevant curves for various numbers of integrated samples and different numbers of maximum permitted 'don't know' states per measurement. Two sets of curves are contained in each diagram:

(1)   The first set of curves was obtained for fixed gain of the IF amplifiers and known noise level at the antenna elements. This means that the noise power present at the bipolar threshold detectors is exactly known and the probability of detection can be calculated using eqn. (5.44), derived in the previous section. This will be referred to as the 'ideal' case or type I.

(2)   The second set of curves was calculated for constant signal plus noise power at the output of the AGC stages. The threshold is set for the highest possible noise level at the bipolar detectors, which is generated in the absence of a signal. As soon as a signal is present, the noise level will be reduced due to the action of the AGC. The curves were calculated with eqns. (5.47a-c) and the modified version of eqn. (5.44). This will be referred to as the 'signal plus noise is constant' situation or type II.

Inspection of Fig. 5-20 shows that the difference between both sets of curves for the two threshold settings becomes smaller for larger number of samples, and is virtually negligible for 10000 or more samples. This can be explained by the fact that the minimum signal-to-noise ratio at the antenna elements is determined by the number of integrated samples. The more samples are integrated, the lower is the minimum required signal-to-noise ratio, and the smaller is the signal component in the output of the AGC stages. Low signal levels cause hardly any reduction of the IF amplifier gain and the system is working close to ideal conditions, i.e. the threshold is set for a known noise level. Comparison of Figs. 5-20a-c also shows that the difference between the curves of type I and II for ϕ = 22.5 ° is always larger than for ϕ = 0°.

The effect of the number of permitted 'don't know' states per measurement is investigated in Fig. 5-21. For up to one 'don't know', the difference in signal-to-noise ratio between the curves of type I and II is
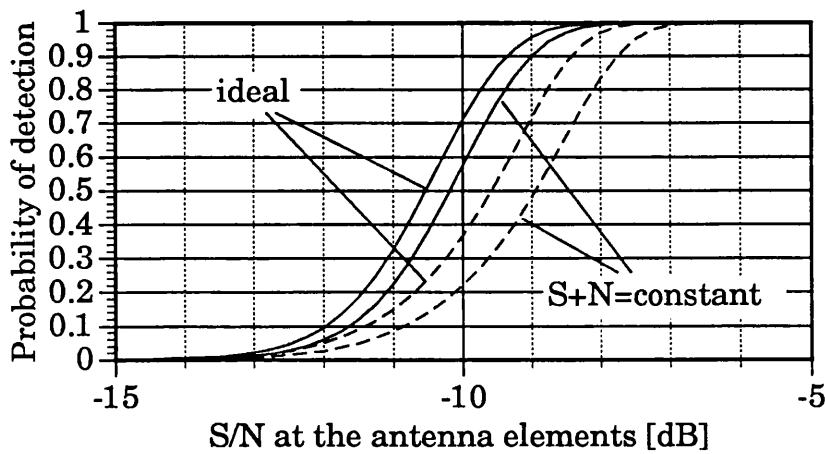
Figure 5-21a

Up to 1 'don't know'

Number of integra-
ted samples:

n=1000

False-alarm rate:

$P_{fa} = 10^{-6}$

———— $\phi = 0°$

– – – – · $\phi = 22.5°$
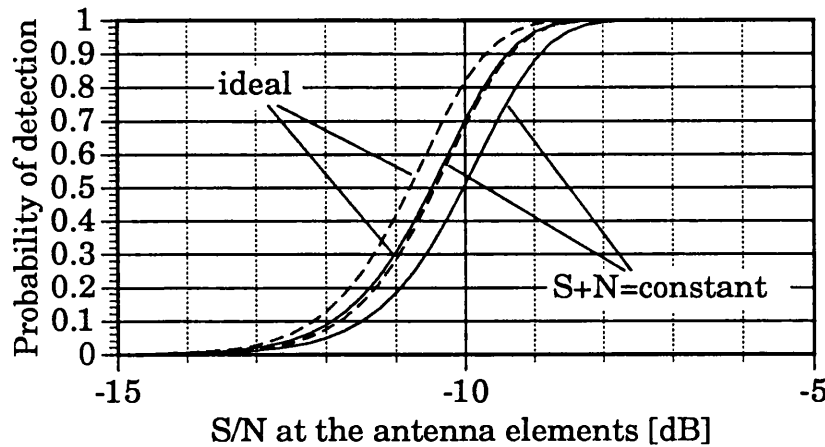
Figure 5-21b

Up to 2 'don't knows'

Number of integra-
ted samples:
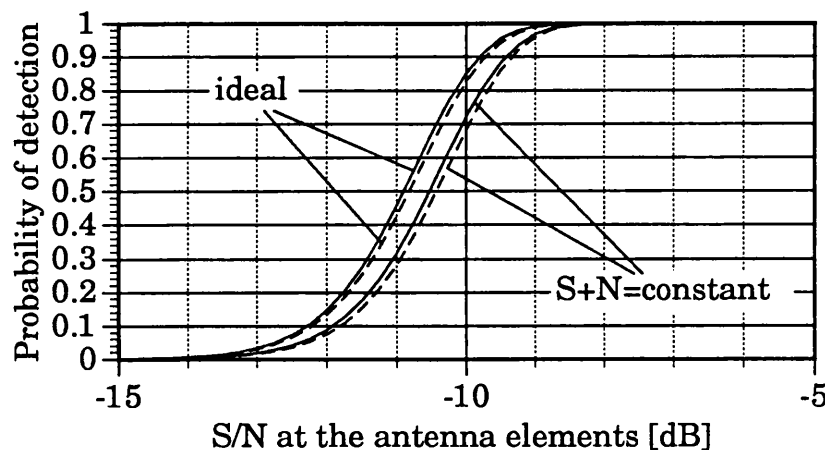
n=1000

False-alarm rate:

$P_{fa} = 10^{-6}$

———— $\phi = 0°$

– – – – · $\phi = 22.5°$

Figure 5-21c

Up to 3 'don't know's

Number of integra-
ted samples:

n=1000

False-alarm rate:

$P_{fa} = 10^{-6}$

———— $\phi = 0°$

– – – – · $\phi = 22.5°$

Figure 5-21: *Probability of detection as a function of the signal-to-noise ratio*
at the antenna elements with ideal threshold and threshold set
for constant signal plus noise power at the output of the AGC
stages for different numbers of maximum permitted 'don't
know' states

negligible for large number of samples (recall Fig. 5-20c), since the dependence of the probability of detection on the angle of arrival, i.e. the difference between the traces for $\phi = 0\,°$ and $\phi = 22.5\,°$, dominates. However, for up to three 'don't knows', the probability of detection is only marginally dependent on the angle of arrival and the difference in signal-to-noise ratio between curves of type I and II can clearly be identified (see Fig. 5-21c). Again, the difference is determined by the number of integrated samples, and is just about 0.4 dB for 1000 samples.

In summary, it can be said that setting the threshold of the symmetric-pair DF system under the assumption that the signal plus noise power at the output of the AGC stages is constant introduces only a small additional loss compared to the ideal case. This loss decreases for growing number of samples and is of the order of 0.5 dB or less for a minimum of 1000 integrated samples and up to one, two or three permitted 'don't know' states. An advantage of this method is that neither the instantaneous gain of the IF amplifiers nor the actual noise level at the antenna elements has to be known for its successful application.

5.3.2. Threshold derived from estimated noise power

The second method to determine the threshold setting tackles the problem from a different point of view. It is assumed that the noise level at the antenna elements and the gain of the IF amplifiers remain unchanged over the time it takes to calibrate the threshold and to take a complete bearing, i.e. stable conditions. The objective is to estimate the actual value of the noise power by sampling the cross-correlator output and set the threshold accordingly.

Since just a finite number of samples can be taken, the estimated noise power will deviate from the correct value and a certain variance or standard deviation can therefore be assigned to it. The larger the number of samples taken, the more precise will the estimate be. The relevant mathematics are fairly well established for the case of Gaussian distributed variates. Since the pdf of the cross-correlator output is non Gaussian, as shown in section 5.1, the concept of Gaussian batches has to be introduced [Leo 89]. A Gaussian batch is the number of samples that

must be averaged in order to obtain a Gaussian distribution by virtue of the central limit theorem. As discussed in section 5.2.1, groups of 100 samples of the cross-correlator output satisfy this condition.

The knowledge of the (in this case estimated) noise power at the antenna elements is necessary to set the threshold and hence the false-alarm rate correctly. As the false-alarm rate is a highly non linear function of the threshold, it is not sufficient to calculate a mean threshold value and insert it into the equations for the false-alarm probability. The mean false-alarm rate has therefore to be evaluated which is defined as follows:

$$
\overline{P}_{fa} = E\left[g_{fa}\left(\hat{m}_{n_G z_{\Sigma}}, \hat{m}_{n_G z_{\Delta}}, S^2_{n_G z_{\Sigma}}, S^2_{n_G z_{\Delta}}; a, n, n_G \right) \right]
$$

$$
= \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} g_{fa}\left(\dot{m}_{n_G z_{\Sigma}}, \dot{m}_{n_G z_{\Delta}}, s^2_{n_G z_{\Sigma}}, s^2_{n_G z_{\Delta}}; a, n \right)
$$

$$
p\left(\dot{m}_{n_G z_{\Sigma}}, \dot{m}_{n_G z_{\Delta}}, s^2_{n_G z_{\Sigma}}, s^2_{n_G z_{\Delta}}; n_G \right) d\dot{m}_{n_G z_{\Sigma}} d\dot{m}_{n_G z_{\Delta}} ds^2_{n_G z_{\Sigma}} ds^2_{n_G z_{\Delta}} \quad ...(5.48)
$$

where a is the threshold setting in standard deviations of the cross-correlator output and n the number of integrated samples. $\hat{m}_{n_G z_{\Sigma/\Delta}}$ and $S^2_{n_G z_{\Sigma/\Delta}}$ are the unbiased estimators of the mean and variance of the Σ and Δ channel:

$$
\hat{m} = \overline{Z} = \frac{1}{n_G}\left(Z_1 + Z_2 + ... + Z_{n_G} \right)
$$

$$...(5.49a)$$

$$
S^2 = \frac{1}{n_G - 1}\sum_{k=1}^{n_G}\left(Z_k - \overline{Z} \right)^2
$$

$$...(5.49b)$$

where n_G is the number of Gaussian batch samples. $\dot{m}_{n_G z_{\Sigma/\Delta}}$ and $s^2_{n_G z_{\Sigma/\Delta}}$ in eqn. (5.48) are specific values of the random variables $\hat{m}_{n_G z_{\Sigma/\Delta}}$ and $S^2_{n_G z_{\Sigma/\Delta}}$.

The above formulation of the problem would allow us to calculate the noise at the antenna elements even in the presence of a signal, i.e. the additional noise components at the cross-correlator output due to the signal x noise cross products are considered. However, it is impossible to solve the relevant integrals in a closed form and multi dimensional numerical integration would be necessary to evaluate them. Since the

outcome of the calculation is the mean false-alarm probability, whose value ranges over several orders of magnitude, rounding errors and problems with the required precision of the numerical algorithm are bound to falsify the results considerably. No attempt has therefore been made to implement the above problem on a computer.

Nevertheless, it is still possible to assess the usefulness of the sampling approach for setting the threshold of the bipolar detectors. The calculations can be simplified by assuming that no signal is present, when the threshold is calibrated. As a result, the output of the cross-correlator does not contain any signal x noise cross terms and can easily be related to the noise power at the antenna elements. The presence of a signal for some of the time during the calibration would contribute to the noise level at the correlator output and a higher threshold than necessary would be the result. A block diagram that illustrates how to obtain the threshold setting for a desired mean false-alarm rate from the sampled output of a cross-correlator is given in Fig. 5-22. The Σ signal has been chosen to estimate the threshold, but it should be noted that either of the two channels could be used. With the above simplifications, the mean false-alarm probability becomes:

$$\overline{P}_{fa} = \int_{-\infty}^{\infty} g_{fa}\left(s^2_{n_G z_\Sigma}, a, n\right) p\left(s^2_{n_G z_\Sigma}; n_G\right) ds^2_{n_G z_\Sigma}$$

...(5.50)

The function $g_{fa}\left(s^2_{n_G z_\Sigma}, a, n\right)$ is the expression for the exact false-alarm rate given in eqn. (5.41) and eqn. (5.42a-c) and the relevant equations for the variances of the Σ and Δ signal after summation of n independent samples from appendix 7 can also be used. Only the formula for the threshold must be modified, which is now defined in terms of the estimated variance of the Σ cross-correlator output:

$$T = a\sqrt{n\ s^2_{n_G z_\Sigma}}$$

...(5.51)

where n is the number of integrated samples and $s^2_{n_G z_\Sigma}$ a possible value of the estimated variance of the correlator output signal. The pdf of the variate $S^2_{n_G z_\Sigma}$ is known in its normalised form, which is a chi-square distribution with n_G - 1 degrees of freedom [Pap 91]:

$$p_{\frac{n_G-1}{\sigma_{z_\Sigma}^2}s_{n_G z_\Sigma}^2}\left(\frac{n_G-1}{\sigma_{z_\Sigma}^2}s_{n_G z_\Sigma}^2\right) = \chi^2(n_G-1)$$

...(5.52)

Note that the subscript gives the random variable under consideration to avoid confusion. $\sigma_{z_\Sigma}^2$ is the variance of the Σ correlator output, which is given in eqn. (A7.9c) with n=1 and A=0:

$$\sigma_{z_\Sigma}^2 = \frac{\left(\sigma^2\right)^2}{16}$$

...(5.53)
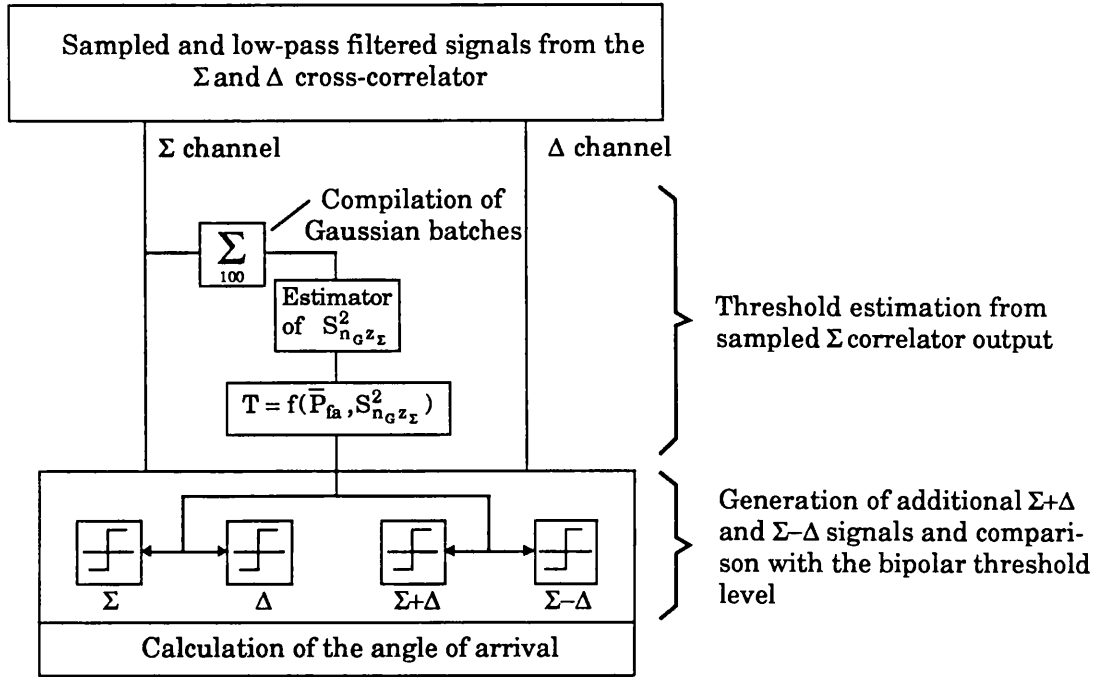
Figure 5-22: Threshold derivation from the estimated noise power

A change of variate to $S_{n_G z_\Sigma}^2$ employing the Jacobian of the transformation leads to the following expression:

$$p_{s_{n_G z_\Sigma}^2}\left(s_{n_G z_\Sigma}^2;n_G\right) = \frac{\left(\frac{(n_G-1)s_{n_G z_\Sigma}^2}{\sigma_{z_\Sigma}^2}\right)^{\left(\frac{n_G-3}{2}\right)}(n_G-1)\ \exp\left[-\frac{(n_G-1)s_{n_G z_\Sigma}^2}{2\ \sigma_{z_\Sigma}^2}\right]}{2^{\left(\frac{n_G-1}{2}\right)}\sigma_{z_\Sigma}^2\ \Gamma\left(\frac{n_G-1}{2}\right)}$$

...(5.54)

where $\Gamma(x)$ is the Gamma function. It is now possible to calculate the mean false-alarm probability by inserting the above formulae in eqn (5.50) and carrying out the required numerical integrations. The mean false-alarm rate for up to one, two or three 'don't knows' as a function of the threshold setting and the number of Gaussian batch samples n_G is given in Figs. 5-23a-c. The traces for $n_G \rightarrow \infty$ represent the case where enough samples have been added so that the estimate is the exact value of the variance at the cross-correlator output. These curves are identical with those in Fig. 5-12. It can be seen that for a given threshold setting, at least 100 Gaussian batch samples should be added to obtain a mean false-alarm rate that is reasonably close to the exact one.

The corresponding mean probability of detection is defined as follows:

$$\overline{P}_d = \int_{-\infty}^{\infty} g_d\left(s^2_{n_G z_r}, a, n\right) p\left(s^2_{n_G z_r}; n_G\right) ds^2_{n_G z_r}$$

$$...(5.55)$$

The function $g_d\left(s^2_{n_G z_r}, a, n\right)$ is defined by equations and (5.43) and (5.44), and $p\left(s^2_{n_G z_r}; n_G\right)$ is given in eqn. (5.54). The expression for the threshold in $g_d\left(s^2_{n_G z_r}, a, n\right)$ is again changed according to eqn. (5.51). Figs. 5-24a/b show a typical set of curves for the mean probability of detection.

It can be seen that the summation of only 10 Gaussian batch samples to estimate the noise power at the correlator output yields an unacceptably high increase in the required minimum signal-to-noise ratio, which is needed to obtain the desired mean false-alarm probability. This result was expected from the earlier mean false-alarm rate calculations. Further calculations yield that the increase in minimum signal-to-noise ratio is in the order of 0.3 dB or less for 100 or more Gaussian batch samples irrespective of the number of permitted 'don't knows'. Fig. 5-24b also shows that the increase in the minimum signal-to-noise ratio is slightly worse for $\phi = 22.5°$, i.e. when probability of detection is lower due to the signal's angle of arrival.

