„Development of a Position Measurement system for model airplane competitions“

Master thesis

by

Knauder Daniel, BSc

Institute of Microwave and Photonic Engineering,
Graz, University of Technology, Austria

Supervisor: Dipl.-Ing. Dr.techn. Schreiber Helmut
Examiner: Univ.-Prof. Dipl.-Ing. Dr.techn. MBA Wolfgang Bösch Graz, Okt. 2017

Zusammenfassung

Wendendetektion in Modellflug Wettbewerben ist ein Problem mit erheblichem Verbesserungspotential. Um die Fairness in Wettbewerben und Weltmeisterschaften zu steigern ist es wichtig den menschlichen Fehler zu beseitigen. Die Arbeit stellt einen ersten Prototypen eines RF-basierten Messsystems vor.

Abstract

Consistently detecting the turns in model airplane competitions is huge problem with a lot of room for improvement. Eliminating human error is an important point to increase the fairness in contests and world championships. This thesis proposes a first prototype for an RF-based measurement system.

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am ……

……..

(Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

……

……..

date (signature)

Table of Contents

List of figures

List of Tables

List of Abbreviations

ADC:	analog-to-digital converter
AFE:	analog frontend
AoA:	angle of arrival
BPSK:	binary phase-shift keying
CCDF:	complementary cumulative distribution function
CDF:	cumulative distribution function
CG:	complex Gaussian
CN:	complex normal
CPW:	coplanar waveguide
CPWG:	coplanar waveguide with ground
CSCG:	circularly-symmetric complex Gaussian
CST:	computer simulation technology (field simulation software)
d:	difference signal
d/s:	difference signal/sum signal (monopulse ratio)
FEM:	finite elements method
FPSL:	free space path loss
FSR:	full-scale range
I:	in-phase component
iid:	independent identically distributed
ISM:	industrial, science and medical (radio band)
MIC:	microwave integrated circuit
ML:	maximum likelihood
MLAT:	Multilateration
MMIC:	monolithic microwave integrated circuit
MoM:	method of moments
NF:	noise figure
NSD:	noise spectral density
PCB:	printed circuit board
PDF:	probability density function
Q:	quadrature-phase component
RCS:	radar cross section
rpm:	revolutions per minute
RV:	random variable
s:	sum signal
SDR:	software defined radio

SNR:	signal to noise ratio
TDOA:	time-difference of arrival
µC:	microcontroller

1 Introduction

This thesis shows a method of measuring if a model airplane passes through 2 hypothetical planes.

To show the requirements, let me explain the rules of the competition, the planes and technical stuff used.
The F5B competition contains 3 tasks. For this measurement system, we are just focused on the first one which is the distance task. To clarify this, F5B is the name of a contest for electric sailplanes.
Let's illustrate this with a diagram and some “Logs out of a Flight” provided by Gernot Tengg. He is the initiator of this project with the need of a training solution for this contest.

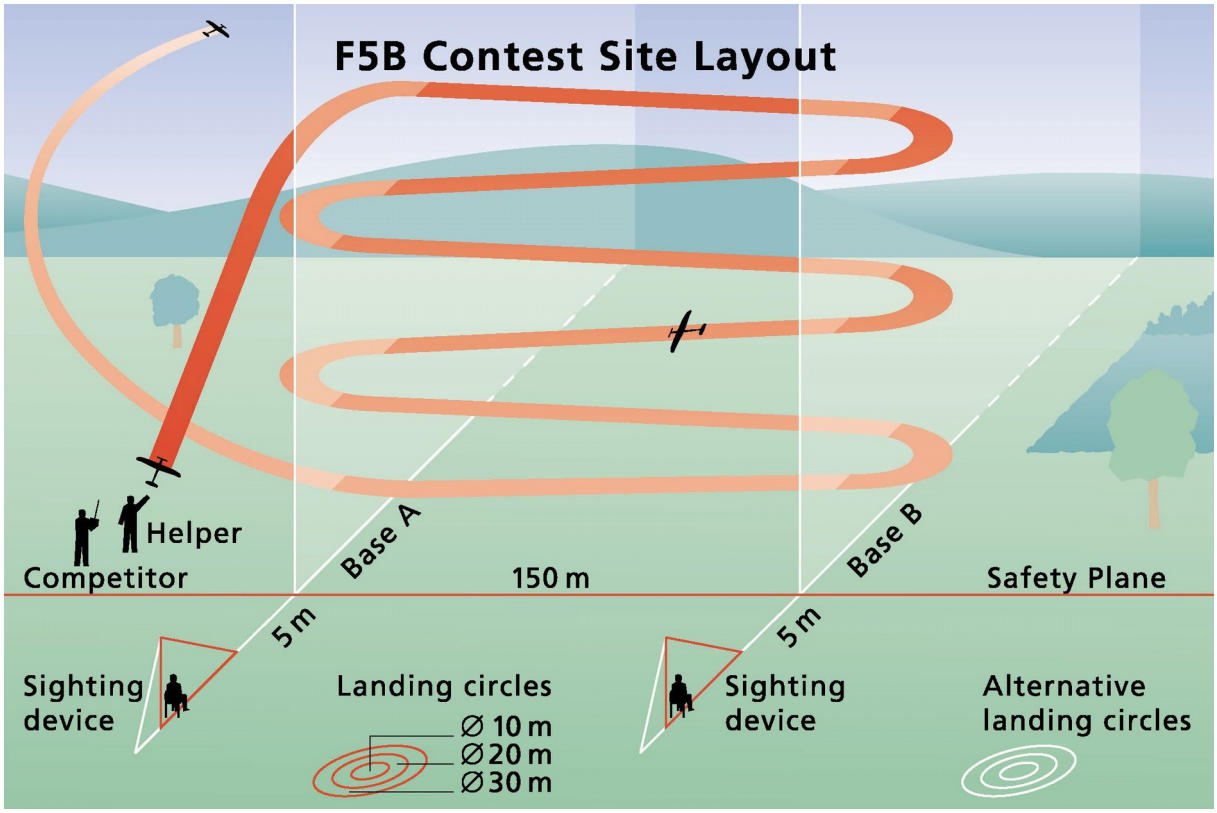

Figure 1: F5B Contest Layout (source: CIAM - FAI ([1]))