The mean probability of detection is therefore very much dependent on the number of integrated Gaussian batch samples. The additional loss in signal-to-noise ratio due to this threshold setting technique is only

Figure 5-23a

Up to 1 'don't know'

Threshold:

$T = a\sqrt{n\ s^2_{n_G z_r}}$

Figure 5-23b

Up to 2 'don't knows'

Threshold:

$T = a\sqrt{n\ s^2_{n_G z_r}}$

Figure 5-23c

Up to 3 'don't knows'

Threshold:

$T = a\sqrt{n\ s^2_{n_G z_r}}$
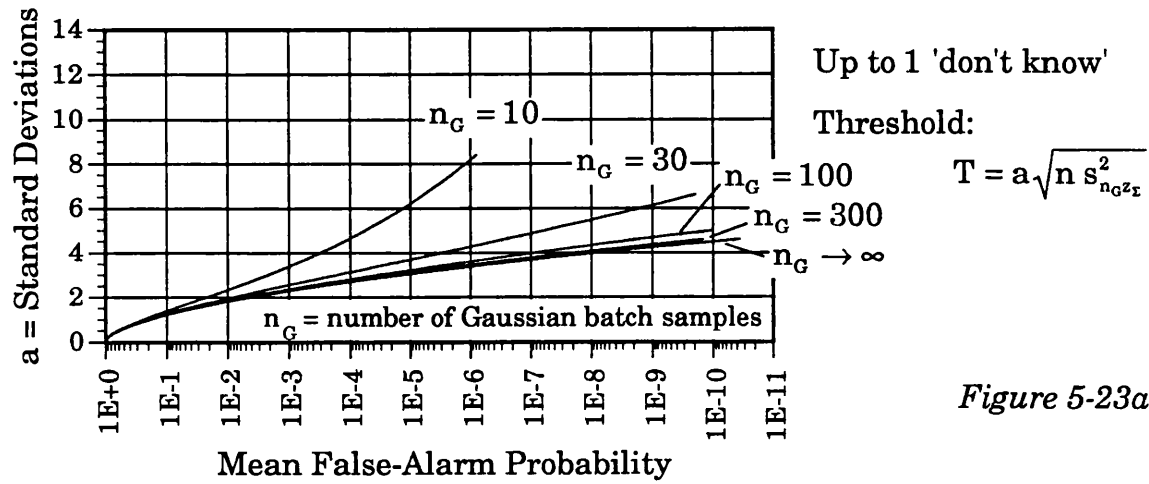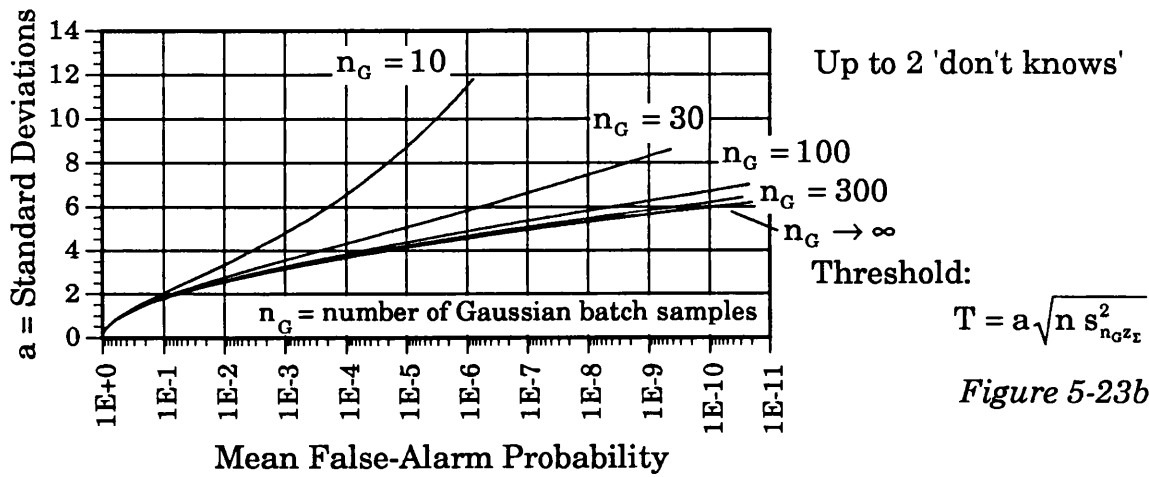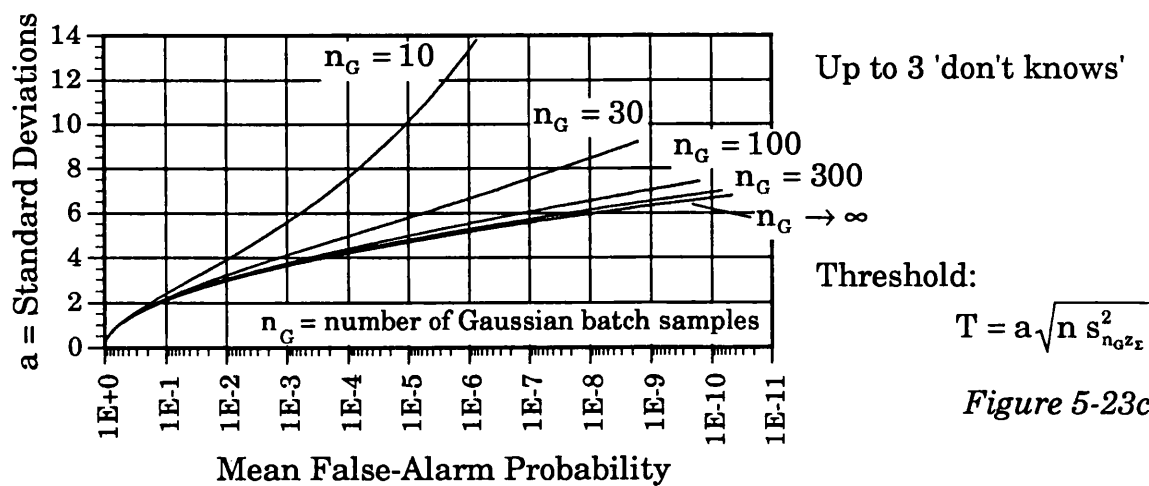
Figure 5-23: *Mean false-alarm probability as a function of threshold setting and number of Gaussian batch samples for up to 1, 2 or 3 'don't know' events per measurement. (The threshold is given in standard deviations of the correlator output.)*

Up to 1 'don't know'

Mean false-alarm
rate: $\overline{P}_{fa} = 10^{-6}$

Phase difference:

$$\phi = 0°$$

Figure 5-24a

Up to 1 'don't know'

Mean false-alarm
rate: $\overline{P}_{fa} = 10^{-6}$

Phase difference:

$$\phi = 22.5°$$
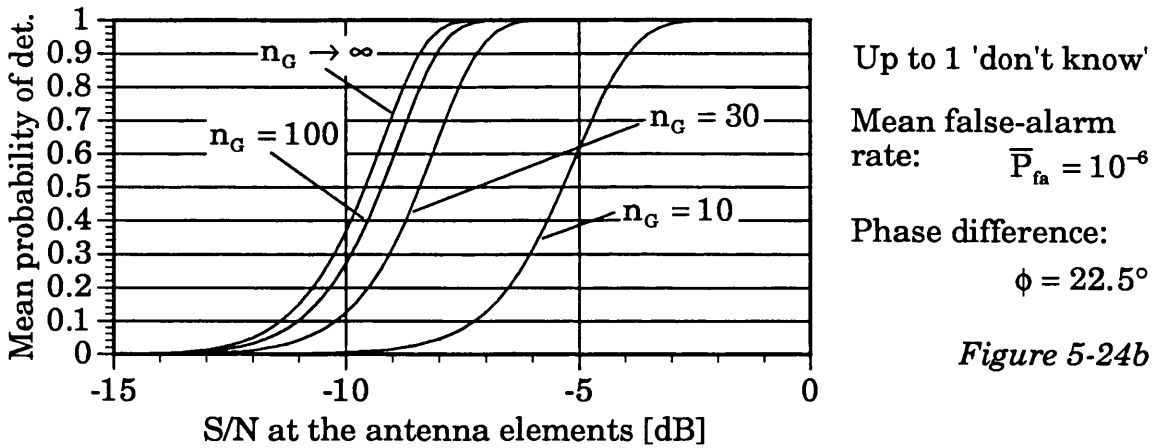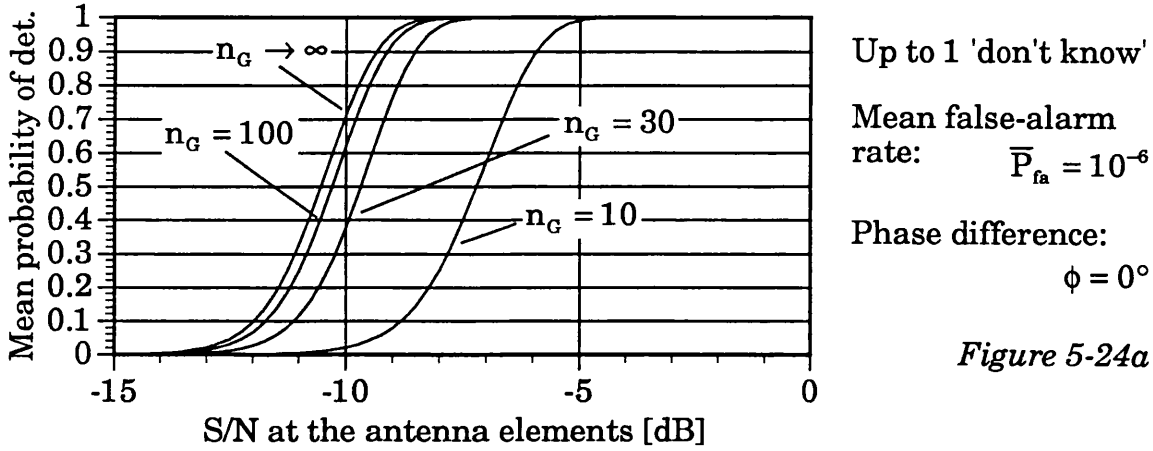
Figure 5-24b

Figure 5-24: Mean probability of detection as a function of the signal-to-noise ratio at the antenna elements for different numbers of Gaussian batch samples n_G.

tolerable if more than 100 Gaussian batch samples are added, which is equivalent to the summation of 10000 independent samples at the cross-correlator output.

5.3.3. Comparison of the two methods

In order to compare the different threshold setting techniques, it is convenient to define a loss factor L. The loss factor for a given false-alarm rate and number of integrated samples is the difference in signal-to-noise ratio between the curves for the approximated and exact threshold for a certain probability of detection:

$$L[dB] = \frac{S}{N}\Bigg|_{P_d=x,approx.\,T} [dB] - \frac{S}{N}\Bigg|_{P_d=x,exact\,T} [dB]$$

...(5.56)

'Approximated threshold' means that the threshold is either set under the assumption that the signal plus noise power is constant at the output of the AGC stage, or it is set for the estimated actual noise power at the correlator output. The signal-to-noise ratio that gives the desired probability of detection was evaluated using the equations for the calculation of P_d and $\overline{P}_d$ (e.g. eqn. (5.43)., eqn. (5.50)) and by interpolating between sets of data. The loss factor for up to one, two or three 'don't knows' and for $\phi = 0°$ or $\phi = 22.5°$ is given in Figs. 5-25a-f. The main difference between the two threshold setting techniques is that the loss factor becomes smaller for an increase of the number of samples when the threshold is set for constant signal plus noise power, whereas for the estimation approach, the loss factor is virtually constant and only dependent on the number of Gaussian batch samples n_G. It can be seen in Figs. 5-25a-f that for more than 1000 integrated samples and $n_G = 100$, both techniques give a loss factor of less than 0.5 dB. It is also apparent that the larger the number of permitted 'don't knows', the smaller is the difference between the loss factors for $\phi = 0°$ and $\phi = 22.5°$ for a certain threshold setting method. This is due to the fact that the difference in signal-to-noise ratio for the two values of ϕ becomes smaller for wider 'don't know' zones (see Figs. 5-16a-c).

As a short summary of the previous discussion, the main properties of the two threshold setting techniques are compared below:

(1) Threshold setting for constant signal plus noise power at the AGC output is easier to implement, adjusts itself inherently and does not need constant recalibration, in contrast to the technique based on the estimated noise power at the correlator output.

(2) At least 100 Gaussian batch samples must be used in the estimation method to obtain a similar loss factor for both techniques for more than 1000 integrated samples. This can prove to be rather time consuming, dependent on how often the threshold must be recalibrated.
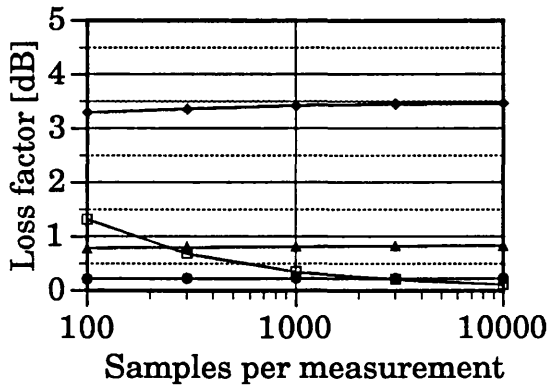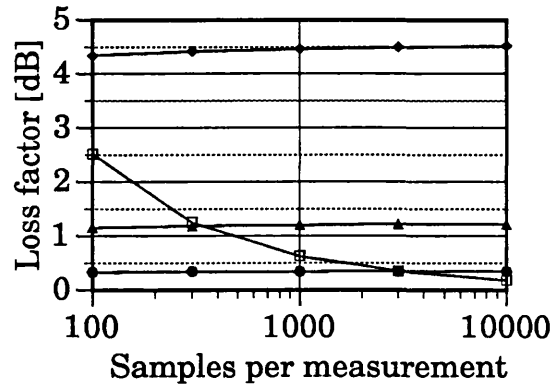
(a) Up to 1 'don't know', $\phi = 0°$ (b) Up to 1 'don't know', $\phi = 22.5°$
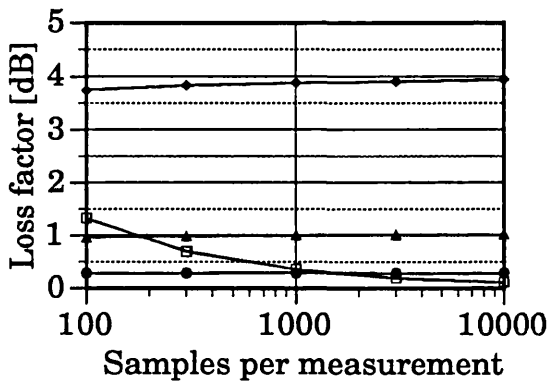
(c) Up to 2 'don't knows', $\phi = 22.5°$ (d) Up to 2 'don't knows', $\phi = 0°$
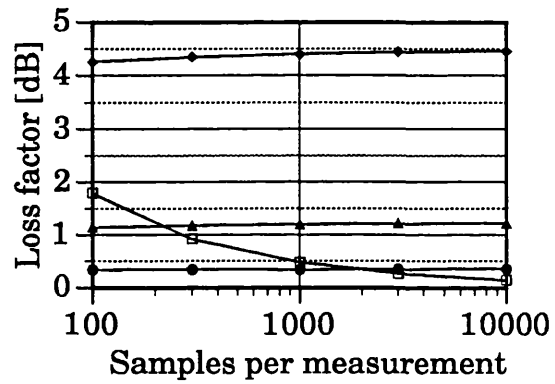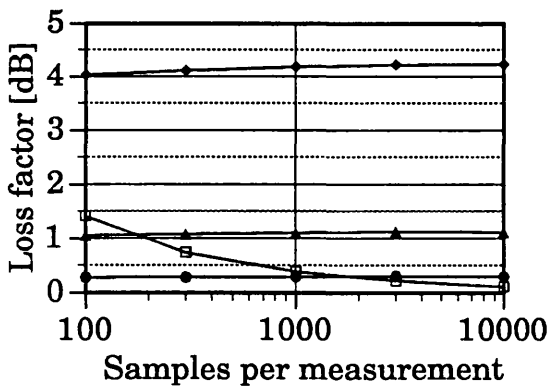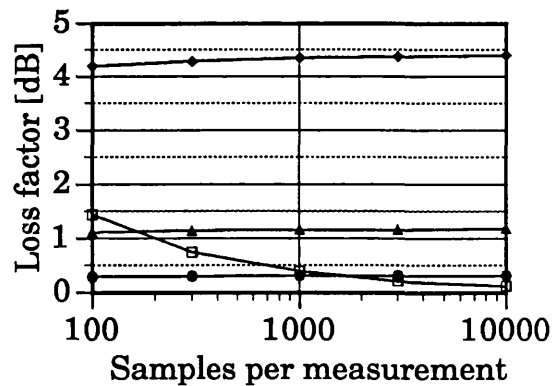
(e) Up to 3 'don't knows', $\phi = 0°$ (f) Up to 3 'don't knows', $\phi = 22.5°$

Legend: ─□─□─ S + N = constant

─♦─♦─ $n_G = 10$ ─▲─▲─ $n_G = 30$ ─●─●─ $n_G = 100$

In all graphs: $P_{fa} = 10^{-6}$ and $P_d = 0.9$

Figure 5-25:   Loss factor for up to 1, 2 or 3 'don't knows' for different threshold setting techniques

(3) The analysis for the threshold derived from the estimated noise power is based on the assumption that no signal is present when the measurements for the threshold calibration are carried out, which is in practice hard to achieve. The presence of a signal during the calibration would cause too high a threshold level.

In view of the above arguments, it is evident that the threshold setting technique that is based on the assumption of constant signal plus noise power at the output of the AGC IF amplifier should be implemented in the symmetric-pair DF system.

6 FURTHER APPLICATIONS OF THE SYMMETRIC-PAIR PRINCIPLE

6.1. SYMMETRIC-PAIRS IN COMMUNICATIONS

The previous three chapters concentrated on the application of symmetric-pair arrays in direction finding. However, it is also possible to use this type of array for communications. An important property of communications antenna arrays is their ability to accomplish agile beam forming in order to adapt to often quickly varying demands. Due to the binary character of the output of symmetric-pairs and the use of equal and opposite phase shifts for beam steering, which halves the number of control parameters to be calculated, symmetric-pair arrays have obvious potential for fast and simple beam forming techniques. The capabilities of symmetric-pair antenna arrays in this field will be investigated in the following section.

Another aspect of communications will be studied in a later section; the idea of an antenna array with isotropic reception gain. This type of array would be useful for mobile communications, where signals with sometimes rapidly varying field strengths from constantly changing directions must be processed. An antenna array with omnidirectional or isotropic reception gain would therefore improve the link budget between the base station and the moving vehicle.

6.2. BEAM FORMING WITH SYMMETRIC-PAIRS

6.2.1. The principle of beam forming with symmetric-pairs

To form a beam with a symmetric-pair array, it is necessary to steer the beams of all pairs in the wanted direction. This is easily possible,

because each symmetric-pair can be treated as an independent subarray, whose beam pattern can be steered by means of equal and opposite phase shifts at the elements (see chapter 2). This kind of beam steering has already been used in land-mobile satcom arrays [Bre 91].

Since the output signals of all pairs are essentially in phase or anti-phase to each other, a beam can be formed by reversing the signs of all signals which are in anti-phase. Reversing the sign of the output signals is in effect equivalent to using ±180° phase shifts at the elements of the relevant pairs. This method of beam-forming yields a maximum deviation of the signal vectors from the ideal co-phasal summation of ±90°. It is reasonable to expect that the phase error of a large array with randomly distributed pairs is uniformly distributed over this range. Hence, the average voltage loss for the Σ or Δ output signal compared to fully co-phasal summation can be defined as:

$$\frac{1}{\pi} \int_{-\pi/2}^{\pi/2} \cos \phi \; d\phi = \frac{2}{\pi} \qquad \qquad ...(6.1)$$

which is equivalent to 3.92 dB. If the signals of the two orthogonal Σ and Δ channels are combined as shown in Fig. 6-1, the average aggregate power is doubled and the total loss reduced to 0.91 dB. (It should be noted that the 90° phase offset between the Σ and Δ signals has to be compensated). Thus it is possible to perform beam forming with symmetric-pair arrays using standard components like 180° hybrids and sign inverters which are broadband and readily available.

The above method is very simple and effective, but has an inherent residual loss of 0.91 dB compared to fully co-phasal summation. To avoid this problem other beam forming methods which

Figure 6-1: Simple Σ and Δ beam forming with a symmetric-pair

do not have these restrictions can be used. In the next two sections, possible alternatives will be introduced and discussed.

6.2.2. Beam forming with phase shifters

As outlined in the previous section, it is necessary to generate equal and opposite phase shifts at the elements of a symmetric-pair in order to steer its beam. The simplest way of doing this is to use variable phase shifters at each element and to add both signals with a power combiner. The output signal of each pair is given by (see section 2.2.2.):

$$V_{eo} = \sqrt{2} V \cos(\phi + \Delta\phi) \, e^{j\omega t} \qquad\qquad ...(6.2)$$

Examining eqn. (6.2), it is obvious that by applying the correct phase shift at the elements it is possible to compensate the phase lead and lag at the pair elements and achieve fully co-phasal summation, which yields maximum output signal. With an n-bit incremental phase choice for each element of a symmetric-pair, the voltage loss L relative to perfect co-phasal summation is equal to:

$$L = \frac{2^n}{2\pi} \int_{-\pi/2^n}^{\pi/2^n} \cos\phi \, d\phi = \frac{\sin(\pi/2^n)}{\pi/2^n} \qquad\qquad ...(6.3)$$

Table 6-1 shows the loss for different numbers of discrete phase steps.

n [2^n: number of discrete phase steps]	1	2	3	4
Loss [eqn. (6.3)]	0.637	0.900	0.974	0.994
Loss dB	3.92	0.91	0.22	0.056
Maximum phase deviation from the condition for co-phasal summation	±90°	±45°	±22.5°	±11.25°

Table 6-1: Loss recovery by use of discrete compensating phase shifts

Note that n=2 defines the phase to the appropriate quadrant, just like the signs of the Σ and Δ output. The corresponding mean reduction in array gain is therefore identical. It is evident that the loss is very small for 3 or 4 bit phase shifters. If vector modulators are used as phase shifters, this arrangement can perform beam forming over a wide frequency range but has an inherent insertion loss of 6 dB due to the type of phase shifter. With switched line phase shifters, beam forming is usually restricted to narrow band applications, and only a small number of line segments per phase shifter should be used to keep the insertion loss low.