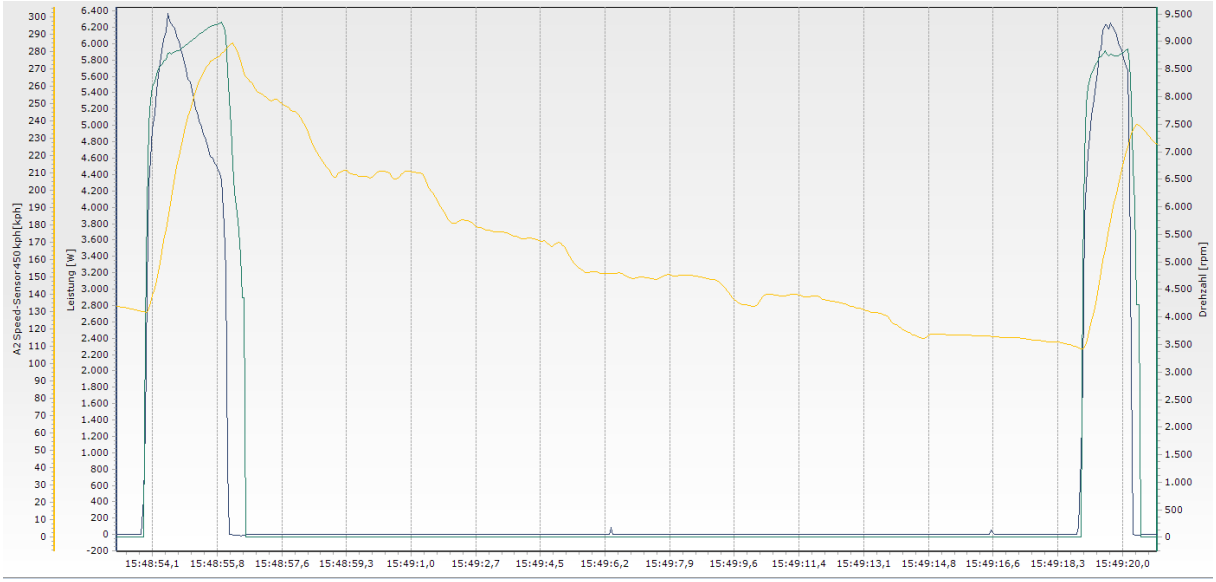

Figure 2: Logs from an F5B competition model airplane showing speed, power and revolutions per minute (rpm) over time (source: Gernot Tengg)

The model airplane uses the electric motor only outside of ***base A***. ***Base A*** and ***base B*** are spaced ***150 meters*** apart. The pilot has to complete as much "legs" as possible from ***A*** to ***B*** and vice versa in ***200 seconds***.

According to the Logs the electric motor gets used for about ***1,5*** to ***2 seconds*** to accelerate the model airplane from about ***130 km/h*** to about ***280 km/h***. So, the aircraft climbs up to about ***100*** to ***150 meters*** in about ***1,5-2 seconds*** outside of ***base A***, comes into ***A*** with up to ***300 km/h*** doing ***4*** or ***6 legs*** and then climbs again repeating the process until the ***200 seconds*** are over. [1]

At the moment, the two bases are checked by a sighting device, which basically consist of a person sitting behind two taut cords and watching if the model airplane passed the hypothetical plane.

The actual "measurement system" is not ideal in terms of angle errors and human errors. So, the aim is to get rid of the system and to provide a training possibility without the need of an additional person to look at ***base B*** (base A can be audited by the pilot himself).

2 Problems of existing systems

There are already optical training systems out there which don't work quite well in some situations.

Let's examine the problems related to the optical solution. One optical solution is called FCD-FxxTrack which is based on USB-webcams and it's problems are described in [2].

If the pilot decides to climb quite high for a so called ***"6-leg"***, there are serious problems detecting the first turn on ***base B***. The aircraft is quite far away (probably over ***150 meters***) and therefore in the same order of magnitude or even smaller than a fly real close to the camera. Additionally the model is in the periphery of the lens, which makes the situation even worse.

Another worst-case situation is if the model is flying directly in the direction of the camera. Hence, the camera system can't detect much motion and the projection surface is pretty low. [2].

So, if an optical solution causes problems, an RF-based solution might be an option. If we can track missiles with radar based solutions we should also be able to track a model airplane.

The angular resolution of a typical radar-system is mainly determined by the half-power beam-width of the antenna, so a single antenna system is insufficient.

One way to achieve an increased accuracy in angle measurement would be sequential lobing or conical scan. In principle, these two techniques use mechanical or electrical movement of the antenna beam to increase the accuracy of the system.

High angular velocities caused by the model being close and fast can cause problems for these types of systems. The angular position can differ significantly during a measurement cycle which causes these types of systems to become infeasible for the requirements given.

The solution to this is its successor called monopulse or simultaneous lobing.

In general, radar systems suffer from low received power due to high path loss and low radar cross sections (RCS) of some targets. This needs to be compensated by high transmit power in the radar system.

To overcome these issues with a small and widely fluctuating radar cross section an

active transmitter onboard the airplane with a mostly omni-directional antenna (e.g. one or two orthogonal dipoles) using one of the ISM bands can be used to achieve the power levels necessary at the receiver.

3 Concept

After defining the requirements, it is time to come up with a rough concept about how such a system is going to look like. One of the most promising techniques in this regard is the above mentioned monopulse.

In a nutshell monopulse uses the difference of two or more antenna patterns to estimate the angle of arrival (AoA). To achieve several antenna patterns a patch array can be used with the feed-network forming the two subarrays.

The monopulse principle is typically used in radar systems. However, without a radio licence the radio bands available for this application are limited to the industrial science medical (ISM) bands and the transmit power is limited. Combined with the low radar cross section of these model airplanes and the considerably higher path loss of passive radar solutions makes a passive radar-based solutions unfeasible.

Hence an active solution with a transmitter in the model airplane is the better solution.

Base A and B can be audited by a monopulse antenna and receiver at each of the bases.

The details and some of the design decisions will be covered in later chapters.

4 Monopulse

4.1 Basic Principle

[3]

First, let's talk about the monopulse principle and why this approach is so powerful.

Where does the name actually come from? The name “monopulse” comes from the fact that the angle of arrival can be estimated with a single radar pulse. This can be a huge game-changer compared to its previously mentioned predecessor conical scan or sequential lobing.

This technique offers huge advantages and very few, if any disadvantages compared to other techniques.

Let's start with the advantages. The main advantage of these systems gave the technique its name. In monopulse the angular information can be obtained from a single pulse compared to conical scan where the angular information can only be obtained after the whole cyclic scan has been completed.

The next advantage is that the monopulse principle is quite resilient to amplitude variations which could be caused by a highly fluctuating radar cross section or highly varying range. This is achieved by the normalization of the difference-voltage by the sum-voltage of the two patterns. This will be covered later in greater detail.

In tracking mode, the sum beam is directly pointed at the target rather than around the target like in conical scan. Hence, the signal to noise ratio (SNR) is greatly increased which improves detection.

In contrast, the technique has no serious drawbacks. The only drawback would be the higher effort in hardware and design and therefore the cost.

4.2 Amplitude monopulse principles

[3]

There are two different forms of monopulse. In order to understand the principle, I'd like to begin with the slightly more intuitive amplitude-comparison monopulse.

To use the amplitude-comparison monopulse technique several antenna patterns (most likely two or four) with the same phase center but slightly squinted beams (Figure 3) pointing in different directions need to be used. This could be realized using a parabolic reflector and several feed horns.

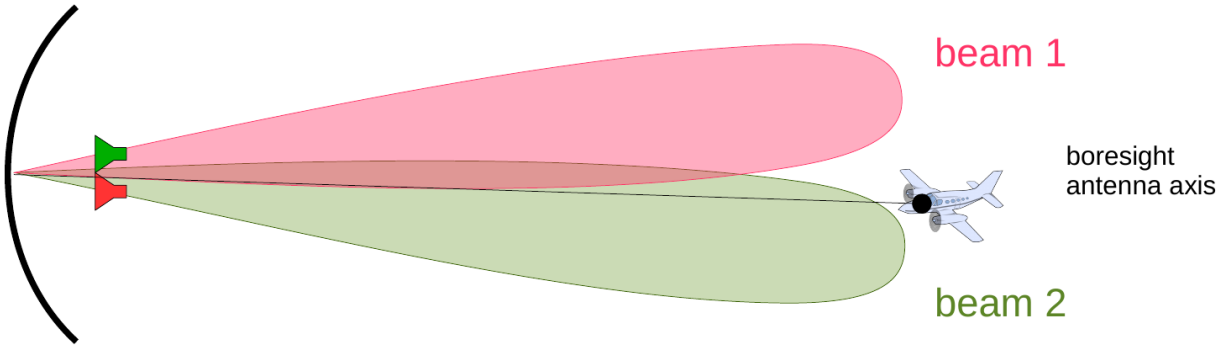

Figure 3: Squinted beams in amplitude-comparison monopulse (source: Helmut Schreiber. Introduction to radar-systems lecture notes)

By comparing the amplitudes of the two beams the angle can be estimated reasonably well.

One approach would be to use a autonomous receiver for each beam. This approach is generally called interferometry. As a matter of fact this approach is less useful compared to the monopulse principle for its designated application. The gain and phase imbalance introduced by the use of autonomous receivers for each of the beams influences zero-angle estimations quite heavily. In addition to that the imbalance can be time variant, so there would be no trivial way to compensate for that.

In order to get around these issues a magic-T, 90-degree-hybrid or a rat-race can be used to form the sum and difference patterns immediately after the antenna in the analog domain before feeding it into the receiver inputs. Hence, this would greatly reduce several issues introduced by imbalances in different receiver chains of the design.

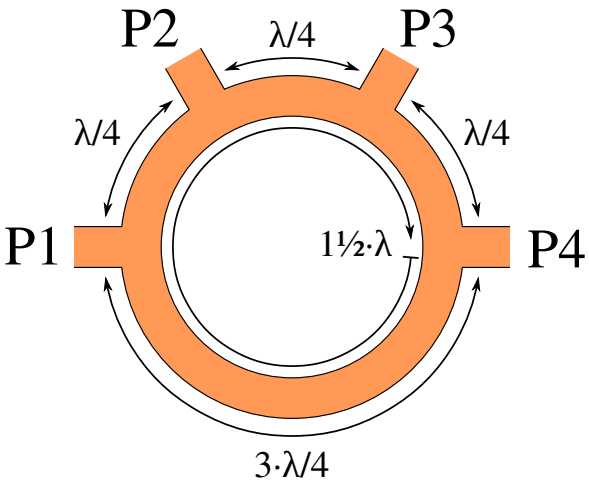

Figure 4: rat-race coupler (source: en.wikipedia.org)