To confirm the beam steering capabilities of symmetric-pair arrays, a 2D prototype with 3 pairs has been built. The antenna elements were monopoles mounted on an aluminium ground plane. The prototype array is shown in Fig. 6-2. The amplitude pattern of the array with equal phasing for all six elements is shown in Fig. 6-3. It can be seen that there is very good accordance between the theoretical and the measured pattern. As a simple demonstration of beam forming, manually controlled line stretchers were used in the experiment to provide equal and opposite phase shifts for each pair. Fig. 6-4 shows the result for a beam steered to -30°. Its mean sidelobe level of -7.43 dB is in good agreement with the calculated value of -7.78 dB for a 6 element array. The high backlobe at 120° is a property of the array configuration, not an artefact of the specific manner of phase steering. The differences between measurement and
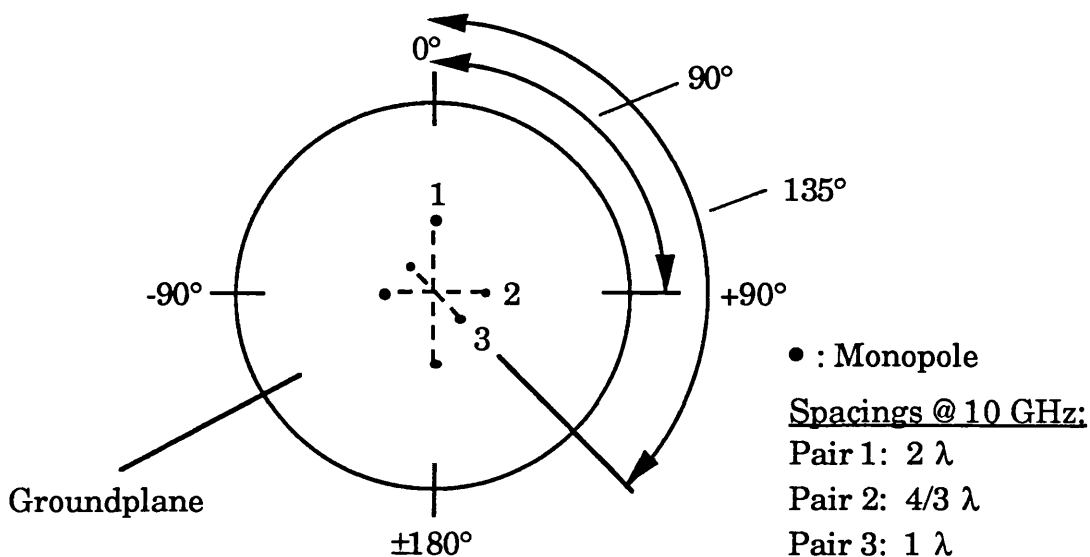
Figure 6-2: Configuration of the prototype array

: calculated pattern   - - - - - -: measured pattern
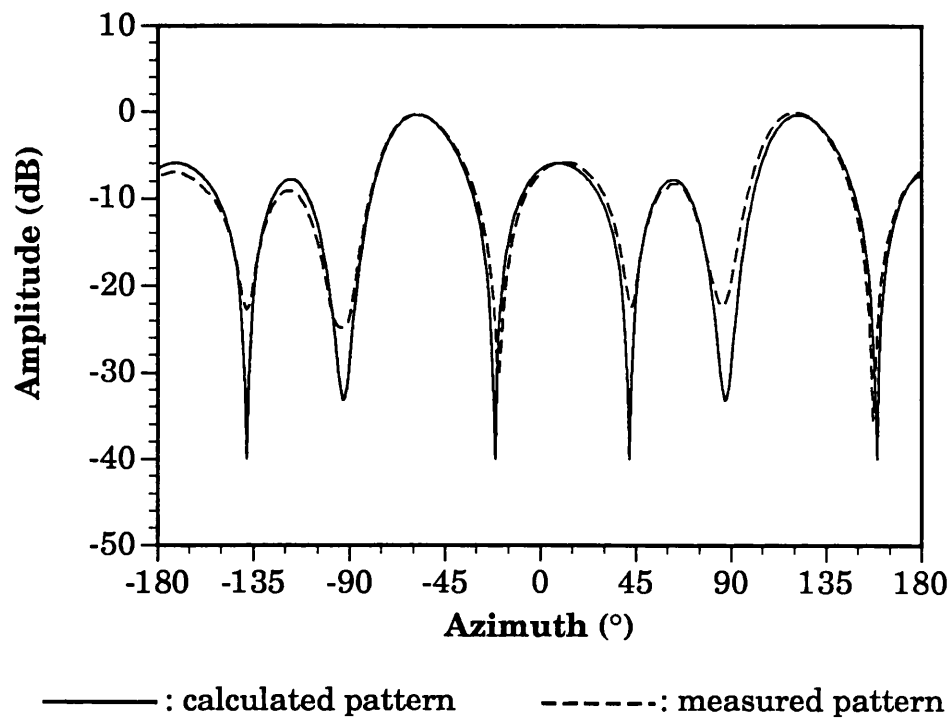
Figure 6-3:   Amplitude response of the 3 pair prototype array with equal phasing at the elements

: calculated pattern   - - - - - -: measured pattern
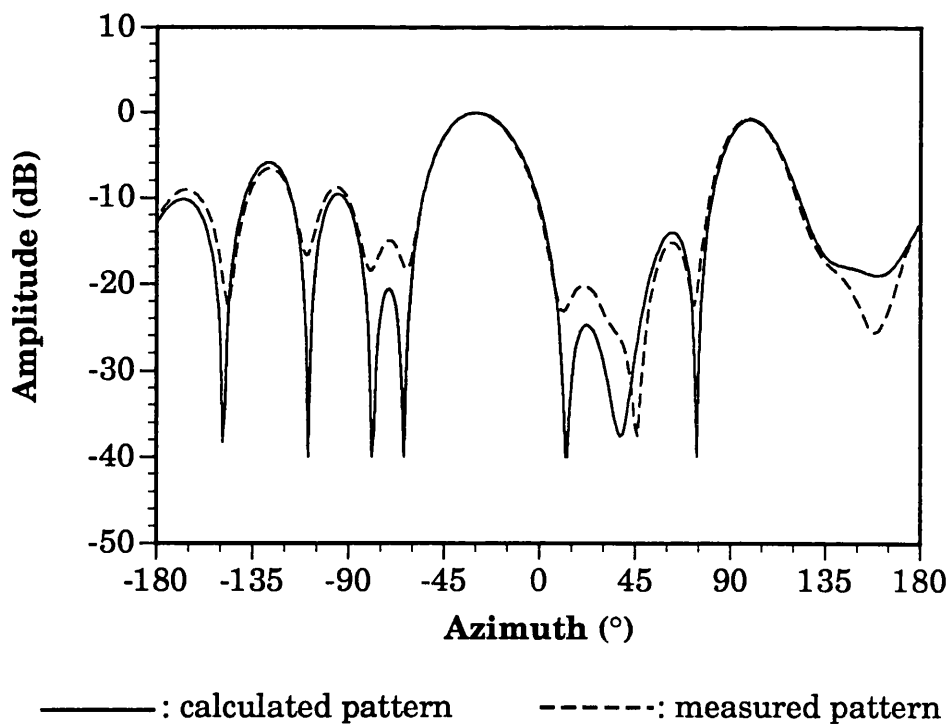
Figure 6-4:   Amplitude response of the 3 pair prototype array with a maximum steered to -30°

calculation for rather low sidelobe levels are due to mutual coupling between the elements and obscuration effects caused by the neighbouring elements.

6.2.3. Beam forming using a bus-bar system and amplitude weighting

A further development of the very simple system introduced in section 6.2.1. is to use bus-bars combined with amplitude weighting of the Σ and Δ signals. This scheme is a modified version of a vector modulator. Following the block diagram in Fig. 6-5 bottom up, it can be seen that the required phase shifts are obtained by generating two quadrature signals, which are weighted in proportion to the sine and cosine of the required phase shift. The 180° hybrid for the summing process at the antenna elements finally provides the equal and opposite phase shifts for each pair. Sign inverters are used to realise the negative amplitude weights. Each Σ output of a symmetric-pair is switched to the +X or -X bus-bar and the Δ signals to the +Y or -Y bus-bar respectively. The output signal of a bus-bar system for one pair is given by (see Fig. 6-5):

$$V_{out} = \frac{V}{\sqrt{2}} \cos(\phi - \Delta\phi)\, e^{j\omega t}$$

...(6.4)

Eqn (6.4) clearly shows that weighting of the Σ and Δ signals of the hybrid with cos $\Delta\phi$ and sin $\Delta\phi$ yields maximum output signal.

If the amplitude weights are stepped in equal increments of ϕ, the mean reduction of array gain compared to fully co-phasal summation is equivalent to the loss of a symmetric-pair array with digitised phase shifters at each element as described in the previous section. On top of that loss, which is dependent on the step size of the weights, is a bus-bar inherent insertion loss of 6 dB. (This assumes that external noise exceeds the noise generated by the system.) Fig. 6-5 illustrates how this loss arises.

Without any amplitude weighting, the bus-bar system is a generalised form of the beam former introduced in section 6.2.1 and matches the phases of the elements of each pair to the nearest quadrant which entails
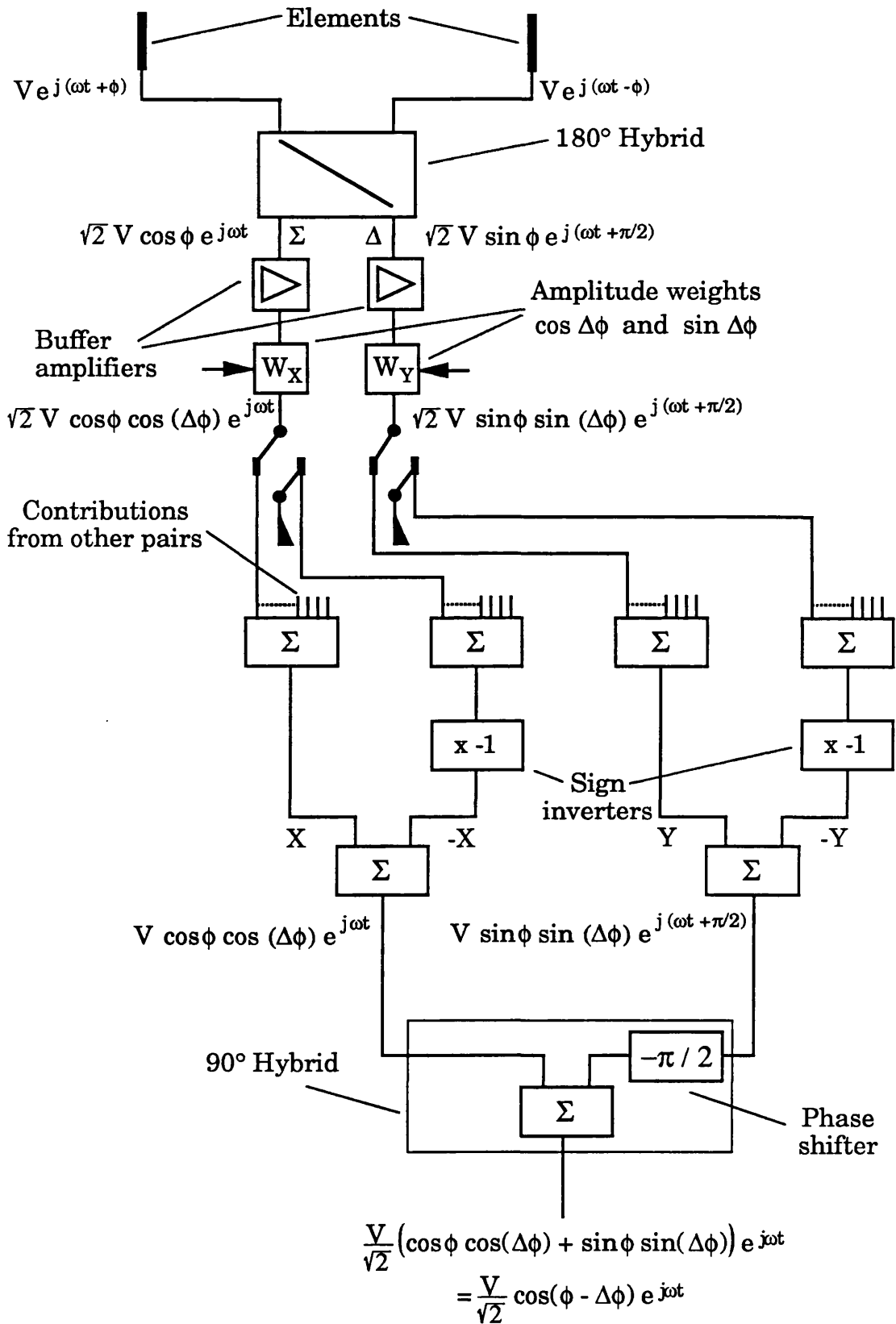
Figure 6-5: Beam forming using a bus-bar system with amplitude weights

a mean reduction in array gain of 0.91 dB. An advantage of this version of the bus-bar system is that 'off the shelf components' can be used for its realisation. All bus-bar beam formers are broad band, whereas in the case of beam forming with phase shifters the bandwidth of the system depends on the type of phase shifters used.

6.2.4. Comparison of the beam forming methods

To allow simple comparison of the two beam forming methods, beam forming networks that match the signal vectors to the nearest quadrant have been chosen. Their corresponding average loss compared to fully co-phasal summation is 0.91 dB. The beam former with equal and opposite phases shifts requires therefore two bit phase shifters, and no amplitude weighting is used in the bus-bar system.

Table 6-2 gives a list of the hardware required for a system with p pairs. The number of PIN diodes needed in the case of switched line phase shifters at the elements is based on the assumption that 4 diodes per bit (phase shift section) per element are needed. With the bus-bar system, 4 PIN diodes are required to switch each of the hybrid output signals to the appropriate bus-bar, since the unused bus-bar input must always be terminated with a load. Examining Table 6-2, one can see that the bus-bar method requires less hardware for large arrays. For small arrays, phase

Item	Phase shifts at elements	Bus-bar system
PIN diodes	16 p	8 p
Power combiners	1 with 2p inputs	4 with p inputs 2 with 2 inputs
180° phase shifters	2 p	2
90° phase shifters	2 p	-
180° hybrids	-	p
90° hybrids	-	1
Loads	-	2 p

Table 6-2:   *Comparison of the two beam forming methods in terms of hardware*

shifters at the elements seem to be an attractive solution, because they allow more flexible use of the phased array and no additional 180° hybrids are needed.

6.3. SYMMETRIC-PAIR ARRAYS FOR GAIN WITH ISOTROPIC COVERAGE

6.3.1. Introduction to gain with isotropic coverage

To understand why conventional arrays cannot have gain with isotropic coverage (i.e. gain irrespective of the received signal's angle of arrival), it is helpful to consider first of all an arbitrary antenna array with a diameter of less than half a wavelength. In such an arrangement, the contribution from each antenna element will always increase the total array power irrespective of the angle of incidence of the signal. This is because the phases of the antenna elements are only distributed over 180°. Enlarging the array diameter increases the proportion of the plane wavefront incident on the array aperture, but it worsens the phase mismatch of the elements. In large arrays with many radiators, the phases of the elements are randomly spread over 360°. As a consequence, signal cancellations due to phase reversals occur when the contributions from all elements are added. Because of these phase mismatches, it is impossible to achieve gain with isotropic coverage with conventional arrays. The proportion of the incident power that couples into such an array varies therefore inversely with the aperture.

6.3.2. Gain with isotropic coverage with symmetric-pairs

In symmetric-pair arrays, irrespective of the size and geometry of the antenna system, all Σ output signals are co-phasal (except for the sign ambiguity), and so are the Δ outputs. After compensating the 90° phase shift between the two channels with a broadband phase shifter, all output signals of a symmetric-pair array are in phase or anti-phase to each other for all angles of incidence.

The sign ambiguity can be removed by squaring these signals, which results in co-phasal double-frequency components. Combining these coherently for N elements (using the Σ and Δ outputs from the pairs), the signal-to-noise ratio would improve by the factor of N. Since the frequency channels are spectrally distinct, the superposition from either the signals or the noise from multiple channels can be avoided. However, the input to any such system must be band-limited, to avoid intermodulation products. Additional frequency bands therefore need their own set of squarers and summing network (Fig. 6-6). To save hardware, the compensation of the phase difference between the Σ and Δ channels can be done after squaring and adding the signals from the different symmetric-pairs.
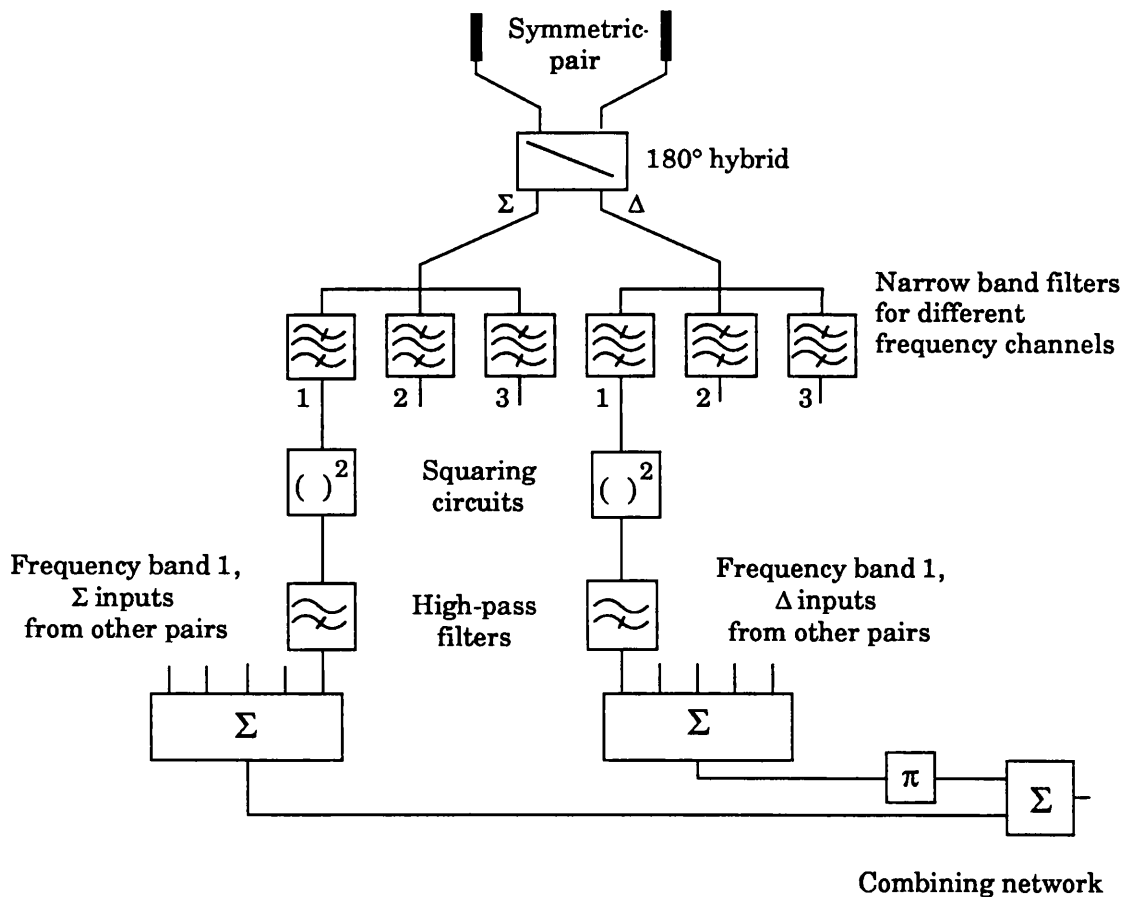
Figure 6-6: Configuration for isotropic reception-gain

Since the received signal will be squared prior to demodulation, not all types of modulation are suitable for this isotropic reception-gain principle. On-off keyed signals can clearly still be decoded, and it is also possible to recover amplitude modulated signals with small modulation indices. AM signals with large modulation indices would be distorted by

the squaring process. Any continuous frequency modulation is preserved in frequency doubling, although the frequency excursions will be doubled and it will be subject to an absolute sign ambiguity for large excursions. The information contained in a phase reversal keyed signal will be virtually lost, since the 180° inter-baud phase transitions will be reduced to cusp-like dips in amplitude.