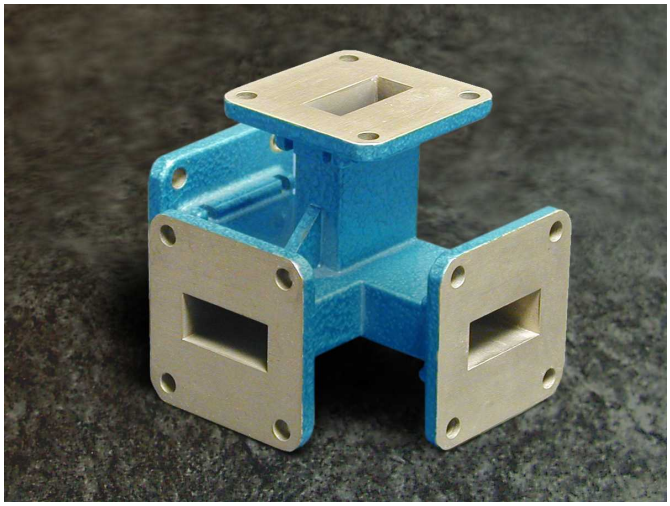

Figure 5: magic-T (source: radartutorial.eu)

The sum and difference can be computed by:

$$\text{Sum:}\quad s=\frac{1}{\sqrt{2}}(v_1+v_2) \tag{1}$$

$$\text{Difference}\quad d=\frac{1}{\sqrt{2}}[v_1-v_2] \tag{2}$$

Where v_1 and v_2 are the output voltages of the two antennas.
The factor $1/\sqrt{2}$ comes from the assumption of lossless 3 dB-splitters or hybrids in order to get the sum and differences.

Plotting the voltage pairs and the sum and difference for a monopulse system would result in:

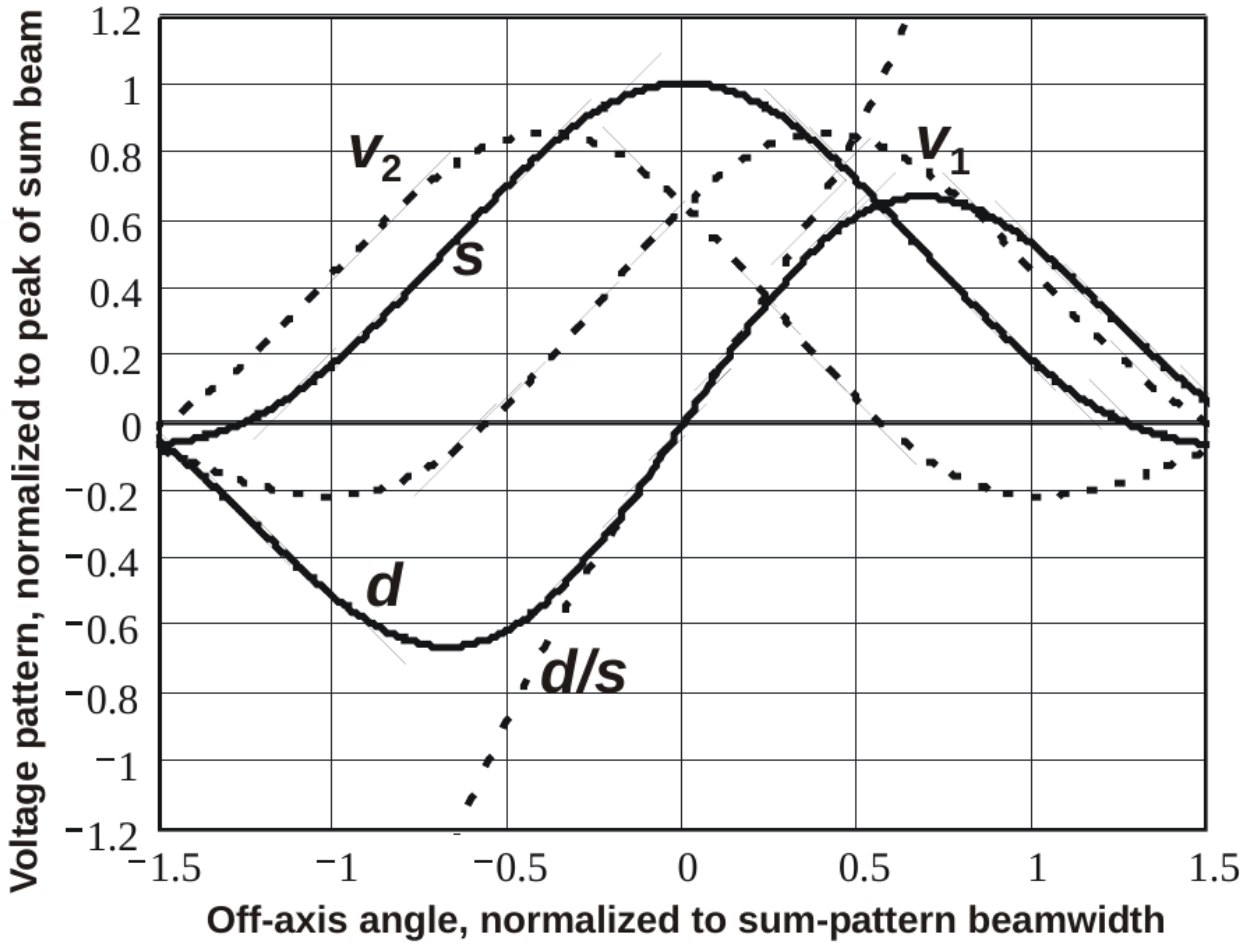

Figure 6: Amplitude-comparison monopulse patterns (source: [3])

The angle at which the corresponding difference pattern (see Figure 6) equals zero is called the monopulse axis or boresight axis. In an ideal world, this axis would be exactly in line with the geometrical axis of the antenna but in practice this may be not the case. To compensate for the non-ideal behaviour of the hardware some sort of calibration is necessary. For example simply measuring the zero-crossing of the difference-pattern and adjust the antenna direction accordingly would be one simple possibility for calibration.

In Figure 6, the ratio of the difference and sum signal (d/s) is also introduced. This quantity offers additional advantages. First it is typically quite linear for small displacements from the boresight axis, so the angle can be determined by measuring d/s and secondly the normalization by s makes it more resilient to amplitude

fluctuations. Additionally the division by the sum-pattern introduces a phase reference. Hence, the phase of the monopulse ratio is the relative phase between the difference and sum pattern. This statement will be examined later in greater detail.