In addition to the squared version of the signal, the squaring process produces two coherent cross products: noise x signal and signal x noise. As appendix 8 shows, their combination results in a 6 dB signal-to-noise degradation for reasonably strong input signals. There is no further loss to alternative solutions, since the output signals of the hybrid after squaring are in the form of $\cos^2 + \sin^2$, which is equivalent to optimum weighting. The net gain of such an antenna (relative to the gain of one of its elements) would thus be:

$$G = 10 \log_{10} N - 6 \text{ dB} \qquad \qquad ...(6.5)$$

According to eqn. (6.5), the number of elements required to compensate the initial loss due to the squaring process is 4 (= 2 pairs). Any additional pair will then contribute to the net array gain. When the input signal-to-noise ratio at the elements is smaller than -3 dB (see eqn. (A8.9)) the so far neglected noise x noise term becomes larger than the noise x signal and signal x noise terms. A square law detector would then further degrade this unfavourable input signal-to-noise ratio, prior to its enhancement by the array gain.

The gain with isotropic coverage is only relevant to the system's internal (thermal) noise. When external noise or interference is dominant, e.g. at low HF, the noise of the different symmetric-pairs is correlated and a logic based selection from the outputs of a set of directional beams would yield a better signal-to-noise ratio at the array output. It should be also noted that the non isotropic radiation patterns of the elements of each pair introduce a certain directivity in the combined array response. However, by using additional pairs and orientating them randomly in different directions, a combined array response which is virtually isotropic can be achieved.

6.3.3. Experimental verification of the principle of gain with isotropic coverage

Experiments have been carried out to prove the validity of the isotropic reception-gain principle. In both experiments, the signals from each pair were fed into a crystal detector. Its DC signal is the 'lower sideband' of the squared signal, which was chosen for experimental convenience. The resulting DC outputs were then amplified, summed and digitised for further data processing.

In the first experiment, the Σ output signals of the three pair prototype (see section 6.2.) were added after detection. The aim of this experiment was to demonstrate that by using just the Σ signals of three symmetric-pairs, it is already possible to avoid deep nulls in the antenna pattern. Fig. 6-7 shows the excellent agreement between simulation and measurement. The simulated pattern was calculated by adding the squared output signals of the three pairs without taking the effects of mutual coupling into account.

————: calculated pattern   − − − − − −: measured pattern
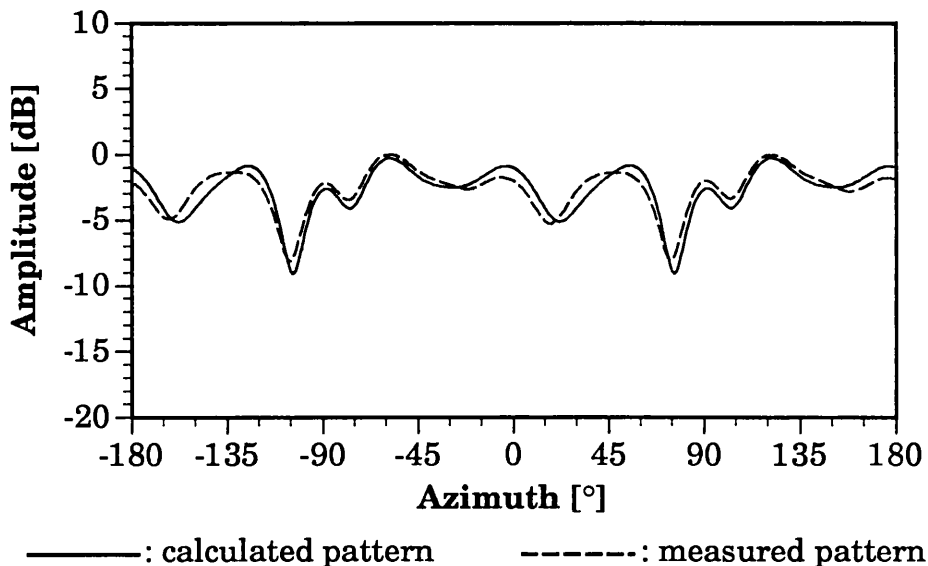
Figure 6-7: Patterns of the 3 pair prototype array after squaring and summation of the Σ signals

The second experiment demonstrates that it is feasible to achieve an omnidirectional pattern from a single pair of elements by squaring its Σ and Δ outputs and adding the two signals. Fig. 6-8 shows the measured

signal after square-law detection and summation of both hybrid outputs
(1) as well as the Σ (2) and Δ (3) outputs of the hybrid prior to summation.
Curve (1) clearly demonstrates the generation of an omnidirectional
pattern from a single pair of elements. In practice, however, mutual
coupling with other symmetric-pairs will introduce a certain amount of
direction dependence in the isotropic pattern for narrow pair spacings.
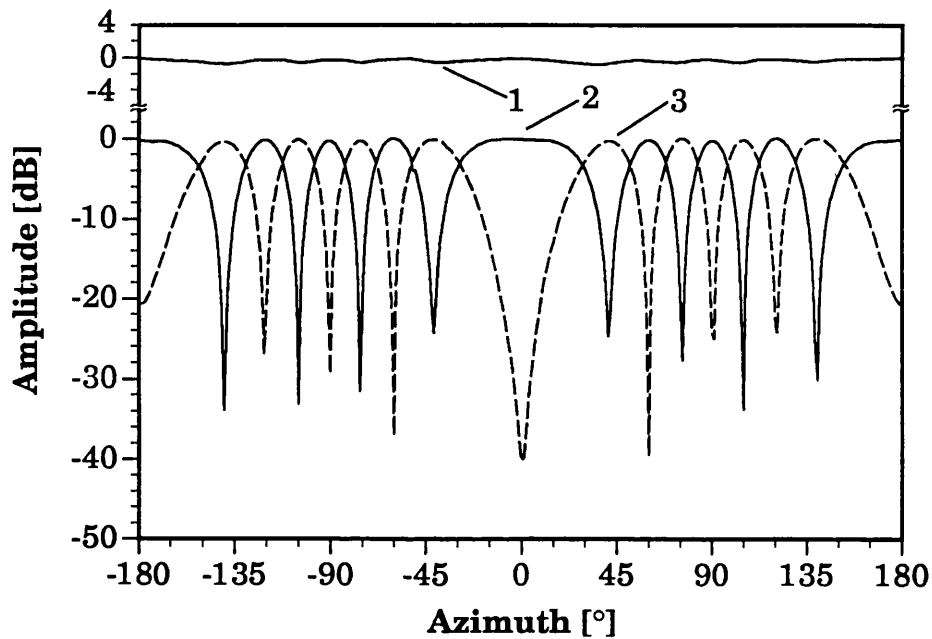
1: Signal after square law detection and summation of both hybrid outputs
2: Signal of the sum output port of the hybrid
3: Signal of the difference output port of the hybrid

*Figure 6-8:   Measured patterns of a symmetric-pair with 2 λ spacing in
isotropic mode*

7 SUMMARY AND RECOMMENDATIONS

7.1. SUMMARY

In summarising the material presented in chapters 2 to 6, the following statements can be made:

7.1.1. Symmetric-pairs in general

It has been shown that the concept of processing the signals of pairs of antenna elements is simple but powerful, and each symmetric-pair can be treated as an independent sub-array which standardises and simplifies the relevant calculations. The symmetric-pair principle is broadband and works therefore over a wide frequency range.

The output signals of symmetric-pairs are ideally just in phase or anti-phase with each other irrespective of the signal's angle of arrival. This inherently binary characteristic eases the extraction and processing of the relevant information and makes symmetric-pair arrays quite robust against phase and amplitude errors.

Depending on the applied processing technique, symmetric-pairs quantise the phase of the received signal to the nearest hemisphere, quadrant or octant.

7.1.2. Symmetric-pairs in direction finding

Fully unambiguous 3 dimensional direction finding can be performed with symmetric-pair arrays, and both azimuth and elevation angle of the

received signal can be determined. Symmetric-pair DF systems are very-wide-aperture direction finders.

The required algorithms and techniques for direction finding with symmetric-pairs have been developed, and it has been found that the most effective arrangement for direction finding is a set up with three mutually orthogonal collinear axes. The accuracy of such a symmetric-pair direction finder varies by a factor of only $\sqrt{2}$ in azimuth and elevation over the entire three dimensional space.

No highly accurate amplitude or phase measurements have to be carried out to locate the angle of arrival, since the *polarity* of the baseband signals contains the relevant information. The effects of mutual coupling are reduced due to the symmetry of the antenna configuration of the DF system.

Symmetric-pair direction finders are susceptible to multi-signal conditions. Additional signals in the same frequency band can cause a shift of the phase transition angles and lead to incorrect bearings.

Three different possibilities to enhance the system sensitivity at baseband have been examined and compared in detail. In conclusion it was decided to use the integrate and dump approach, which allows fast measurements combined with the best possible signal-to-noise ratio improvement and is entirely hardware based.

The introduction of the 'don't know' option as a third possible state for the baseband signal allows the system to cope with angles of arrival that are close to the nulls in the grating lobe pattern of a pair. As a consequence, the probability of detection of the DF system is virtually independent of the angle of arrival.

Two different strategies to set the threshold of the bipolar phase detector bank have been developed and assessed. The method where the threshold is set for constant signal plus noise power at the AGC output proved to be the better solution because it is easy to implement, adapts itself inherently to varying noise levels and does not need constant recalibration.

7.1.3. Symmetric-pairs in communications

Symmetric-pair arrays can be used for agile beam forming. Discrete equal and opposite phase shifts at the elements or a bus-bar system yield an array gain that is close to the gain achievable with fully co-phasal summation.

The use of equal and opposite phase shifts to improve the DF accuracy or the array gain in beam forming halves the number of control parameters to be calculated.

Gain with isotropic coverage for reception can be obtained with symmetric-pairs, which was not possible with any of the previously known array configurations. The experimental results have shown conclusively that it is possible to generate omnidirectional coverage from a single pair of elements.

7.2. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

In view of the above summary it can be said that the symmetric-pair concept for antenna arrays is certainly worth pursuing. It exhibits features such as the binary character of the pair output or the relative robustness against phase and amplitude errors that make it attractive for many applications. Apart from the investigation of the general properties of symmetric-pairs, the emphasis of this thesis has been on the examination of the potential of this novel type of array for direction finding. Initial experiments and theory strongly suggest that symmetric-pair direction finders could be an interesting alternative to current DF systems. This is mainly due to the fact that all known multi-channel direction finders need extremely well matched channel responses. (The amplitude and phase responses of different channels must often be matched to fractions of a dB or degree.) Symmetric-pair DF systems do not have to meet such stringent requirements since they are more tolerant of these errors.

The graphs and tables provided in this thesis allow the design of a symmetric-pair DF system to given specifications, such as minimum

accuracy, average time for a complete bearing, system false-alarm rate or system sensitivity.

To see whether the predictions for the performance of symmetric-pair direction finders are correct, it is necessary to build a fully operational prototype as soon as possible. Since symmetric-pair DF systems are very-wide-aperture direction finders, it is recommended to choose an operating frequency around X-band, to obtain reasonable dimensions for the prototype, which eases construction and testing. Higher frequencies are not advisable, because technology becomes more complicated and vulnerable, and unwanted side effects are bound to lead to unnecessary complications. The initial prototype should only be two dimensional and cover not more than one hemisphere, since this is sufficient to prove the principles. Because the test array does not need to be very sensitive, short non-resonant monopoles are a possible choice for the antenna elements, as they are easy to manufacture, yield reproducible identical patterns and reduce mutual coupling effects.

A combination of the direction finding and beam forming capabilities of symmetric pair-arrays is also worth considering. In communications it is possible to have situations where the exact location of a transmitter is unknown. The same array could then be used to locate the emitter in the first place and subsequently form a beam in its direction to obtain a high enough signal-to-noise ratio that allows reliable communication between the two stations.

Further investigations of the communications aspect of symmetric-pair arrays could also include a more detailed analysis of the effects of squaring on different types of modulation. This would help to characterise the signal requirements for symmetric-pair arrays that are used for gain with isotropic coverage.

References

CHAPTER 1:

Ben 89   **Benjamin, R., and Griffiths, H.D.:** "Aperture-domain signal processing"; *IEE Electronics & Communication Engineering Journal,* Vol. 1, No. 2, pp. 71-80, March/April 1989.

Dze 78   **Dzekov, T.A., and Berkowitz, R.S.:** "Parameters of a spherical random antenna array"; *Electronics Letters,* Vol. 14, No. 16, pp 495-496, August 1978.

End 81   **Ender, J., and Wilden, H.:** "The crow's-nest antenna - a spatial array in theory and experiment"; *Proc. 2nd IEE Int. Conf. on Antennas and Propagation, ICAP 81,* York, IEE Conf. Publ. 195, pp 25-27, April 1981.

Hal 90   **Hall, P.S., and Vetterlein, S.J.:** "Review of radio frequency beamforming techniques for scanned and multiple beam antennas"; *IEE Proc. Pt. H,* Vol. 137, No. 5, pp 293-303, October 1990.

Hay 92   **Haykin, S., Reilly, J.P., Kezys, V., and Vertatschitsch, E.:** "Some aspects of array signal processing"; *IEE Proc. Pt. F,* Vol. 139, No. 1, pp 1-26, February 1992.

Lev 79   **Levin, L.:** "Comment; analytic expression for the frequency response and bandwidth of a spherical random array" *Electronics Letters,* Vol. 15, No. 19, pp 585-586, September 1979.

Lo 64   **Lo, Y.T.:** "A mathematical theory of antenna arrays with randomly spaced elements"; *IEEE Transactions on Antennas and Propagation,* Vol. AP-12, pp 257-268, May 1964.

Mai 82   **Mailloux, R.J.:** "Phased array theory and technology"; *Proceedings of the IEEE,* Vol. 70, No. 3, pp. 246-291, March 1982.

Rud 83   **Rudge, A.W., Milne, K., Olver, A.D., and Knight, P. (Eds.):** "The handbook of antenna design"; Peter Peregrinus Ltd, London, 1983.

Sta 74   **Stark, L.:** "Microwave theory of phased-array antennas - a review"; *Proceedings of the IEEE,* Vol. 62, No. 12, pp. 1661-1701, December 1974.

Ste 76   **Steinberg, B.D.:** "Principles of aperture and array system design"; John Wiley & Sons, New York, 1976.

Wil 87   **Wilden, H., and Ender, J.:** "The crow's-nest antenna - special aspects and results"; *Proc. 17th European Microwave Conference*, Rome, pp 509-514, September 1987.

Wil 90   **Wilden, H., and Ender, J.:** "The crow's nest antenna - experimental results"; *Proc. IEEE 1990 International Radar Conference*, Arlington Virginia, pp 280-285, May 1990.

CHAPTER 2:

Fra 67   **Frank, J., and Ruze, J.:** "Beam steering increments for a phased array"; *IEEE Transactions on Antennas and Propagation*, Vol. AP-15, No. 6, pp 820-821, November 1967.

Gly 68   **Glynn, T.:** "Beam steering of a phased array by conjugate element phase shifting"; *IEEE Transactions on Antennas and Propagation*, Vol. AP-16, No. 5, pp 597-598, September 1968.

CHAPTER 3:

Adc 19   **Adcock, F.:** "Improvements in means for determining the direction of a distant source of electro-magnetic radiation"; British Patent No 130490, 1919.

Bel 31   **Bell, R.J.T.:** "An elementary treatise on coordinate geometry of three dimensions"; Macmillan & Co., London, 1931.

Bel 08   **Bellini, E., and Tosi, A.:** "System einer gerichteten drahtlosen Telegraphie"; Jahrbuch der drahtlosen Telegraphie und Telephonie, Band 1, Verlag S. Hirzel, Leipzig, pp 598-608, 1908.

Bre 90   **Brennan, P.V.:** "Antennas and radar", lecture notes of the antennas part, MSc course at UCL, London, 1990.

Bro 83   **Bronstein, I.N., and Semendjajew, K.A.:** "Taschenbuch der Mathematik"; Verlag Harri Deutsch, Frankfurt/Main, 1983.

Cve 88   **Cvetkovic, M.:** "HF direction finding and null steering using circular arrays"; PhD thesis, University of London, 1988.

Don 76   **Donvito, M., and Kassam, S.A.:** "Random arrays for source parameter estimation", *Electronics Letters*, Vol. 12, No. 24, pp 643-644, November 1976.

Get 89   **Gething, P. J. D.:** "Radio direction finding (and the resolution of multicomponent wavefields)"; Peter Peregrinus Ltd, London, 1989.

Gra 89   **Grabau, R., and Pfaff, K.:** "Funkpeiltechnik"; Franckh'sche Verlagshandlung, Stuttgart, 1989.

Gri 90   **Griffiths, H.D., and Eiges, R.:** "Calibration techniques for phased arrays"; Part 1 of a report from UCL for Philips Research Labs, London, November 1990.

Hil 90   **Hill, D.S.:** "Multiple signal DF using superresolution: a practical assessment"; *IEE Electronics & Communication Engineering Journal*, Vol. 2, No. 6, pp 221-232, December 1990.

Joh 86   **Johnson, R.L., and Miner, G.E.:** "Comparison of superresolution algorithms for radio direction finding"; *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-22, No. 4, pp 432-441, July 1986.

Kas 78   **Kassam, S.A., and Shamash, Y.:** "Range estimation from random array phase measurements"; *Electronics Letters*, Vol. 14, No. 2, pp 40-42, January 1978.

Kra 88   **Kraus, J.D.:** "Antennas"; McGraw-Hill, New York, 1988.

Lip 87   **Lipsky, S.E.:** "Microwave passive direction finding"; John Wiley & Sons, New York, 1987.

Nic 87   **Nickel, U.:** "Angular superresolution with phased array radar: a review of algorithms and operational constraints"; *IEE Proc. Pt. F*, Vol. 134, No. 1, pp 53-59, February 1987.

Rin 56   **Rindfleisch, H.:** "Die Großbasispeilanlage 'Wullenwever'", *NTZ*, Jahrgang 9, pp 119-123, 1956.

Smi 26   **Smith-Rose, R.L., and Barfield, R.H.:** "The cause and elimination of night errors in radio direction-finding"; *Journal of the Institution of Electrical Engineers*, Vol. 64, pp 831-843, London, 1926.

Wat 26   **Watson Watt, R.A., and Herd, J.F.:** "An instantaneous direct-reading radiogoniometer", *Journal of the Institution of Electrical Engineers*, Vol. 64, pp 611-622, London, 1926.

Zhu 90   **Zhu, N.Y., Fallegger, H., and Landstorfer, F.M.:** "Einfluß der Verkopplung auf die Peilgenauigkeit einfacher Interferometer"; *ITG-Fachberichte*, Vol. 111, pp 13-17, Wiesbaden, 20.-23. March 1990.

CHAPTER 4:

And 73   **Andrews, L.C.:** "The probability density function for the output of a cross-correlator with bandpass inputs" *IEEE Transactions on Information Theory*, Vol. IT-19, No. 1, pp 13-19, January 1973.

Cos 50   **Costas, J.P.:** "Periodic sampling of stationary time series"; Technical Report No. 156, Research Laboratory of Electronics, Massachusetts Institute of Technology, 16. May, 1950.

Rob 82   **Robins, W.P.:** "Phase noise in signal sources"; Peter Peregrinus Ltd, London, 1982.

Tau 86   **Taub, H., and Schilling, D.L.:** "Principles of communication systems", McGraw-Hill, New York, 1986.

CHAPTER 5:

Abr 70   **Abramowitz, M., and Stegun, I.A.:** "Handbook of mathematical functions"; Dover Publications, New York, 1970.

And 73   **Andrews, L.C.:** "The probability density function for the output of a cross-correlator with bandpass inputs" *IEEE Transactions on Information Theory,* Vol. IT-19, No. 1, pp 13-19, January 1973.

And 74a   **Andrews, L.C.:** "Output probability density functions for cross correlators utilizing sampling techniques"; *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-10, No. 1, pp 78-81, January 1974.

And 74b   **Andrews, L.C.:** "The output pdf of a polarity coincidence correlation detector"; *IEEE Transactions on Aerospace and Electronic Systems,* Vol. AES-10, pp 712-715, September 1974.

And 80   **Andrews, L.C.:** "Analysis of a cross correlator with a clipper in one channel"; *IEEE Transactions on Information Theory,* Vol. IT-26, No. 6, pp 743-746, November 1980.

And 83   **Andrews, L.C., and Brice, C.S.:** "The pdf and cdf for the sum of N filtered outputs of an analog cross correlator with bandpass inputs"; *IEEE Transactions on Information Theory,* Vol. IT-29, No. 2, pp 299-306, March 1983.

Bri 82   **Brice, C.S., and Andrews, L.C.:** "New expressions for the pdf and cdf of the filtered output of an analog cross correlator"; *IEEE Transactions on Information Theory,* Vol. IT-28, No. 4, pp 668-677, July 1982.

Bro 67   **Brown, J.L., and Piper, H.S.:** "Output characteristic function for an analog crosscorrelator with bandpass inputs" *IEEE Transactions on Information Theory,* Vol. IT-13, No. 1, pp 6-10, January 1967.

Che 68   **Cheng, M.C.:** "The clipping loss in correlation detectors for arbitrary input signal-to-noise ratios"; *IEEE Transactions on Information Theory,* Vol. IT-14, No. 3, pp 382-389, May 1968.

Coo 65   **Cooper, D.C.:** "The probability density function for the output of a correlator with band-pass input waveforms" *IEEE Transactions on Information Theory,* Vol. IT-11, No. 2, pp 190-195, April 1965.

Dav 51   **Davenport, W.B.:** "Correlator errors due to finite observation intervals"; Technical Report No. 191, Research Laboratory of Electronics, Massachusetts Institute of Technology, 8. March, 1951.

Ekr 63   **Ekre, H.:** "Polarity coincidence correlation detection of a weak noise source"; *IEEE Transactions on Information Theory,* Vol. IT-9, No. 1, pp 18-23, January 1963.

Fan 51   **Fano, R.M.:** "On the signal-to-noise ratio in correlation detectors"; Technical Report No. 186, Research Laboratory of Electronics, Massachusetts Institute of Technology, 19. February, 1951.

Gre 57   **Green, P.E.:** "The output signal-to-noise ratio of correlation detectors"; *IRE Transactions on Information Theory,* Vol. IT-3, No. 1, pp 10-18, March 1957.

Hor 55   **Horowitz, M., and Johnson, A.A.:** "Theory of noise in a correlation detector"; *IRE Transactions on Information Theory* , Vol. IT-1, No. 3, pp 3-5, December 1955.

Joh 72   **Johnson, N.L., and Kotz, S.:** "Distributions in Statistics: Continuous multivariate distributions"; John Wiley & Sons, New York, 1972.

Kai 66   **Kaiteris, C.:** "The density and distribution functions for the product of two narrow-band waveforms"; *Proceedings of the IEEE,* Vol. 54, No. 11, pp 1588-1589, November 1966.

Lam 56   **Lampard, D.G.:** "The probability distribution for the filtered output of a multiplier whose inputs are correlated, stationary, Gaussian time-series"; *IRE Transactions on Information Theory,* Vol. IT-2, No. 1, pp 4-11, March 1956.

Lee 50a   **Lee, Y.W., and Wiesner, J.B.:** "Correlation functions and communication applications"; *Electronics,* Vol. 23, pp 86-92, June 1950.

Lee 50b   **Lee, Y.W., Cheatham, T.P., and Wiesner, J.B.:** "Application of correlation analysis to the detection of periodic signals in noise"; *Proceedings of the IRE,* Vol. 38, No. 10, pp 1165-1171, October 1950.

Leo 89   **Leon-Garcia, A.:** "Probability and random processes for electrical engineering"; Addison-Wesley, New York, 1989.

Mar 60   **Marcum, J.I., and Swerling, P.:** "Studies of target detection by pulsed radar"; (Reprint of RAND Research Memos: Marcum, J.I., 'A statistical theory of target detection by pulsed radar', RM-754, December 1, 1947; Marcum, J.I., 'Mathematical Appendix', RM-753, July 1, 1948; Swerling, P., 'Probability of detection for fluctuating targets', RM-1217, March 17, 1954) *IRE Transactions on Information Theory*, Vol. IT-6, pp 59-308, April 1960.

Mil 74   **Miller, L.E., and Lee, J.S.:** "The probability density function for the output of an analog cross-correlator with correlated bandpass inputs" *IEEE Transactions on Information Theory*, Vol. IT-20, No. 4, pp 433-440, July 1974.

Mil 91   **Milne, K.:** private communication, September 1991.

Mur 71   **Murarka, N.P.:** "The probability density function for correlators with correlated noisy reference inputs; *IEEE Transactions on Communication Technology*, Vol. COM-19, No. 5, pp 711-714, October 1971.

Pap 91   **Papoulis, A.:** "Probability, random variables and stochastic processes"; McGraw-Hill, New York, 1991.

Ric 44   **Rice, S.O.:** "Mathematical analysis of random noise"; *Bell System Technical Journal*, Vol. 23, pp 282-332, July 1944.

Ric 45   **Rice, S.O.:** "Mathematical analysis of random noise"; *Bell System Technical Journal*, Vol. 24, pp 46-156, January 1945.

Roe 61   **Roe, G.M., and White, G.M.:** "Probability density functions for correlators with noisy reference signals"; *IRE Transactions on Information Theory*, Vol. IT-7, No. 1, pp 13-18, January 1961.

Sko 70   **Skolnik, M.I.:** "Radar Handbook", McGraw-Hill, New York, 1970.

Str 70   **Stremler, F.G., and Jensen, T.:** "Probability density function for the output of an analog crosscorrelator with bandpass inputs"; *IEEE Transactions on Information Theory*, Vol. IT-16, No. 5, pp 627-629, September 1970.

Wha 71   **Whalen, A.D.:** "Detection of signals in noise"; Academic Press, New York, 1971.

Wil 70   **Wilson, R.:** "The probability distribution of the sum of N sampled baseband outputs of a product detector whose inputs are equal power coherent narrow-band signals in a Gaussian channel"; *IEEE Transactions on Communications*, Vol. COM-27, No. 5, pp 796-801, May 1979.

Wol 91   **Wolfram, S.:** "Mathematica, a system for doing mathematics by computer"; Addison-Wesley, New York, 1991.

CHAPTER 6:

Bre 91   **Brennan, P.V.:** "A low-cost phased array antenna for land-mobile satcom applications"; *IEE Proc. Pt. H*, Vol. 138, No. 2, pp 131-136, April 1991.

Ste 76   **Steinberg, B.D.:** "Principles of aperture and array system design"; John Wiley & Sons, New York, 1976.

APPENDICES:

Abr 70   **Abramowitz, M., and Stegun, I.A.:** "Handbook of mathematical functions"; Dover Publications, New York, 1970.

And 83   **Andrews, L.C., and Brice, C.S.:** "The pdf and cdf for the sum of N filtered outputs of an analog cross correlator with bandpass inputs"; *IEEE Transactions on Information Theory*, Vol. IT-29, No. 2, pp 299-306, March 1983.

Bey 78   **Beyer, W.H. (Editor):** "CRC Handbook of mathematical sciences"; CRC Press, West Palm Beach, Florida, 1978.

Bre 90   **Brennan, P.V.:** "A Carrier-Phased Satellite Receiving Array"; *International Journal of Satellite Communications*, Vol. 8, pp 1-9, 1990.

Bri 82   **Brice, C.S., and Andrews, L.C.:** "New expressions for the pdf and cdf of the filtered output of an analog cross correlator"; *IEEE Transactions on Information Theory*, Vol. IT-28, No. 4, pp 668-677, July 1982.

Bro 83   **Bronstein, I.N., and Semendjajew, K.A.:** "Taschenbuch der Mathematik"; Verlag Harri Deutsch, Frankfurt/Main, 1983.

Cla 85   **Clarke, A.B., and Disney, R.L.:** "Probability and random processes: A first course with applications"; John Wiley & Sons, New York, 1985.

Gra 80   **Gradshteyn, I.S., and Ryzhik, I.M.:** "Tables of integrals, series, and products"; Academic Press, London, 1980.

Kra 88   **Kraus, J.D.:** "Antennas"; McGraw-Hill, New York, 1988.

Mil 91   **Milne, K.:** private communication, September 1991.

Spi 88   **Spiegel, M.R.:** "Höhere Mathematik für Ingenieure und Natur-wissenschaftler", McGraw-Hill, Hamburg, 1988.

Wha 71   **Whalen, A.D.:** "Detection of signals in noise"; Academic Press, New York, 1971.

APPENDICES

A 1 Second Method to Derive the Output Phase of a Symmetric-Pair with Phase and Amplitude Errors

The alternative derivation of the phase ϕ_{pae} of a symmetric-pair with phase and amplitude errors explores the relationships between the vectors of the symmetric pair as shown in Fig. A1-1.
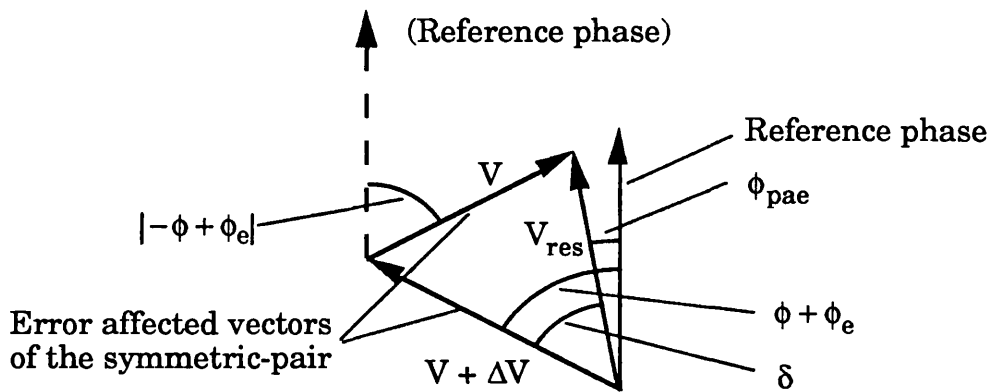
Figure A1-1: Vector representation of phase and amplitude errors of a symmetric-pair

To calculate ϕ_{pae}, the following relation can be used [Bro 83]:

$$\tan(\gamma) = \frac{c\,\sin(\alpha)}{b - c\,\cos(\alpha)}, \qquad \text{with:} \qquad \qquad \qquad \qquad \qquad \text{...(A1.1)}$$

With Fig. A1-1 and eqn. (A1.1), ϕ_{pae} is now defined as:

$$\phi_{pae} = \phi + \phi_e - \delta = \phi + \phi_e - \arctan\left(\frac{V\sin(\pi - 2\phi)}{V + \Delta V - V\cos(\pi - 2\phi)}\right)$$

$$= \phi + \phi_e - \arctan\left(\frac{V\sin(2\phi)}{\Delta V + 2V\cos^2(\phi)}\right) \qquad \qquad \text{...(A1.2)}$$

199

Using the following formula [Bro 83],

$$\arctan(x) \pm \arctan(y) = \arctan\left(\frac{x \pm y}{1 \mp xy}\right)$$

...(A1.3)

eqn. (A1.2) can be modified to:

$$\phi_{pae} = \arctan\left(\frac{\tan(\phi + \phi_e) - \dfrac{V\sin(2\phi)}{\Delta V + 2V\cos^2(\phi)}}{1 + \tan(\phi + \phi_e)\ \dfrac{V\sin(2\phi)}{\Delta V + 2V\cos^2(\phi)}}\right)$$

...(A1.4)

After some manipulations, it is possible to bring eqn. (A1.4) in the following form:

$$\phi_{pae} = \arctan\left(\frac{(2V + \Delta V)\cos(\phi)\sin(\phi_e) + \Delta V\cos(\phi_e)\sin(\phi)}{(2V + \Delta V)\cos(\phi)\cos(\phi_e) - \Delta V\sin(\phi)\sin(\phi_e)}\right)$$

...(A1.5)

To be able to compare the above derivation with the result obtained in section 2.5.3, it is convenient to rearrange eqn. (2.18). With the help of eqn. (A1.3), eqn. (2.18), which was obtained by superposition of the results for amplitude and symmetrical phase error only, can be rewritten in the following form:

$$\phi_{pae} = \phi_e + \arctan\left(\frac{\Delta V}{2V + \Delta V}\tan(\phi)\right) = \arctan\left(\frac{\tan(\phi_e) + \dfrac{\Delta V}{2V + \Delta V}\tan(\phi)}{1 - \tan(\phi_e)\dfrac{\Delta V}{2V + \Delta V}\tan(\phi)}\right)$$

$$= \arctan\left(\frac{(2V + \Delta V)\cos(\phi)\sin(\phi_e) + \Delta V\cos(\phi_e)\sin(\phi)}{(2V + \Delta V)\cos(\phi)\cos(\phi_e) - \Delta V\sin(\phi)\sin(\phi_e)}\right)$$

...(A1.6)

A comparison of eqn. (A1.5) and eqn. (A1.6) shows that the two expressions are identical, which means that both derivations yield the same result.

A2 DERIVATION OF THE AZIMUTH AND ELEVATION ANGLE IN THE LOCAL COORDINATE SYSTEM

In this appendix, the angles α_l and ε_l are derived in a local coordinate system for given ϕ_{cone1}, ϕ_{cone2} and γ. It can be seen in Fig. A2-1 and Fig. A2-2 that the calculations are based on the assumption that the intersections take place on the surface of a unity sphere. Furthermore, the variables a_x and o_x stand for the adjacent and opposite of a right-angled triangle with the angle $\phi_{cone\,x}$. a, b and c are auxiliary parameters.

Using Fig. A2-1, the following expression for the elevation angle ε_l can be obtained:

$$\sin(\varepsilon_l) = h / 1 \qquad \text{...(A2.1)}$$

As can be seen in Fig. A2-1, h can be expressed as:

$$h^2 = o_1^{\;2} - c^2 \qquad \text{...(A2.2)}$$

where:
$$o_1 = 1\sin(\phi_{cone1}) \qquad \text{...(A2.3a)}$$

$$c = \frac{b}{\sin(\gamma)} \qquad \text{...(A2.3b)}$$

With Fig. A2-2, b can be derived as follows:

$$b = a_2 - a \qquad \text{...(A2.4)}$$

where:

$$a_2 = 1\cos(\phi_{cone2})$$ ...(A2.5a)

$$a = a_1\cos(\gamma)$$ ...(A2.5b)

and:

$$a_1 = 1\cos(\phi_{cone1})$$ ...(A2.6)

Using eqns. (A2.3b) to (A2.6), the following expression for c can now be derived:

$$c = \left(\cos(\phi_{cone2}) - \cos(\phi_{cone1})\cos(\gamma)\right)/\sin(\gamma)$$ ...(A2.7)

Manipulating eqn. (A2.1) with the help of eqn. (A2.2), eqn. (A2.3a) and eqn. (A2.7) leads to the formula for the local elevation angle ε_ℓ:

$$\varepsilon_{\ell1/2} = \pm\arcsin\sqrt{\left(\sin(\phi_{cone1})\right)^2 - \left(\frac{\cos(\phi_{cone2}) - \cos(\phi_{cone1})\cos(\gamma)}{\sin(\gamma)}\right)^2}$$ ...(A2.8)

The $\pm$ sign for the elevation angle takes the ambiguity of the intersection into account (see Fig. A2-1).

Fig. A2-1 defines the azimuth angle α_ℓ as:

$$\tan(\alpha_\ell) = \frac{c}{\cos(\phi_{cone1})}$$ ...(A2.9)

After substitution of eqn. (A2.7) into eqn. (A2.9), the local azimuth angle is equal to:

$$\alpha_\ell = \arctan\left(\frac{\cos(\phi_{cone2}) - \cos(\phi_{cone1})\cos(\gamma)}{\sin(\gamma)\cos(\phi_{cone1})}\right)$$ ...(A2.10)
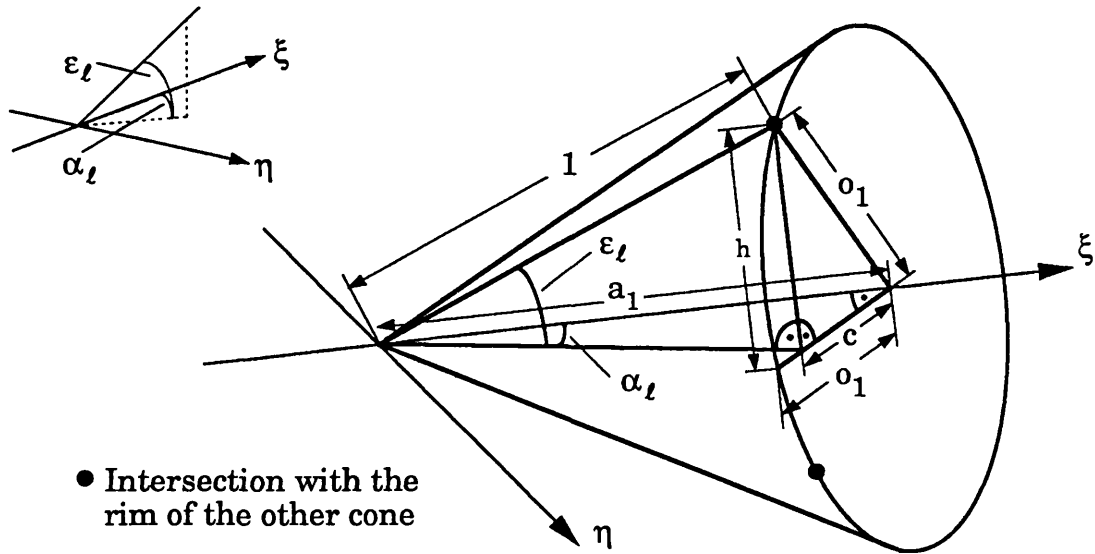
Figure A2-1: Derivation of the local azimuth and elevation angle
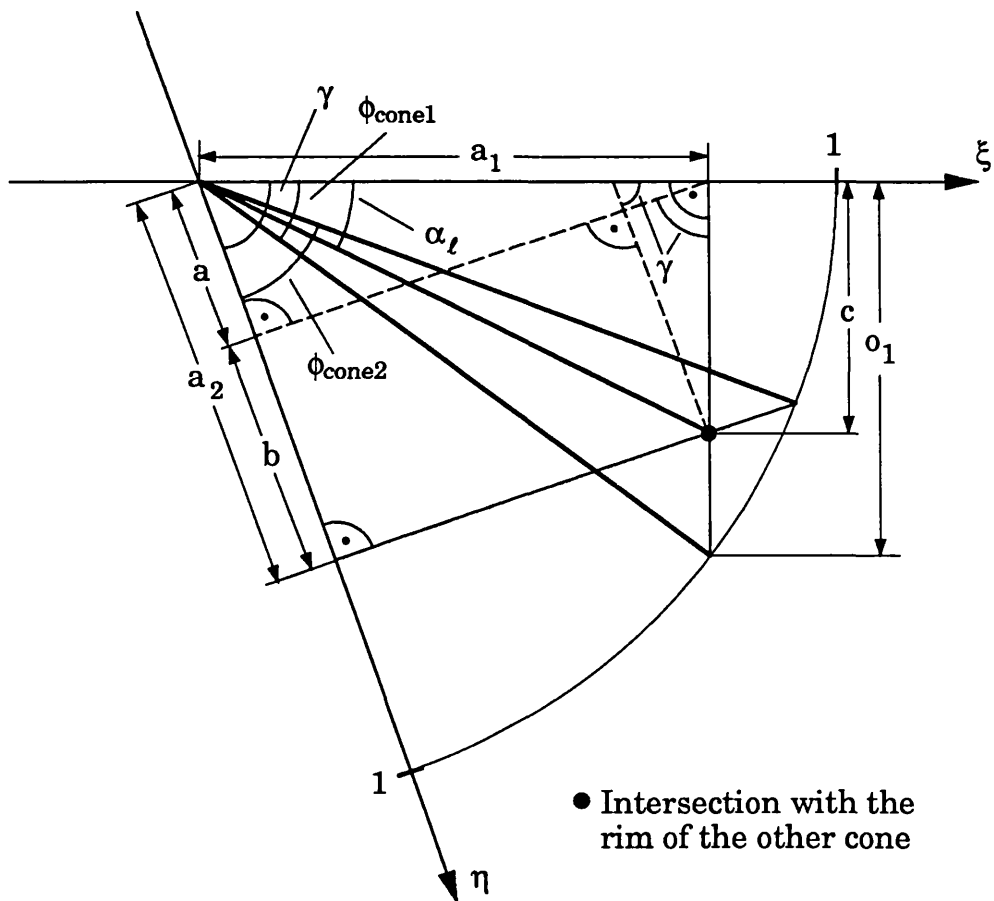
Figure A2-2: Derivation of the local azimuth and elevation angle (Top view)