The peak of the sum pattern beam should be coincident with the monopulse axis. Of course, in practice this might not truly be the case. All sorts of deviations of the two antennas would cause the sum pattern axis to be slightly off the geometrical monopulse-axis.

The sum and difference signals at RF frequency are normally downconverted to IF or even DC for further processing.

4.3 Phase-comparison monopulse

[3]

This method is quite similar to amplitude-comparison monopulse but there are some minor differences. The main difference to amplitude-comparison monopulse is that the antennas have different phase centers but the two beam axes are now in parallel. This is illustrated in Figure 7.

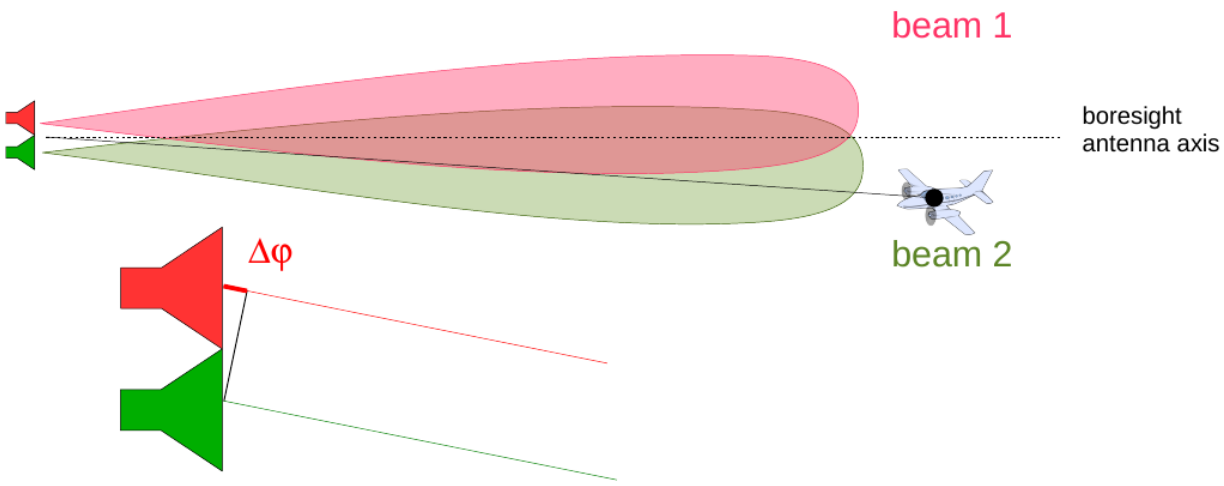

Figure 7: Phase-comparison monopulse (source: Helmut Schreiber. Introduction to radar-systems lecture notes)

In Figure 7, the plane wave in the far-field reaches the two different phase centers with different phase if the target/transmitter is off the boresight axis. Compared to amplitude-comparison monopulse, sum and difference patterns are not in-phase any more but have a phase-shift of 90-degrees. This can be easily shown by a short straightforward derivation.

In order to show this let's start with the two signals at the antennas. Let's assume a phase-shift of $\Delta\phi$ between them:

$$v_1 = A \cdot \sin(\omega_0 t) \tag{3}$$

$$v_2 = A \cdot \sin(\omega_0 t + \Delta\phi) \tag{4}$$

The sum and difference patterns are already described in (1) and (2):

$$d = (v_1 - v_2)/\sqrt{2} \tag{5}$$

$$s = (v_1 + v_2)/\sqrt{2} \tag{6}$$

Using trigonometric relations gives:

$$d = \frac{1}{\sqrt{2}}[A \cdot \sin(\omega_0 t) - A \cdot \sin(\omega_0 t + \Delta\phi)] = -\sqrt{2} \cdot A \cdot \sin(\frac{\Delta\phi}{2}) \cdot \cos(\omega_0 t + \frac{\Delta\phi}{2}) \tag{7}$$

$$s = \frac{1}{\sqrt{2}}[A \cdot \sin(\omega_0 t) + A \cdot \sin(\omega_0 t + \Delta\phi)] = \sqrt{2} \cdot A \cdot \cos(\frac{\Delta\phi}{2}) \cdot \sin(\omega_0 t + \frac{\Delta\phi}{2}) \tag{8}$$

The sine and cosine functions of the time-dependent part of the equations indicate that the sum and difference patterns of phase-comparison monopulse are orthogonal to each other.

In terms of processing this means working with the imaginary part instead of the real part or shifting the phase in one of the two signals. But without knowing the hardware used for building the sum and difference patterns there is no way to know if the phase-shift is caused by the principle (amplitude-comparison vs. phase-comparison) or the combining device itself.

For example, using 90°-hybrids for splitting and combining the phase monopulse system, the 90-degree phase-shift introduced by the hybrid would compensate for the 90-degree shift introduced by the phase-comparison principle.

4.4 The complex monopulse ratio (d/s)

[3]

In the preceding chapters the monopulse principle, differences between amplitude and phase monopulse and sum and difference signals were discussed to some extent. However the d/s ratio is important for various reasons.

The phasors of the difference and sum signal can be expressed by pointer representation:

$$d = |d| \exp(j\delta_d) \tag{9}$$

$$s = |s| \exp(j\delta_s) \tag{10}$$

where the corresponding deltas (δ_d , δ_s) indicate the phase angles to a common arbitrary reference.