A3 Solid Angle Covered by the Intersection of Two Cones from Orthogonal Axes

The area of a (mathematically) smooth surface is defined as [Spi 88]:

$$\iint \sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\; dx\, dy$$

...(A3.1)

where $z = f(x,y)$ is the equation that describes the surface. In this calculation, the surface under consideration is the surface of a unity sphere, and z is therefore equal to:

$$z = \sqrt{1 - x^2 - y^2}$$

...(A3.2)

Hence, the integral to be solved has the following form:

$$\int_{y_1}^{y_2} \int_{x_1}^{x_2} \frac{1}{\sqrt{1 - y^2 - x^2}}\; dx\, dy$$

...(A3.3)

Integrating eqn. (A3.3) with respect to x gives:

$$\int_{y_1}^{y_2} \arcsin\left(\frac{x_2}{\sqrt{1-y^2}}\right) dy - \int_{y_1}^{y_2} \arcsin\left(\frac{x_1}{\sqrt{1-y^2}}\right) dy$$

...(A3.4)

These are now two integrals of the type:

$$\int \arcsin\left(\frac{c}{\sqrt{1-z^2}}\right) dz$$

...(A3.5)

After integration by parts, eqn. (A3.5) becomes:

$$z\arcsin\left(\frac{c}{\sqrt{1-z^2}}\right) - \int \frac{c\, z^2}{(1-z^2)\sqrt{1-z^2-c^2}}\; dz$$

...(A3.6)

With the substitution u = 1 - z² , eqn. (A3.6) can be rewritten as:

$$z \arcsin\left(\frac{c}{\sqrt{1-z^2}}\right) - \int \frac{-c\sqrt{1-u}}{2u\sqrt{u-c^2}}\,du$$

$$\dots (A3.7)$$

Expanding eqn. (A3.7) with $\sqrt{1-u}$ finally gives:

$$z \arcsin\left(\frac{c}{\sqrt{1-z^2}}\right) - \int \frac{c}{2\sqrt{-u^2+\left(c^2+1\right)u-c^2}}\,du + \int \frac{c}{2u\sqrt{-u^2+\left(c^2+1\right)u-c^2}}\,du$$

$$\dots (A3.8)$$

These are integrals containing $\sqrt{ax^2+bx+c}$ that can be solved with the help of standard integral tables (e.g. [Gra 80], integrals No. 2.261 and No. 2.266, or [Bro 83], integrals No. 241 and No. 258).

Equation (A3.4) can now be manipulated with the help of eqns. (A3.5) to eqn. (A3.8), which leads to the following expression for the surface integral:

$$\int_{y_1}^{y_2}\int_{x_1}^{x_2} \frac{1}{\sqrt{1-y^2-x^2}}\,dx\,dy =$$

$$y_2 \arcsin\left(\frac{x_2}{\sqrt{1-y_2^2}}\right) + \frac{x_2}{2}\arcsin\left(\frac{x_2^2-1-2y_2^2}{\left|x_2^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_2^2+1\right)\left(1-y_2^2\right)-2x_2^2}{\left(1-y_2^2\right)\left|x_2^2-1\right|}\right) -$$

$$y_1 \arcsin\left(\frac{x_2}{\sqrt{1-y_1^2}}\right) + \frac{x_2}{2}\arcsin\left(\frac{x_2^2-1-2y_1^2}{\left|x_2^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_2^2+1\right)\left(1-y_1^2\right)-2x_2^2}{\left(1-y_1^2\right)\left|x_2^2-1\right|}\right) -$$

$$y_2 \arcsin\left(\frac{x_1}{\sqrt{1-y_2^2}}\right) + \frac{x_1}{2}\arcsin\left(\frac{x_1^2-1-2y_2^2}{\left|x_1^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_1^2+1\right)\left(1-y_2^2\right)-2x_1^2}{\left(1-y_2^2\right)\left|x_1^2-1\right|}\right) +$$

$$y_1 \arcsin\left(\frac{x_1}{\sqrt{1-y_1^2}}\right) + \frac{x_1}{2}\arcsin\left(\frac{x_1^2-1-2y_1^2}{\left|x_1^2-1\right|}\right) + \frac{1}{2}\arcsin\left(\frac{\left(x_1^2+1\right)\left(1-y_1^2\right)-2x_1^2}{\left(1-y_1^2\right)\left|x_1^2-1\right|}\right)$$

$$\dots (3.14)$$

where: | ... | means absolute value

A4 MAXIMUM ERROR CAUSED BY MUTUAL COUPLING AND FEED NETWORK IMPERFECTIONS

To assess the impact of systematic errors on the accuracy of the DF system, the magnitude of the towards the feed network travelling corrupted signals a_{F_2} (= b_{M_1}) must be known (see Fig. 3-18). For the calculation of b_{M_1}, it is necessary to determine the vector b_{F_2}, and evaluate the mutual coupling coefficients.

A 4.1. CALCULATION OF b_{F_2}:

The scattering parameters of a 180° hybrid are as follows:

b = H a

or:

$$
\begin{pmatrix} b_{F_{1,1}} \\ b_{F_{1,2}} \\ b_{F_{2,1}} \\ b_{F_{2,2}} \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{13} \\ h_{12} & h_{22} & h_{23} & -h_{23} \\ h_{13} & h_{23} & h_{33} & h_{34} \\ h_{13} & -h_{23} & h_{34} & h_{44} \end{pmatrix} \begin{pmatrix} a_{F_{1,1}} \\ a_{F_{1,2}} \\ a_{F_{2,1}} \\ a_{F_{2,2}} \end{pmatrix}
$$

$$...(A4.1)$$

By assuming a reflection coefficient of ρ at the Σ and Δ output port of the hybrid, two additional equations can be obtained:

$$a_{F_{1,1}} = \rho \, b_{F_{1,1}} \qquad\qquad ...(A4.2a)$$

$$a_{F_{1,2}} = \rho \, b_{F_{1,2}} \qquad\qquad ...(A4.2b)$$

Manipulation of the above equations yields the formulae for $b_{F_{1,1}}$ and $b_{F_{1,2}}$. Since no information about the exact phases of either the hybrid coefficients or the reflection coefficient ρ is available, $b_{F_{1,1}}$ and $b_{F_{1,2}}$ must be estimated with their absolute values:

$$\left| b_{F_{1,1}} \right| \le \left| \frac{\left(a_{F_{2,1}} - a_{F_{2,2}} \right) h_{23} h_{12} \rho + \left(a_{F_{2,1}} + a_{F_{2,2}} \right) h_{13} (1 - h_{22} \rho)}{(1 - h_{11} \rho)(1 - h_{22} \rho) - (h_{12} \rho)^2} \right| \qquad \text{...(A4.3a)}$$

$$\left| b_{F_{1,2}} \right| \le \left| \frac{\left(a_{F_{2,1}} - a_{F_{2,2}} \right) h_{13} h_{12} \rho + \left(a_{F_{2,1}} + a_{F_{2,2}} \right) h_{23} (1 - h_{11} \rho)}{(1 - h_{11} \rho)(1 - h_{22} \rho) - (h_{12} \rho)^2} \right| \qquad \text{...(A4.3b)}$$

where according to the data sheet of an x-band hybrid:

$|h_{13}| \approx |h_{23}| \approx |h_{14}| \approx |h_{24}| \approx 1/\sqrt{2}$, $|h_{xx}| < 0.2$ and
$|h_{12}| \approx |h_{34}| \approx |\rho| < 0.1$

The expressions for $b_{F_{2,1}}$ and $b_{F_{2,2}}$ can be derived using eqn. (A4.1) and eqns. (A4.2a/b):

$$b_{F_{2,1}} = h_{13} \rho \, b_{F_{1,1}} + h_{23} \rho \, b_{F_{1,2}} + h_{33} \, a_{F_{2,1}} + h_{34} \, a_{F_{2,2}} \qquad \text{...(A4.4a)}$$

$$b_{F_{2,2}} = h_{13} \rho \, b_{F_{1,1}} - h_{23} \rho \, b_{F_{1,2}} + h_{34} \, a_{F_{2,1}} + h_{44} \, a_{F_{2,2}} \qquad \text{...(A4.4a)}$$

After insertion of eqns. (A4.3a/b) into eqns. (A4.4a/b) and with the numbers from above, it is possible to approximate the expressions for $b_{F_{2,1}}$ and $b_{F_{2,2}}$ as a function of $a_{F_{2,1}}$ and $a_{F_{2,2}}$:

$$\left| b_{F_{2,1}} \right| \le \left| 0.3 \, a_{F_{2,1}} + 0.1 \, a_{F_{2,2}} \right| \qquad \text{...(A4.5a)}$$

$$\left| b_{F_{2,2}} \right| \le \left| 0.3 \, a_{F_{2,1}} + 0.1 \, a_{F_{2,2}} \right| \qquad \text{...(A4.5b)}$$

It should be mentioned that the finite isolation between the two element ports of the hybrid and contributions from signals reflected at these ports dominate $b_{F_{2,1}}$ and $b_{F_{2,2}}$. Hence, low reflection coefficients and high isolation would decrease the magnitude of $b_{F_{2,1}}$ and $b_{F_{2,2}}$.

A 4.2. CALCULATION OF THE MUTUAL COUPLING COEFFICIENTS:

The mutual coupling matrix of this five element arrangement has the following form (compare with Fig. 3-19 and eqn 3.17):

$$
\begin{pmatrix} b_{M1,1} \\ b_{M1,2} \\ b_{M1,3} \\ b_{M1,4} \\ b_{M1,5} \end{pmatrix} =
\begin{pmatrix}
m_{11} & m_{12} & m_{12} & m_{12} & m_{12} \\
m_{12} & m_{22} & m_{23} & m_{24} & m_{24} \\
m_{12} & m_{23} & m_{22} & m_{24} & m_{24} \\
m_{12} & m_{24} & m_{24} & m_{22} & m_{23} \\
m_{12} & m_{24} & m_{24} & m_{23} & m_{22}
\end{pmatrix}
\begin{pmatrix} a_{M1,1} \\ a_{M1,2} \\ a_{M1,3} \\ a_{M1,4} \\ a_{M1,5} \end{pmatrix}
\qquad ...(A4.6)
$$

Since the array elements are assumed to be dipoles, the mutual coupling coefficients are calculated utilising the data for the mutual impedances of two parallel dipoles given by Kraus [Kra 88]. The impedances can be converted into scattering coefficients with the following formula:

$$ S = (Z - I)\,(Z + I)^{-1} \qquad ...(A4.7) $$

where Z is the impedance coupling matrix, normalised to 50Ω, S is the corresponding scattering matrix and I the identity matrix.

Spacing	Mutual impedance coefficients (normalised to 50Ω)	Scattering parameters for different reflection coefficients at the elements	
		$\mid m_{xx} \mid = 0.1$	$\mid m_{xx} \mid = 0.2$
$\dfrac{\lambda}{2}$	$z_{12} = -0.254 - j0.56$	$\mid m_{12} \mid = 0.143$	$\mid m_{12} \mid = 0.169$
$\dfrac{\lambda}{\sqrt{2}}$	$z_{24} = -0.5$	$\mid m_{24} \mid = 0.048$	$\mid m_{24} \mid = 0.056$
λ	$z_{23} = 0.076 + j0.36$	$\mid m_{23} \mid = 0.017$	$\mid m_{23} \mid = 0.021$

Table A4-1:   Impedance and scattering parameters for the coupling between two parallel dipoles

Since the exact reflection coefficients m_{xx} of the elements are unknown, they had to be approximated by assumed maximum values. The mutual coupling coefficients are then calculated for the phase values

of the reflection coefficient that yield maximum coupling. Table A4-1 gives the mutual impedance coefficients and the corresponding worst case scattering parameters for the coupling between two parallel dipoles for different spacings.

A 4.3. CALCULATION OF THE MAXIMUM ERROR:

To describe the influence of mutual coupling and an imperfect feed network on the performance of the system, the signals travelling towards the ports of each hybrid and the signal of the reference element must be calculated. With eqn. A4.6 and knowing that M_{21} is an identity matrix (see section 3.6.1.), the following equations can be derived:

$$b_{M_{1,1}} = m_{11}\, a_{M_{1,1}} + m_{12}\, (a_{M_{1,2}} + a_{M_{1,3}} + a_{M_{1,4}} + a_{M_{1,5}}) + 1 \qquad ...(A4.8a)$$

$$b_{M_{1,2}} = m_{12}\, a_{M_{1,1}} + m_{22}\, a_{M_{1,2}} + m_{23}\, a_{M_{1,3}} + m_{24}\, (a_{M_{1,4}} + a_{M_{1,5}}) + e^{\Psi_1(\theta)}$$
$$...(A4.8b)$$

$$b_{M_{1,3}} = m_{12}\, a_{M_{1,1}} + m_{23}\, a_{M_{1,2}} + m_{22}\, a_{M_{1,3}} + m_{24}\, (a_{M_{1,4}} + a_{M_{1,5}}) + e^{\Psi_2(\theta)}$$
$$...(A4.8c)$$

where the last term in every equation represents the signal incident at each element. The equations for the second hybrid are similar to eqn. (A4.8b) and eqn. (A4.8c).

Because of the symmetry of the arrangement, it can be assumed that the signals travelling towards the elements of the two pairs have the same absolute value:

$$|a_{M_{1,2}}| = |a_{M_{1,3}}| = |a_{M_{1,4}}| = |a_{M_{1,5}}| \qquad ...(A4.9)$$

Accepting that the signal of the reference element encounters a reflection coefficient of ρ and using eqns. (A4.5) with $a_{M_1} = b_{F_2}$, it is possible to obtain two additional equations:

$$|a_{M_{1,1}}| = |\rho\, b_{M_{1,1}}| \qquad ...(A4.10a)$$

$$|a_{M_{1,2}}| = |a_{M_{1,3}}| = |a_{M_{1,4}}| = |a_{M_{1,5}}| = 0.4 \ |b_{M_{1,2}}| = 0.4 \ |b_{M_{1,3}}|$$

$$...(A4.10b)$$

Insertion of eqns. (A4.10a/b) into eqns. (A4.8a-c) leads to expressions of the following form:

$$|b_{M_{1,1}}| \leq |1\ 1| + |m_{11} \ \rho \ b_{M_{1,1}}| + |4 \ m_{12} \ 0.4 \ b_{M_{1,2}}| \qquad ...(A4.11a)$$

$$|b_{M_{1,2}}| \leq |e^{\Psi_1(\theta)}| + |(m_{22} + m_{23}) \ 0.4 \ b_{M_{1,2}}| +$$
$$|m_{24} \ 0.8 \ b_{M_{1,1}} + m_{12} \ \rho \ b_{M_{1,1}}| \quad ...(A4.11b)$$

$$|b_{M_{1,3}}| \leq |e^{\Psi_2(\theta)}| + |(m_{22} + m_{23}) \ 0.4 \ b_{M_{1,3}}| +$$
$$|m_{24} \ 0.8 \ b_{M_{1,1}} + m_{12} \ \rho \ b_{M_{1,1}}| \quad ...(A4.11c)$$

With $|\rho| = 0.1$ and the numbers from above, these equations can be solved for $b_{M_{1,x}}$ and the worst case magnitudes of the signals travelling towards the feed network can be calculated. The knowledge of the magnitude of $\mathbf{b}_{M_1}$ allows us now to explain the influence of mutual coupling and the imperfections of the feed network on the DF accuracy in more detail.

The examination of the two signals travelling towards the element port of a hybrid shows that part of the error signal comes from the interaction between the elements of a pair and the other elements of the array. Due to the symmetry of the array, this portion has identical influence on both elements of a pair and does not affect the DF accuracy, it only offsets the phase of the Σ and Δ processed pair signal relative to the reference phase. This part is given by:

$$|b_{M_{1,2}}|_{(a)} = |e^{\Psi_1(\theta)}| + |m_{24} \ 0.8 \ b_{M_{1,1}} + m_{12} \ \rho \ b_{M_{1,1}}| \qquad ...(A4.12a)$$

$$|b_{M_{1,3}}|_{(a)} = |e^{\Psi_2(\theta)}| + |m_{24} \ 0.8 \ b_{M_{1,1}} + m_{12} \ \rho \ b_{M_{1,1}}| \qquad ...(A4.12b)$$

The other part, which accounts for the coupling between the elements of a pair and the reflection coefficient of the elements can act as opposite phase shift between the signals of a pair as well as in phase component for both elements, dependent on the phases of $|a_{M_{1,2}}|$ and $|a_{M_{1,3}}|$. Since the phase of most of the error parameters is not known, it is impossible to

predict the portions of these two components. For a worst case scenario, it is assumed that the error vectors are in quadrature to the signal and opposite for the input signals of the hybrid. Hence, the maximum portion which can cause opposite phase deviation is given as follows:

$$|b_{M_{1,2}}|_{(b)} = |(m_{22} + m_{23})\ 0.4\ b_{M_{1,2}}| \qquad \qquad ...(A4.13a)$$

$$|b_{M_{1,3}}|_{(b)} = |(m_{22} + m_{23})\ 0.4\ b_{M_{1,3}}| \qquad \qquad ...(A4.13b)$$

Each of the signals in eqns. (A4.12a/b) and eqns. (A4.13a/b) consists of the original signal vector with unity length and an error vector on top of it. To calculate the maximum phase deviation of each signal, it is assumed that the error vector is in quadrature with the signal vector. If the maximum amplitude difference between two signals is of interest, it is assumed that the error vector is in phase with one and in anti-phase with the other signal. Table 3-6 in Chapter 3 summarises the effects of an imperfect feed network and mutual coupling on the DF system for different reflection coefficients of the dipole elements.

A5 NOISE POWER AT THE PHASE DETECTOR OUTPUT AFTER INTEGRATION

The expression for the noise power at the phase detector output after integration over a fixed period T is given by (see eqn. (4.30) in chapter 4):

$$\sigma_\infty^2 = \frac{2}{T} \int_{\tau=0}^{\tau=T} \left(1 - \frac{\tau}{T}\right) R_N(\tau) \, d\tau$$

$$...(A5.1)$$

where $R_N(\tau)$ is the autocorrelation function of the noise at the correlator output, which is (see eqn.(4.20) in chapter 4):

$$R_N(\tau) = \frac{\sin(\pi B_{IF}\tau)}{\pi B_{IF}\tau}(y_1 - y_2)B_{IF} + \left(\frac{\sin(\pi B_{IF}\tau)}{\pi B_{IF}\tau}\right)^2 y_2 B_{IF}$$

$$...(A5.2)$$

Inserting eqn. (A5.2) into eqn. (A5.1) and with the substitution $x = \pi B_{IF}\tau$, the following expression can be obtained:

$$\sigma_\infty^2 = \int_0^{T\pi B_{IF}} \left(c_1 \frac{\sin(x)}{x} - c_2 \sin(x) + c_3 \frac{\sin^2(x)}{x^2} - c_4 \frac{\sin^2(x)}{x}\right) dx$$

$$...(A5.3)$$

with:

$$c_1 = \frac{y_1 - y_2}{\pi}; \quad c_2 = \frac{y_1 - y_2}{\pi^2 B_{IF}T}; \quad c_3 = \frac{y_2}{\pi}; \quad c_4 = \frac{y_2}{\pi^2 B_{IF}T}$$

Applying the relation $\sin^2(x) = \frac{1}{2}(1 - \cos(2x))$, eqn. (A5.3) can be modified to:

$$\sigma_\infty^2 = \int_0^{T\pi B_{IF}} \left(c_1 \frac{\sin(x)}{x} - c_2 \sin(x) + c_3 \frac{1 - \cos(2x)}{2x^2} - c_4 \frac{1 - \cos(2x)}{2x}\right) dx$$

$$...(A5.4)$$

The above integral can be solved using

$$\int \frac{\cos(ax)}{x^2} dx = -\frac{\cos(ax)}{x} - a \int \frac{\sin(ax)}{x} dx$$

([Bey 78], integral No. 405)

and the series representations of two of the integrals:

$$\int \frac{\sin(ax)}{x} dx = \sum_{n=0}^{\infty} (-1)^n \frac{(ax)^{2n+1}}{(2n+1)(2n+1)!}$$

([Bey 78], integral No. 397)

$$\int \frac{\cos(ax)}{x} dx = \ln(x) + \sum_{n=1}^{\infty} (-1)^n \frac{(ax)^{2n}}{2n(2n)!}$$

([Bey 78], integral No. 398)

Integrating eqn. (A5.4) and after few manipulations, σ_{∞}^2 has the following form:

$$\sigma_{\infty}^n = \left[\sum_{n=0}^{\infty} (-1)^n \frac{c_1 (x)^{2n+1} + c_3 \, 2(x)^{2n+1}}{(2n+1)(2n+1)!} + \right.$$

$$\left. c_2 \cos(x) - c_3 \frac{\sin^2(x)}{x} + \frac{c_4}{2} \sum_{n=1}^{\infty} (-1)^n \frac{(2x)^{2n}}{2n(2n)!} \right]_0^{T\pi B_{IF}}$$

...(A5.5)

After insertion of the bounds of the integral and with:

$$\lim_{x \to 0} \frac{\sin^2(x)}{x} \overset{\text{l'Hospital}}{=} \lim_{x \to 0} \frac{2\sin(x)\cos(x)}{1} = 0 \quad,$$

the noise power at the output of the phase detector after integration over a fixed period of time T is equal to:

$$\sigma_{\infty}^2 = \frac{2}{T} \left[u(c_1 + 2c_3) + \sum_{n=1}^{\infty} (-1)^n \left[\frac{c_1(u)^{2n+1} + c_3(2u)^{2n+1}}{(2n+1)(2n+1)!} + \frac{c_4(2u)^{2n}}{4n(2n)!} \right] + \right.$$

$$\left. c_2 (\cos(u) - 1) - c_3 \left(\frac{\sin^2(u)}{u} \right) \right]$$

...(A5.6)

where $u = \pi B_{IF} T$.

A6 INTEGRATED OUTPUT OF AN ANALOGUE CROSS-CORRELATOR FED WITH GAUSSIAN NOISE

The first five sections of this appendix are based on an unpublished note by Milne [Mil 91]. The main derivations and results are reproduced here with some additional comments and references. In the remaining appendix, it is shownthat Milne's results can be related to expressions obtained by Andrews and Brice [And 83].

A cross correlator is fed with zero mean narrow band Gaussian noise and its output signal is sampled after zonal low-pass filtering. Firstly, the probability density function (pdf) of a single sample is derived, before the pdf for the summation of n independent samples will be obtained. A formula for the probability to exceed a certain threshold will also be presented. The examined arrangement is shown in Fig. A6-1.
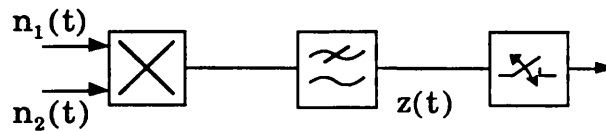
Figure A6-1: Examined cross-correlator

A6.1 PROBABILITY DENSITY FUNCTION FOR A SINGLE SAMPLE

The narrow band noise in the two input channels of the correlator is represented using the quadrature representation of narrow band noise where [Wha 70]:

$$n_1(t) = x_1(t)\cos(\omega t) - y_1 \sin(\omega t) \qquad \qquad ...(A6.1a)$$

$$n_2(t) = x_2(t)\cos(\omega t) - y_2 \sin(\omega t) \qquad \qquad \text{...(A6.1b)}$$

The signals $x_x(t)$ and $y_x(t)$ are quadrature components of the noise, which themselves have Gaussian distribution, since $n_x(t)$ is Gaussian. The noise powers of the two inputs are:

$$E\left[n_1{}^2(t)\right] = E\left[x_1{}^2(t)\right] = E\left[y_1{}^2(t)\right] = \sigma_1^2 \qquad \qquad \text{...(A6.2a)}$$

$$E\left[n_2{}^2(t)\right] = E\left[x_2{}^2(t)\right] = E\left[y_2{}^2(t)\right] = \sigma_2^2 \qquad \qquad \text{...(A6.2b)}$$

For the rest of the analysis, the time dependence of the variables x, y, z and n will be dropped for simplification. There is no correlation between x_1 and y_1 or between x_2 and y_2, and the modulus of the overall correlation coefficient ρ between the inputs is defined by:

$$\rho^2 = \rho_0{}^2 + \lambda_0{}^2 \qquad \qquad \text{...(A6.3)}$$

Whereby the in-phase correlation coefficient is

$$\rho_0 = \frac{E\left[x_1 x_2\right] + E\left[y_1 y_2\right]}{2\sigma_1\sigma_2} = \frac{E\left[x_1 x_2\right]}{\sigma_1\sigma_2} = \frac{E\left[y_1 y_2\right]}{\sigma_1\sigma_2} \qquad \qquad \text{...(A6.4a)}$$

and the quadrature correlation coefficient is given by:

$$\lambda_0 = \frac{E\left[x_1 y_2\right] - E\left[y_1 x_2\right]}{2\sigma_1\sigma_2} = -\frac{E\left[y_1 x_2\right]}{\sigma_1\sigma_2} = \frac{E\left[x_1 y_2\right]}{\sigma_1\sigma_2} \qquad \qquad \text{...(A6.4b)}$$

After zonal low-pass filtering to remove double frequency components, the probability density function for a single sample of the output z of the analogue correlator is given by Brice's and Andrews's eqn. (16) (for a multiplier constant $\kappa = 1$) [Bri 82]:

$$p_1(z) = \frac{1}{\sigma_1\sigma_2\sqrt{1-\lambda_0{}^2}} \exp\left[\frac{-2z}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0{}^2} + \rho_0\right)}\right] \quad \text{for } z > 0 \qquad \qquad \text{...(A6.5a)}$$

or

$$p_1(z) = \frac{1}{\sigma_1\sigma_2\sqrt{1-\lambda_0^2}}\exp\left[\frac{2z}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}-\rho_0\right)}\right] \quad \text{for } z < 0$$

...(A6.5b)

The subsequent mathematics is simplified by introducing a change of variables. Writing

$$\alpha = \frac{2}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}+\rho_0\right)}$$

...(A6.6a)

and

$$\beta = \frac{2}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}-\rho_0\right)}$$

...(A6.6b)

the pdf for one sample can be rewritten as:

$$p_1(z) = \frac{\beta\,\alpha}{\beta+\alpha}\exp[-\alpha z] \quad \text{for } z > 0$$

...(A6.7a)

or

$$p_1(z) = \frac{\beta\,\alpha}{\beta+\alpha}\exp[\beta z] \quad \text{for } z < 0$$

...(A6.7b)

A6.2   PDF FOR THE ADDITION OF TWO INDEPENDENT SAMPLES

The PDF for the summation of two independent samples is obtained by convolving $p_1(z)$ with itself. Thus

$$p_2(z) = \int_{-\infty}^{\infty} p_1(x)p_1(z-x)dx$$

...(A6.8)

For positive values of the output z, this becomes

$$p_2(z) = \frac{(\beta\,\alpha)^2}{(\beta+\alpha)^2}\exp(-\alpha z)\left[\int_{-\infty}^{0}\exp[(\beta+\alpha)x]dx + \int_{0}^{z}dx + \int_{x=z}^{\infty}\exp[(\beta+\alpha)(z-x)]dx\right]$$

$$= \frac{(\beta\,\alpha)^2}{(\beta+\alpha)^3}\exp(-\alpha z)\,[2+(\beta+\alpha)z] \quad \text{for } z>0$$

...(A6.9a)

Similar calculations for negative z give:

$$p_2(z) = \frac{(\beta\,\alpha)^2}{(\beta+\alpha)^3}\exp(\beta z)\,[2-(\beta+\alpha)z] \quad \text{for } z<0$$

...(A6.9b)

A6.3 PDF FOR THE ADDITION OF n INDEPENDENT SAMPLES

Further convolutions involve complete Gamma function type integrals such as:

$$\int_{0}^{\infty} v^{m-1}\exp(-av)dv = \frac{\Gamma(m)}{a^m} = \frac{(m-1)!}{a^m} \qquad ([\text{Bro }83],\ \S1.1.3.4.\ \text{No. }1)$$

and

$$\int_{-\infty}^{0} v^{m-1}\exp(av)dv = \frac{(-1)^{m-1}(m-1)!}{a^m}$$

After continuing the convolution process for the first few values of n, it is possible to deduce a general expression for the PDF for the summation of n independent samples of the form:

$$p_n(z) = \frac{(\beta\,\alpha)^n}{(\beta+\alpha)^{2n-1}}f_n(|z|)\begin{cases}\exp(-\alpha z) & \text{for } z>0\\ \exp(\beta z) & \text{for } z<0\end{cases}$$

...(A6.10)

where $f_n(z)$ is a polynomial of degree n-1, i.e.:

$$f_n(z) = a_0 + a_1[(\beta + \alpha)z] + a_2[(\beta + \alpha)z]^2 + \ ...a_{n-1}[(\beta + \alpha)z]^{n-1}$$

Further inspection suggests that the coefficient corresponding to $[(\beta + \alpha)z]^k$ is:

$$a_k = \frac{(2n - 2 - k)!}{k! \ (n - k - 1)! \ (n-1)!}$$

In particular

$$a_0 = \frac{(2n - 2)!}{(n-1)! \ (n-1)!}, \quad a_1 = \frac{a_0}{2}, \quad a_{n-1} = \frac{1}{(n-1)!}$$

It is now possible to write $f_n(|z|)$ in the following form:

$$f_n(|z|) = \sum_{k=0}^{n-1} \frac{(2n - 2 - k)! \ [(\beta + \alpha)|z|]^k}{k! \ (n - 1 - k)! \ (n-1)!} \quad\quad ...(A6.11)$$

Milne verified by induction the formula for the addition of n independent samples (eqn. (A6.10)).

A6.5 PROBABILITY OF EXCEEDING A THRESHOLD

The probability of the output exceeding a positive threshold T is:

$$P(z > T) = \int_{T}^{\infty} p_n(z)dz \quad\quad ...(A6.12)$$

which becomes with eqn. (A6.10) and eqn. (A6.11):

$$P(z > T) = \frac{(\beta\,\alpha)^n}{(\beta + \alpha)^{2n-1}} \int_{T}^{\infty} \exp(-\alpha z) \sum_{k=0}^{n-1} \frac{(2n - 2 - k)! \ [(\beta + \alpha)|z|]^k}{k! \ (n - 1 - k)! \ (n-1)!} \, dz \quad\quad ...(A6.13)$$

The integral can be evaluated as a series of incomplete Gamma functions

$$Q(m,y) = \int_y^\infty \frac{v^{m-1}}{(m-1)!} \exp(-v) \, dv = \frac{1}{\Gamma(m)} \int_y^\infty v^{m-1} \exp(-v) \, dv$$

$$= \frac{\Gamma(m,y)}{\Gamma(m)} = \exp(-y) \sum_{k=0}^{m-1} \frac{y^k}{k!}$$

The above transformations have been made using relations 6.1.6, 6.5.3, 26.4.19 and 26.4.21 in Abramowitz and Stegun [Abr 70]. The result has the following form:

$$P(z > T) = \left(\frac{\beta}{\beta+\alpha}\right)^n \sum_{k=0}^{n-1} \left(\frac{\alpha}{\beta+\alpha}\right)^{n-k-1} \frac{(2n-2-k)!}{(n-1-k)! \, (n-1)!} \, Q(k+1,\alpha T)$$

$$...(A6.14a)$$

Alternatively, writing $p = (n-1-k)$, the result becomes:

$$P(z > T) = \left(\frac{\beta}{\beta+\alpha}\right)^n \sum_{p=0}^{n-1} \left(\frac{\alpha}{\beta+\alpha}\right)^p \frac{(n-1+p)!}{(p)! \, (n-1)!} \, Q(n-p,\alpha T)$$

$$...(A6.14b)$$

A6.6 PROOF THAT MILNE'S AND ANDREWS'S EXPRESSIONS FOR THE PDF AND CDF ARE IDENTICAL

Andrews and Brice derive in their paper expressions for the pdf and cdf (cumulative distribution function) at the output of an analogue cross-correlator after the summation of n independent samples [And 83]. For the case of no input signals, their results can be related to Milne's. The difference between the two derivations is that Andrews and Brice calculate the pdf and cdf using the characteristic function, whereas Milne repeatedly convolves the pdf for a single sample with itself. The proof that both results are identical is just given for z > 0, it can be similarly shown

for z < 0. Andrews's eqn. (43a) for z > 0 with his "starred" parameters inserted and a correction factor of $1/\sigma_1\sigma_2\sqrt{1-\lambda_0^2}$ has the following form:

$$p_n(z) = \frac{1}{(n-1)!}\exp\left[\frac{-2z}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}+\rho_0\right)}\right]$$

$$\sum_{m=0}^{n-1}\frac{(n+m-1)!\left[1-\left(\frac{\rho_0}{\sqrt{1-\lambda_0^2}}\right)^2\right]^m z^{n-m-1}}{m!(n-m-1)!4^m\left(\sigma_1\sigma_2\sqrt{1-\lambda_0^2}\right)^{n-m}} \qquad ...(A6.15)$$

With the index manipulation k = n-m-1 and exploiting that $\rho^2 = \lambda_0^2+\rho_0^2$ the previous equation becomes:

$$p_n(z) = \frac{1}{(n-1)!}\exp\left[\frac{-2z}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}+\rho_0\right)}\right]$$

$$\sum_{k=0}^{n-1}\frac{(2n-k-2)!\left(1-\rho^2\right)^{n-k-1}z^k}{(n-k-1)!k!4^{n-k-1}\left(\sigma_1\sigma_2\right)^{k+1}\left(\sqrt{1-\lambda_0^2}\right)^{2n-k-1}} \qquad ...(A6.16)$$

Milne's corresponding result for z > 0 is eqn. (A6.11) inserted in eqn. (A6.10) which leads to the following expression:

$$p_n(z) = \frac{(\beta\,\alpha)^n}{(\beta+\alpha)^{2n-1}}\exp\left(\frac{-2z}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}+\rho_0\right)}\right)\sum_{k=0}^{n-1}\frac{(2n-2-k)!\,(\beta+\alpha)^k z^k}{k!\,(n-1-k)!\,(n-1)!} \qquad ...(A6.17)$$

α and β are given in eqn.(A6.6). Inspection suggests that eqn. (A6.16) and eqn. (A6.17) are identical if it can be shown that:

$$\frac{(\beta\,\alpha)^n(\beta+\alpha)^k}{(\beta+\alpha)^{2n-1}} \equiv \frac{\left(1-\rho^2\right)^{n-k-1}}{\left(\sigma_1\sigma_2\right)^{k+1}\left(\sqrt{1-\lambda_0^2}\right)^{2n-k-1}4^{n-k-1}} \qquad ...(A6.18)$$

Insertion of α and β in the above equation shows that the two equations are identical. It should be noted, that the correction factor of $1/\sigma_1\sigma_2\sqrt{1-\lambda_0^2}$ for Andrews's expression of the pdf must be also used for $z < 0$.

The cdf can be derived from the pdf by evaluating the integral:

$$F(z) = \int_{-\infty}^{z} p(z)dz = 1 - \int_{z}^{\infty} p(z)dz$$

$$...(A6.19)$$

Andrews's expression for the cdf for $z > 0$ (eqn. 44a [And 83]) with his "starred" parameters inserted is:

$$F_n(z) = 1 - \left(\frac{1+\rho_0/\sqrt{1-\lambda_0^2}}{2}\right)^n \frac{1}{(n-1)!}\exp\left[\frac{-2z}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}+\rho_0^2\right)}\right]$$

$$\sum_{p=0}^{n-1}\sum_{k=0}^{n-p-1}\frac{(n-1+p)!\left(1-\dfrac{\rho_0^2}{\left(\sqrt{1-\lambda_0^2}\right)^2}\right)^p z^k}{p!\, k!\, 2^{p-k}\left(1-\dfrac{\rho_0}{\sqrt{1-\lambda_0^2}}\right)^{p+k}\left(\sigma_1\sigma_2\sqrt{1-\lambda_0^2}\right)^k}$$

$$...(A6.20)$$

Milne did not exactly derive the cdf, but the complement of it. Using eqn. (A6.19), it is possible to calculate the corresponding cumulative distribution function. With the series representation of the incomplete Gamma function inserted in eqn. (A6.14b), Milne's cdf is equal to:

$$F_n(z) = 1 - \left(\frac{\beta\,\alpha}{\beta+\alpha}\right)^n \exp\left(\frac{-2z}{\sigma_1\sigma_2\left(\sqrt{1-\lambda_0^2}+\rho_0\right)}\right)$$

$$\sum_{p=0}^{n-1}\sum_{k=0}^{n-p-1}\left(\frac{\alpha}{\beta+\alpha}\right)^p\frac{(n-1+p)!\,\alpha^k z^k}{p!\,(n-1)!\,k!}$$

$$...(A6.21)$$

Again, inspection suggests that eqn. (A6.20) and eqn. (A6.21) are identical, if:

$$\left(\frac{\beta\,\alpha}{\beta+\alpha}\right)^n \equiv \left(\frac{1+\rho_0 \big/ \sqrt{1-\lambda_0^2}}{2}\right)^n$$

...(A6.22a)

and:

$$\frac{\alpha^{p+k}}{(\alpha+\beta)^p} \equiv \frac{\left(1-\dfrac{\rho_0^2}{\left(\sqrt{1-\lambda_0^2}\right)^2}\right)^p}{2^{p-k}\left(1+\dfrac{\rho_0}{\sqrt{1-\lambda_0^2}}\right)^{p+k}\left(\sigma_1\sigma_2\sqrt{1-\lambda_0^2}\right)^k}$$

...(A6.22b)

Inserting the expressions for α and β in the above equations and after some algebra, it can be shown that relations (A6.22a) and (A6.22b) hold.

The expressions for both the pdf and cdf of the sum of n sampled outputs of an analogue cross-correlator obtained by Andrews et al. and Milne are therefore identical.

A7 MEAN, VARIANCE AND CORRELATION COEFFI-CIENT OF THE Σ AND Δ BASEBAND SIGNALS

A7.1 PRELIMINARY REMARKS

This appendix contains the derivation of the mean and variance of the output of the Σ and Δ cross-correlator of the symmetric-pair DF system after integration of n independent samples. The correlation coefficient of the Σ and Δ baseband signal will also be derived, and some analysis that cross checks the expressions obtained for the mean and variance is given in the last section of this appendix.