Hence, the monopulse ratio can be computed:

$$\frac{d}{s} = \left|\frac{d}{s}\right| \exp[j(\delta_d - \delta_s)] = \left|\frac{d}{s}\right| \exp[j\delta] \tag{11}$$

with δ indicating the phase difference between d and s:

$$\delta = \delta_d - \delta_s \tag{12}$$

The Euler function can be expressed in terms of sine and cosine functions:

$$\frac{d}{s} = \left|\frac{d}{s}\right| \exp(j\delta) = \left|\frac{d}{s}\right| \cos(\delta) + j\left|\frac{d}{s}\right| \sin(\delta) \tag{13}$$

Therefore, the real and imaginary parts are:

$$\Re\{d/s\} = \left|\frac{d}{s}\right| \cos(\delta) \tag{14}$$

$$\Im\{d/s\} = \left|\frac{d}{s}\right| \sin(\delta) \tag{15}$$

In practice, an I/Q-demodulator is often used on the sum and difference signals. Hence, it is suitable to express the sum and difference separately in “I/Q”-form where the phase of the complex quantity is referenced to the same reference phase (e.g. the

same oscillator).

Expressing the sum and difference signal by in- and quatrature-phase components results in:

$$d_I = |d| \cos(\delta_d) = \Re\{d\} \tag{16}$$

$$d_Q = |d| \sin(\delta_d) = \Im\{d\} \tag{17}$$

$$s_I = |s| \cos(\delta_s) = \Re\{s\} \tag{18}$$

$$s_Q = |s| \sin(\delta_s) = \Im\{s\} \tag{19}$$

In complex form:

$$d = \Re\{d\} + j\Im\{d\} = d_I + jd_Q \tag{20}$$

$$s = \Re\{s\} + j\Im\{s\} = s_I + js_Q \tag{21}$$

To achieve the real and imaginary part of the d/s (monopulse) ratio, the denominator can be expanded by its complex conjugate:

$$\frac{d}{s} = \frac{ds^*}{ss^*} = \frac{(d_I + j\,d_Q)(s_I - js_Q)}{(s_I + js_Q)(s_I - js_Q)} = \frac{d_I s_I + d_Q s_Q + j(d_Q s_I - d_I s_Q)}{s_I^2 + s_Q^2} \tag{22}$$

Hence, the real and imaginary parts are:

$$\Re\{\frac{d}{s}\} = \frac{d_I s_I + d_Q s_Q}{s_I^2 + s_Q^2} \tag{23}$$

$$\Im\{\frac{d}{s}\} = \frac{d_Q s_I - d_I s_Q}{s_I^2 + s_Q^2} \tag{24}$$

The arbitrary phase reference doesn't matter in that case. The real and imaginary part of the monopulse ratio stay the same as long as the receivers for the sum and difference signal use the same oscillator and therefore use the same reference phase.

5 Design Considerations

Now that the most basic theory about monopulse is covered, it is time to take a look at some design considerations of a monopulse system. As already briefly mentioned, the angle of arrival will be estimated using a monopulse antenna and receiver. The monopulse antenna is a one-dimensional patch array consisting of two subarrays forming the two antenna beams. The monopulse antenna will be covered later in greater detail. For now it is all about the necessary SNR for a given angular accuracy.

5.1 Angular accuracy

Considering the first turn at base B should be measured with an accuracy of ***+/-1 m*** at about 170 m altitude and in a horizontal distance of 50 m.

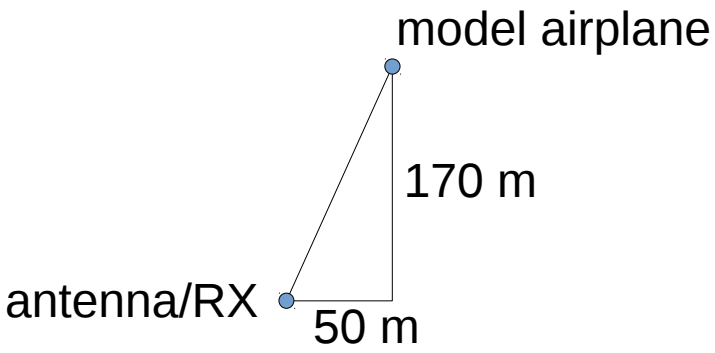

Figure 8: Schematical overview of first base B turn

From Figure 8 the diagonal distance can be calculated by

$$c = \sqrt{170^2 + 50^2} = 177{,}2\,m \approx 180\,m \tag{25}$$

From there one can easily approximate the angular accuracy:

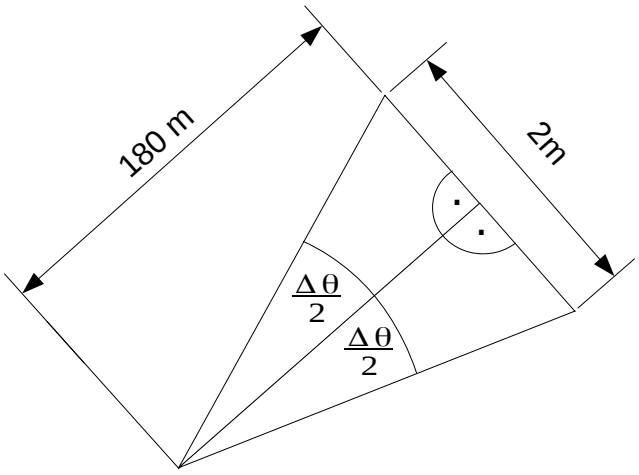

Figure 9: angle estimate

To get a simple angle estimate for small angles the angle can be approximated by a right-angled triangle for small angles:

$$\frac{\Delta\theta}{2} \approx \arctan\left(\frac{1\,m}{180\,m}\right) = 0{,}32\,° \tag{26}$$

Hence

$$\Delta\theta \approx 0{,}64\,° \tag{27}$$

The phase-shift can be estimated by the deviation from the boresight axis. To illustrate this further let's look at Figure 10:

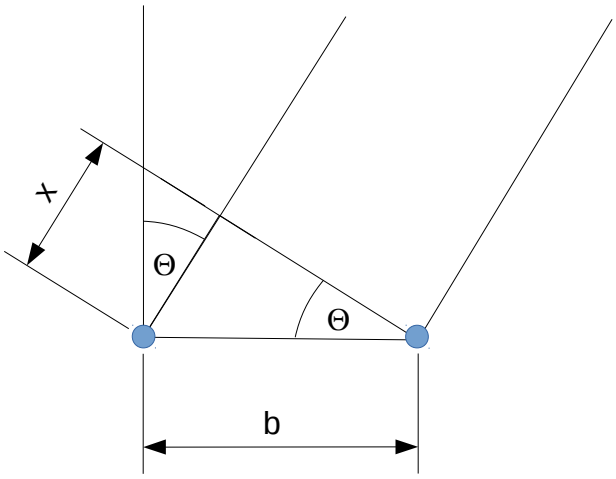

Figure 10: Deviation caused by off-boresight angle

Where Θ is the off-boresight angle, ***b*** the distance between the phase centers of the two subarrays, and ***x*** is the length difference for the two phase centers in the off-axis case.