The output signals of the Σ and Δ correlators have been derived in chapter 5 and are given by eqn. (5.21) and eqn. (5.22). The Σ output has the following form:

$$z_\Sigma(t) = \frac{1}{8} \left[2A^2 \cos(\phi) + 2A \cos(\phi) n_{c3}(t) + A\big(n_{c1}(t) + n_{c2}(t)\big) \right.$$
$$\left. + \big(n_{c1}(t) + n_{c2}(t)\big)n_{c3}(t) + \big(n_{s1}(t) + n_{s2}(t)\big)n_{s3}(t) \right] \qquad ...(A7.1)$$

and the Δ output is given by:

$$z_\Delta(t) = \frac{1}{8} \left[2A^2 \sin(\phi) + 2A \sin(\phi) n_{c3}(t) + A\big(n_{s1}(t) - n_{s2}(t)\big) \right.$$
$$\left. + \big(n_{s1}(t) - n_{s2}(t)\big)n_{c3}(t) + \big(n_{c2}(t) - n_{c1}(t)\big)n_{s3}(t) \right] \qquad ...(A7.2)$$

All the Gaussian distributed quadrature noise components are uncorrelated and hence independent of each other [Cla 85]. To simplify the calculation of the mean and variance, it is therefore sensible to rearrange the terms of the correlator outputs in groups of mutually independent

processes. After the separation of the variables, the Σ and Δ outputs have the following generalised form:

$$z_x(t) = \frac{1}{8}\left[\underbrace{a_x}_{I} + \underbrace{b_x X_x + c_x U_x + X_x U_x}_{II} + \underbrace{Y_x V_x}_{III}\right]$$

...(A7.3)

where the index x denotes whether it is the Σ or the Δ signal. The definition of the constants a, b, c and the variates X, Y, U, V is as follows:

$$a_\Sigma = 2A^2\cos(\phi), \qquad a_\Delta = 2A^2\sin(\phi),$$
$$b_\Sigma = 2A\cos(\phi), \qquad b_\Delta = 2A\sin(\phi),$$
$$c_\Sigma = A, \qquad c_\Delta = A,$$
$$X_\Sigma = n_{c3}(t), \qquad X_\Delta = n_{c3}(t),$$
$$Y_\Sigma = n_{s3}(t), \qquad Y_\Delta = n_{s3}(t),$$
$$U_\Sigma = n_{c1}(t) + n_{c2}(t), \qquad U_\Delta = n_{s1}(t) - n_{s2}(t),$$
$$V_\Sigma = n_{s1}(t) + n_{s2}(t), \qquad V_\Delta = n_{c2}(t) - n_{c1}(t),$$

The following points are a summary of useful rules and relations, which have been applied to calculate the mean, variance and correlation coefficient of the Σ and Δ baseband signals:

(1) The expectation of any of the variates X, Y, U and V is zero, since they have zero mean.

(2) Since the variates X, Y, U and V have zero mean, the second moment and the second central moment are identical. (The second central moment is by definition the variance of a process.)

(3) The following rules for calculations with random variables are used [Cla 85]:

$E[aX + b] = aE[X] + b$ (holds always)

$E[X + Y + ...] = E[X] + E[Y] + ...$ (holds always)

$E[X \cdot Y \cdot ...] = E[X] \cdot E[Y] \cdot ...$ (only for *independent* variates)

$\sigma^2_{X+Y+...} = \sigma^2_X + \sigma^2_Y + ...$ (only for *independent* variates)

A7.2 MEAN AND VARIANCE OF THE Σ AND Δ CROSS-CORRELATOR OUTPUT AFTER SUMMATION OF N INDEPENDENT SAMPLES

Applying the above rules, the mean and variance of the three independent random processes in eqn. (A7.3) can now be calculated as follows:

Process I:

Since the process of type I is a constant, its mean and variance are:

$$m_I = E[a_x] = a_x \qquad\qquad ...(A7.4a)$$

$$\sigma_I^2 = E\left[(a_x - E[a_x])^2\right] = 0 \qquad\qquad ...(A7.4b)$$

Process II:

$$m_{II} = E[b_x X_x + c_x U_x + X_x U_x] = 0 \qquad\qquad ...(A7.5a)$$

$$\begin{aligned}
\sigma_{II}^2 &= E\left[\left((b_x X_x + c_x U_x + X_x U_x) - E[b_x X_x + c_x U_x + X_x U_x]\right)^2\right] \\
&= E\left[\left(b_x(X_x - E[X_x]) + c_x(U_x - E[U_x]) + (X_x U_x - E[X_x U_x])\right)^2\right] \\
&= b_x^{\,2} E\left[X_x^{\,2}\right] + c_x^{\,2} E\left[U_x^{\,2}\right] + E\left[X_x^{\,2}\right] E\left[U_x^{\,2}\right] \\
&= b_x^{\,2}\sigma_{X_x}^2 + c_x^{\,2}\sigma_{U_x}^2 + \sigma_{X_x}^2 \sigma_{U_x}^2 \qquad\qquad ...(A7.5b)
\end{aligned}$$

Process III:

$$m_{III} = E[Y_x V_x] = 0 \qquad\qquad ...(A7.6a)$$

$$\begin{aligned}
\sigma_{III}^2 &= E\left[(Y_x V_x - E[Y_x V_x])^2\right] \\
&= E\left[Y_x^{\,2}V_x^{\,2} - 2Y_x V_x E[Y_x]E[V_x] + (E[Y_x]E[V_x])^2\right] \\
&= E\left[Y_x^{\,2}\right] E\left[V_x^{\,2}\right] \\
&= \sigma_{Y_x}^2 \sigma_{V_x}^2 \qquad\qquad ...(A7.6b)
\end{aligned}$$

Utilising the independence of the three separated processes, the mean and variance of the cross-correlator output are then:

$$m_x = m_I + m_{II} + m_{III} \qquad \qquad ...(A7.7a)$$

$$\sigma_x^2 = \sigma_I^2 + \sigma_{II}^2 + \sigma_{III}^2 \qquad \qquad ...(A7.7b)$$

After the summation of n independent samples, the mean and variance are equal to:

$$m_{nx} = \sum_{i=1}^{n} m_x = n\, m_x \qquad \qquad ...(A7.8a)$$

$$\sigma_{nx}^2 = \sum_{i=1}^{n} \sigma_x^2 = n\, \sigma_x^2 \qquad \qquad ...(A7.8b)$$

Resubstituting the actual values of the constants and variates into eqns. (A7.8a/b), the mean and variance of the Σ cross-correlator output become:

$$m_{nz_\Sigma} = \frac{n}{4} A^2 \cos(\phi) \qquad \qquad ...(A7.9a)$$

$$\sigma_{nz_\Sigma}^2 = \frac{n}{64} \left[A^2 \left(4\cos^2(\phi)\, \sigma_{n_{c3}}^2 + \sigma_{n_{c1}}^2 + \sigma_{n_{c2}}^2 \right) \right.$$
$$\left. + \sigma_{n_{c3}}^2 \left(\sigma_{n_{c1}}^2 + \sigma_{n_{c2}}^2 \right) + \sigma_{n_{s3}}^2 \left(\sigma_{n_{s1}}^2 + \sigma_{n_{s2}}^2 \right) \right] \qquad ...(A7.9b)$$

Assuming a noise power of σ^2 at all antenna elements, the variances of the quadrature components simplify to:

$$\sigma^2 = \sigma_{n_{c1}}^2 = \sigma_{n_{c2}}^2 = \sigma_{n_{c3}}^2 = \sigma_{n_{s1}}^2 = \sigma_{n_{s2}}^2 = \sigma_{n_{s3}}^2$$

The variance at the cross-correlator output can therefore be modified to:

$$\sigma_{nz_\Sigma}^2 = \frac{n}{16} \left[A^2\, \sigma^2 \left(\cos^2(\phi) + 0.5 \right) + \left(\sigma^2 \right)^2 \right] \qquad \qquad ...(A7.9c)$$

For the Δ cross-correlator, the results are:

$$m_{nz_\Delta} = \frac{n}{4} A^2 \sin(\phi) \qquad \qquad ...(A7.10a)$$

$$\sigma_{nz_\Delta}^2 = \frac{n}{64} \left[A^2 \left(4\sin^2(\phi)\, \sigma_{n_{c3}}^2 + \sigma_{n_{s1}}^2 + \sigma_{n_{s2}}^2 \right) \right.$$
$$\left. + \sigma_{n_{c3}}^2 \left(\sigma_{n_{s1}}^2 + \sigma_{n_{s2}}^2 \right) + \sigma_{n_{s3}}^2 \left(\sigma_{n_{c1}}^2 + \sigma_{n_{c2}}^2 \right) \right] \qquad ...(A7.10b)$$

With equal noise power at the antenna elements, eqn. (A7.10b) reduces to:

$$\sigma_{nz_\Delta}^2 = \frac{n}{16} \left[A^2\, \sigma^2 \left(\sin^2(\phi) + 0.5 \right) + \left(\sigma^2 \right)^2 \right] \qquad \qquad ...(A7.10c)$$

A7.3 CORRELATION COEFFICIENT BETWEEN THE Σ AND Δ SIGNAL

The correlation coefficient between the Σ and Δ signal of the symmetric-pair DF system is defined as follows:

$$\rho_{\Sigma\Delta} = \frac{E\left[\left(Z_\Sigma - E[Z_\Sigma] \right) \left(Z_\Delta - E[Z_\Delta] \right) \right]}{\sqrt{\sigma_{Z_\Sigma}^2\, \sigma_{Z_\Delta}^2}} = \frac{\sigma_{Z_\Sigma Z_\Delta}}{\sqrt{\sigma_{Z_\Sigma}^2\, \sigma_{Z_\Delta}^2}} \qquad \qquad ...(A7.11)$$

This coefficient can be calculated with the results obtained in the previous section by letting n = 1. It should be noted that the above correlation coefficient ought not be confused with the sample correlation coefficient, which is defined slightly differently and used to establish the correlation between two *unknown* processes by obtaining their sample mean, variance and covariance. The sample correlation coefficient is obviously a function of the number of samples and converges in the limit for a large number of samples to the correlation coefficient defined in eqn. (A7.11). In the case investigated here, both processes are well known and the definition of eqn. (A7.11) can be applied.

With eqn. (A7.3) and eqn. (A7.4a), the covariance $\sigma_{\Sigma\Delta}$ of the two processes becomes:

$$\sigma_{Z_\Sigma Z_\Delta} = E\big[(Z_\Sigma - E[Z_\Sigma])\,(Z_\Delta - E[Z_\Delta])\big]$$

$$= E\Big[\frac{1}{8}(a_\Sigma + b_\Sigma X_\Sigma + c_\Sigma U_\Sigma + X_\Sigma U_\Sigma + Y_\Sigma V_\Sigma - a_\Sigma)$$
$$\frac{1}{8}(a_\Delta + b_\Delta X_\Delta + c_\Delta U_\Delta + X_\Delta U_\Delta + Y_\Delta V_\Delta - a_\Delta)\Big]$$

Applying the rules given in section A7.1, the expression for the covariance reduces to:

$$\sigma_{Z_\Sigma Z_\Delta} = \frac{b_\Sigma b_\Delta}{64} E[X_\Sigma X_\Delta] = \frac{b_\Sigma b_\Delta}{64} E\Big[\big(n_{c3}(t)\big)^2\Big] = \frac{b_\Sigma b_\Delta}{64}\,\sigma_{n_{c3}}^2$$

After replacement of the constants by their actual values and for equal noise power at the antenna elements, the numerator of eqn. (A7.11) is given by:

$$\sigma_{Z_\Sigma Z_\Delta} = \frac{2A\cos(\phi)\,2A\sin(\phi)}{64}\sigma^2 = \frac{A^2}{32}\sin(2\phi)\sigma^2 \qquad \qquad \text{...(A7.12)}$$

The denominator of the correlation coefficient can be calculated with the help of eqn. (A7.9c) and eqn. (A7.10c), letting n = 1:

$$\sqrt{\sigma_{Z_\Sigma}^2\,\sigma_{Z_\Sigma}^2} =$$

$$\sqrt{\frac{1}{16}\Big[A^2\,\sigma^2\big(\cos^2(\phi)+0.5\big)+\big(\sigma^2\big)^2\Big]\,\frac{1}{16}\Big[A^2\,\sigma^2\big(\sin^2(\phi)+0.5\big)+\big(\sigma^2\big)^2\Big]}$$
$$\text{...(A7.13)}$$

Inserting eqn. (A7.12) and eqn. (A7.13) into eqn. (A7.11), the correlation coefficient $\rho_{\Sigma\Delta}$ has the following form:

$$\rho_{\Sigma\Delta} = \frac{0.5\,A^2\sin(\phi)}{\sqrt{\Big[A^2\,\big(\cos^2(\phi)+0.5\big)+\sigma^2\Big]\Big[A^2\,\big(\sin^2(\phi)+0.5\big)+\sigma^2\Big]}} \qquad \text{...(A7.14)}$$

A7.4 CROSS CHECK OF THE EXPRESSIONS OBTAINED FOR THE MEAN AND VARIANCE OF THE CORRELATOR OUTPUT

To test the correctness of the formulae for the mean and variance, the signal-to-noise ratio at the cross-correlator output is calculated using two different approaches. The first method derives the signal-to-noise ratio from the statistics of the signals, which means it uses the expressions for the mean and variance. The second approach is to utilise the formula for the signal-to-noise ratio of the output of a phase detector (see chapter 4), which has been derived in the time domain. The proof will just be undertaken for the Σ signal, the calculations for the Δ signal follow the same arguments.

(1) The statistical approach:

The total power at the cross-correlator output is given by its second moment:

$$E\left[(Z_\Sigma)^2\right] = E[Z_\Sigma]^2 + \sigma_{Z_\Sigma}^2$$

...(A7.15)

As shown before, the noise at the correlator output has zero mean and the mean of the output is its signal x signal component. The noise power of the output signal is therefore its variance, and the signal power its mean squared. Hence, using eqn. (A7.9a) and eqn. (A7.9c) and letting n = 1, the signal-to-noise ratio at the cross correlator output is defined as:

$$\frac{S}{N}_{out} = \frac{E[Z_\Sigma]^2}{\sigma_{Z_\Sigma}^2} = \frac{\left(A^2/4 \ \cos(\phi)\right)^2}{\left(A^2 \ \sigma^2\left(\cos^2(\phi)+0.5\right)+\left(\sigma^2\right)^2\right)/16}$$

...(A7.16)

With the signal-to-noise ratio of the antenna elements defined as

$$\frac{S}{N}_{ant} = \frac{A^2/2}{\sigma^2} ,$$

...(A7.17)

eqn. (A7.16) for the signal-to-noise ratio at the correlator output can be modified to:

$$\frac{S}{N_{out}} = \frac{4\left(\frac{S}{N_{ant}}\right)^2 \cos^2(\phi)}{\frac{S}{N_{ant}}\left(1+2\cos^2(\phi)\right)+1}$$

...(A7.18)

(2)   The time domain approach:

The signal-to-noise ratio derived in the time domain of the output of a phase detector (cross-correlator) for uncorrelated input noise is given by eqn. (4.11) in chapter 4:

$$\frac{S}{N_{out}} = \frac{2\frac{S}{N_{in1}}\frac{S}{N_{in2}}\cos^2(\phi_1 - \phi_2)}{\frac{S}{N_{in1}} + \frac{S}{N_{in2}} + 1}$$

...(A7.19)

where S/N_1 and S/N_2 are the signal-to-noise ratios at the inputs of the cross-correlator and ϕ_1 and ϕ_2 the respective relative phases. In the case of a 4 channel symmetric-pair DF system, as introduced in chapter 4.4, the signal-to-noise ratio of the Σ input to the correlator is:

$$\frac{S}{N_\Sigma} = \frac{S}{N_{in1}} = \frac{\left(\sqrt{2}A\cos(\phi)\right)^2 / 2}{\sigma^2} = \frac{S}{N_{ant}} 2\cos^2(\phi)$$

...(A7.20)

The signal-to-noise ratio at the reference element input is:

$$\frac{S}{N_r} = \frac{S}{N_{in2}} = \frac{S}{N_{ant}}$$

...(A7.21)

Since the signals of the symmetric-pair and the reference element are ideally always just in phase or anti phase, $\phi_1 - \phi_2$ is only 0 or π and $\cos^2(\phi_1 - \phi_2)$ consequently always one.

After insertion of eqn. (A7.20) and eqn. (A7.21) into eqn. (A7.19), the following expression for the signal-to-noise ratio at the cross-correlator output can be obtained:

$$\frac{S}{N}_{out} = \frac{4\left(\frac{S}{N}_{ant}\right)^2 \cos^2(\phi)}{\frac{S}{N}_{ant}\left(1 + 2\cos^2(\phi)\right) + 1}$$

...(A7.22)

Comparison of eqn. (A7.18) and eqn. (A7.22) shows that the two expressions for the signal-to-noise ratios are identical. The formulae for the mean and variance of the cross-correlator output are therefore correct.

A8 DEGRADATION OF THE NOISE PERFORMANCE OF A SIGNAL SQUARER

The analysis presented in this appendix is based on similar work done by Brennan [Bre 90]. Fig. A8-1 shows the input and output signals of the squarer.
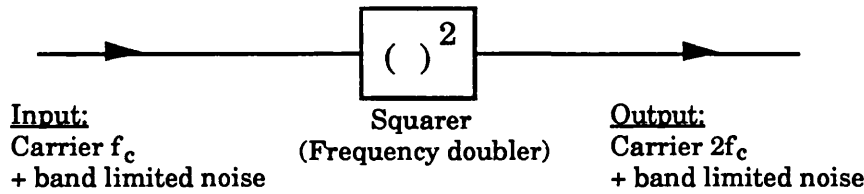

Input: Squarer   Output:
Carrier f_c (Frequency doubler) Carrier $2f_c$
+ band limited noise   + band limited noise

Figure A8-1: Signals of the squarer

The input to the squarer consists of a carrier signal at f_c with amplitude A_c and band limited noise around the carrier frequency with bandwidth B_n and a double sided noise power spectral density $\eta/2$. The input signal $s_{in}(t)$ can be written as:

$$s_{in}(t) = A_c \cos(\omega_c t) + \sqrt{2\eta B_n} \cos(\omega_c t + \phi_n(t))$$

...(A8.1)

The input signal-to-noise ratio is therefore:

$$\frac{S}{N}_{in} = \frac{\left(A_c/\sqrt{2}\right)^2}{\left(\sqrt{2\eta B_n}/\sqrt{2}\right)^2} = \frac{A_c^2}{2\eta B_n}$$

...(A8.2)

The squared output signal can then be written as:

$$s_{out}(t) = \left(A_c \cos(\omega_c t)\right)^2 + 2A_c\sqrt{2\eta B_n} \cos(\omega_c t + \phi_n(t))\cos(\omega_c t) +$$

$$2\eta B_n \left(\cos(\omega_c t + \phi_n(t))\right)^2 \quad ...(A8.3)$$

Ignoring the 'lower sideband' DC terms, eqn. (A8.3) becomes:

$$S_{out}(t) = \frac{A_c}{2}\cos(2\omega_c t) + A_c\sqrt{2\eta B_n}\left[\cos(2\omega_c t + \phi_n(t)) + \cos(\phi_n(t))\right] +$$
$$\eta B_n \cos(2\omega_c t + 2\phi_n(t)) \quad ...(A8.4)$$

The output signal power is now equal to:

$$S_{out} = \frac{A_c{}^4}{8} \qquad ...(A8.5)$$

and the output noise power is given as:

$$N_{out} = A_c{}^2 \eta B_n + (\eta B_n)^2 \qquad ...(A8.6)$$

The first term in eqn. (A8.6) is the contribution from the two signal x noise cross products in the output signal of the squarer and the second term is the noise x noise component.

With the help of eqn. (A8.2) it is possible to write the signal-to-noise ratio at the output of the squarer in terms of the input signal-to-noise ratio:

$$\frac{S}{N}_{out} = \frac{A_c{}^4/8}{A_c{}^2 \eta B_n + (\eta B_n)^2} = \frac{A_c{}^4/8}{\dfrac{A_c{}^4}{2S/N_{in}} + \dfrac{A_c{}^4}{4(S/N_{in})^2}} = \frac{(S/N_{in})^2}{4 S/N_{in} + 2} \qquad ...(A8.7)$$

It can be seen from eqn. (A8.7) that there is a 6 dB degradation in signal-to-noise ratio for reasonably strong input signals:

$$\frac{S}{N}_{out} \approx \frac{S/N_{in}}{4} \qquad ...(A8.8)$$

The extra 3 dB loss in signal-to-noise ratio, compared to a mixer with equal signal-to-noise ratio at its inputs, is due to the fact that the noise sources of the squarer are correlated which entails the total loss of 6 dB.

As a corollary, it is interesting to see for which signal-to-noise ratio at the input of the squarer the two noise terms in eqn. (A8.6) are equal. With eqn. (A8.2), eqn. (A8.6) can be modified to:

$$N_{out} = 2\frac{S}{N_{in}}\left(\eta B_n\right)^2 + \left(\eta B_n\right)^2 \qquad \text{...(A8.9)}$$

Inspection of eqn. (A8.9) shows that the two noise terms are equal for an input signal-to-noise ratio of 0.5 or -3 dB respectively.