The adjacent can be calculated by its hypotenuse:

$$\sin\Theta = \frac{x}{b} \tag{28}$$

$$x = b \cdot \sin\Theta \tag{29}$$

The size ***D*** of the two antennas was roughly estimated for a desired beamwidth:

$$D \approx 0{,}89 \cdot \frac{c}{f \cdot \Delta\phi_{3dB}} = 0{,}89 \cdot \frac{3 \cdot 10^8 \frac{m}{s}}{5{,}8 \cdot 10^9 \frac{1}{s} \cdot 12° \cdot \frac{\pi}{180°}} \approx 0{,}22\,m \tag{30}$$

From the size, the distance of the two phase centers can be estimated by:

$$d \approx \frac{D}{2} = \frac{0{,}22\,m}{2} \approx 0{,}1\,m \tag{31}$$

The distance of the phase centers can be set in relation to the wavelength to obtain the phase-shift:

$$\Delta\phi = \frac{360° \cdot x}{\lambda} = \frac{360° \cdot b \cdot \sin\Theta}{\lambda} = \frac{360° \cdot b \cdot \sin\Theta \cdot f}{c} = \frac{360° \cdot 0{,}1m \cdot \sin(0{,}32°) \cdot 5{,}8 \cdot 10^9}{3 \cdot 10^8} \tag{32}$$

$$\Delta\phi = 3{,}89°$$

From this phase-shift, the necessary SNR can be estimated. This is the maximum allowable phase-shift introduced by the deviation. This phase shift is introduced at one subarray in the simulation. Computing the cumulative distribution function (CDF) for a given SNR gives the percentage at which the desired accuracy is met.

5.2 Estimation of necessary SNR

The necessary SNR to achieve the desired angle accuracy can be estimated using simulation software like AWR.

In order to get an estimate for the necessary SNR, I introduced the phase-shift in one of the paths and tweaked the SNR value in the receiver model until the CDF reaches the desired probability at for a d/s ratio of zero. Due to the introduced phase shift, the target would be off axis by the desired maximum deviation. If the histogram shows values beyond d/s=0, the accuracy requirements are not met. Computing the CDF at a d/s ratio gives the probability at which the accuracy requirements are met. Hence, the system with this SNR satisfies the resolution requirement of +/-1 m at about 180 m distance with the given probability.

Figure 11, Figure 12 and Figure 13 show the probability that the angle error is smaller than 0,32°, which gives us a first estimation about the necessary SNR to achieve a given angular resolution. This approach only considers the effects of noise on the angular accuracy.

Other effects like multipath and in-band interference can have a significant impact on the angular accuracy. It is hard to consider these effects beforehand but later chapters will cover the effects introduced by interference and multipath and the possible countermeasures.

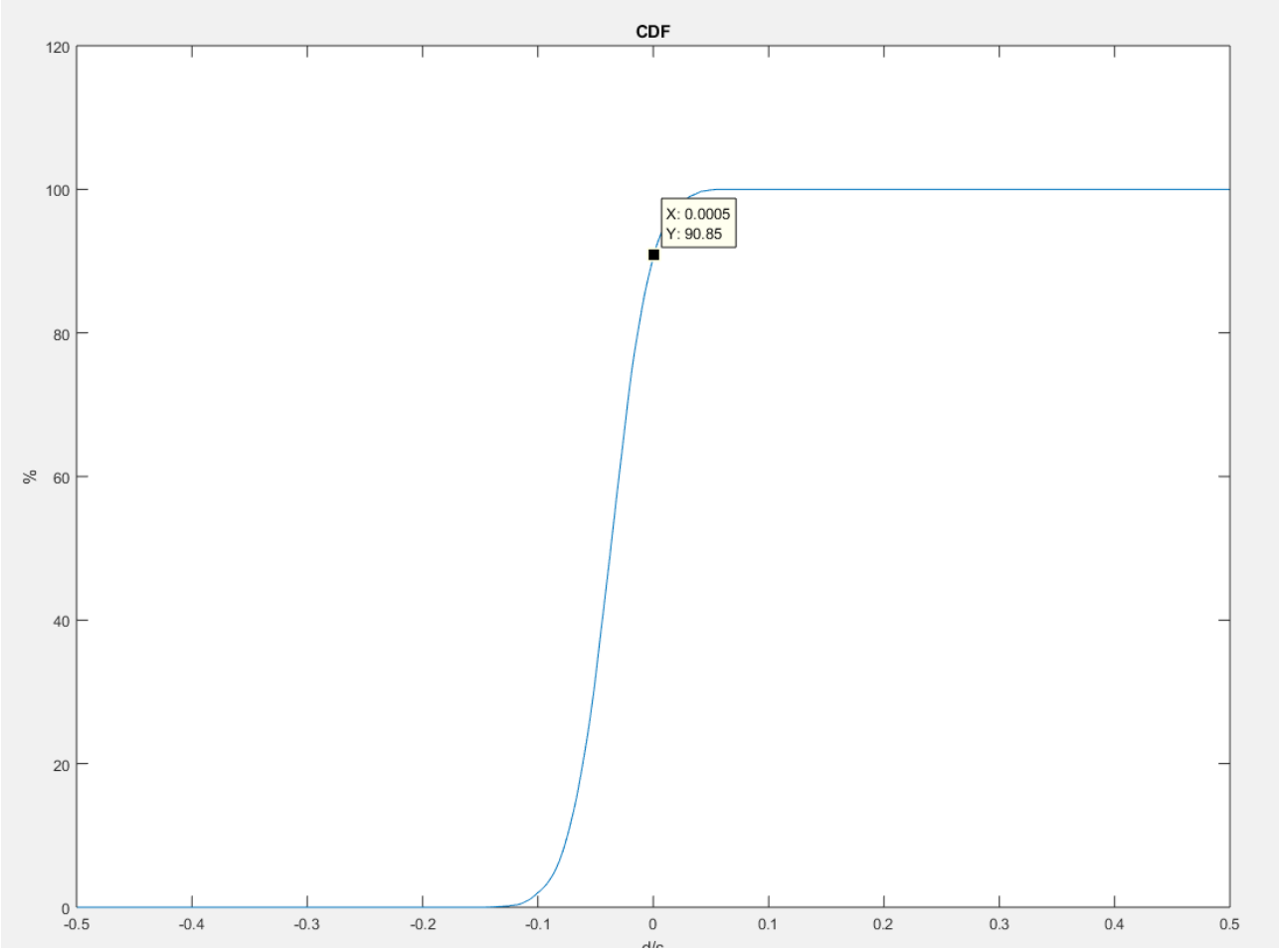

Figure 11: CDF for a SNR of 18 dB

Figure 11 shows the CDF for a SNR of 18 dB. At 18 dB SNR the resolution requirements are met with a probability of 90.85%. Hence 90.85% of the time the requirements are met.

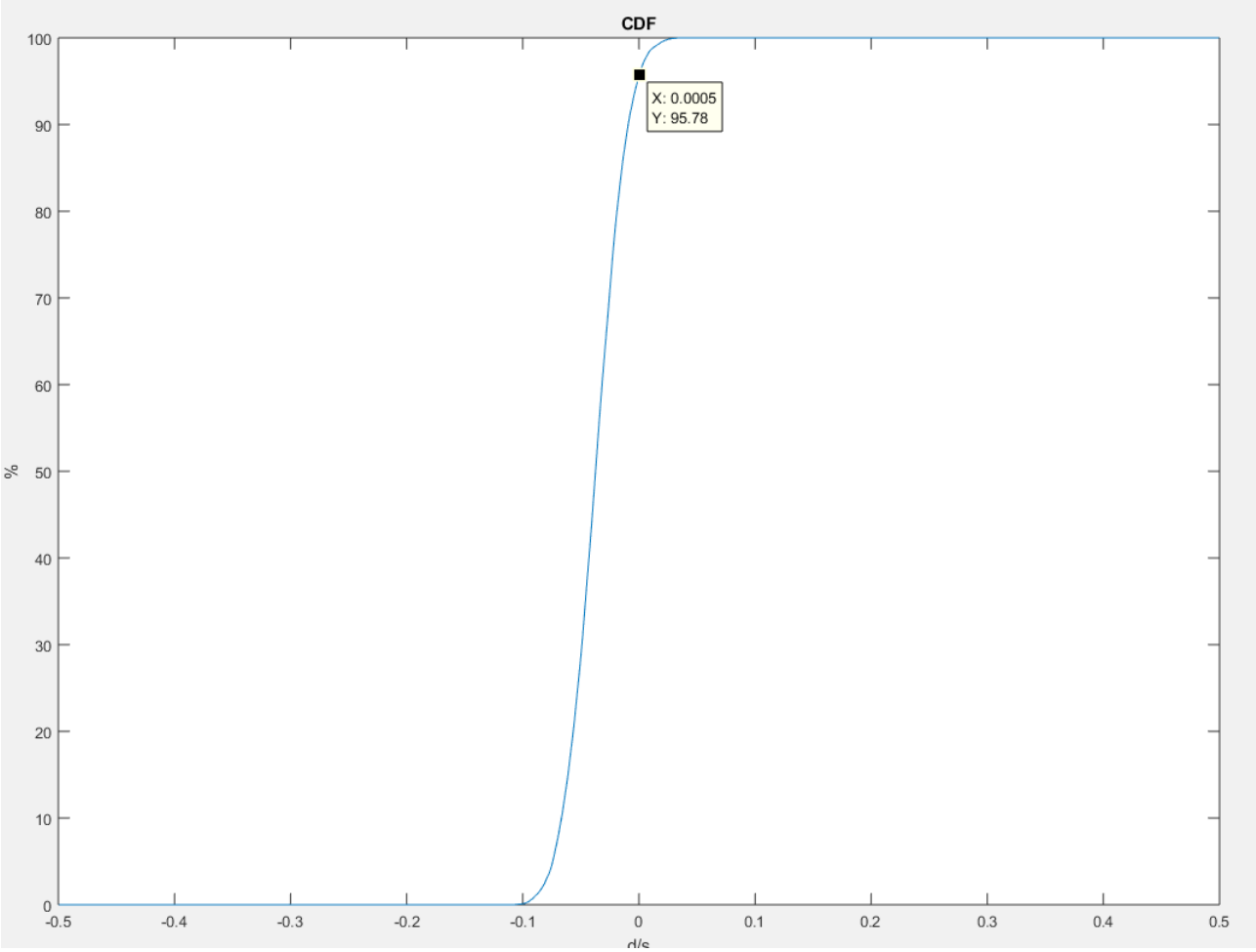

Figure 12: CDF for a SNR of 20 dB

Figure 12 shows the CDF for a SNR of 20 dB. At 20 dB SNR the resolution requirements are met with a probability of 95.78%. Hence 95.78% of the time the requirements are met.

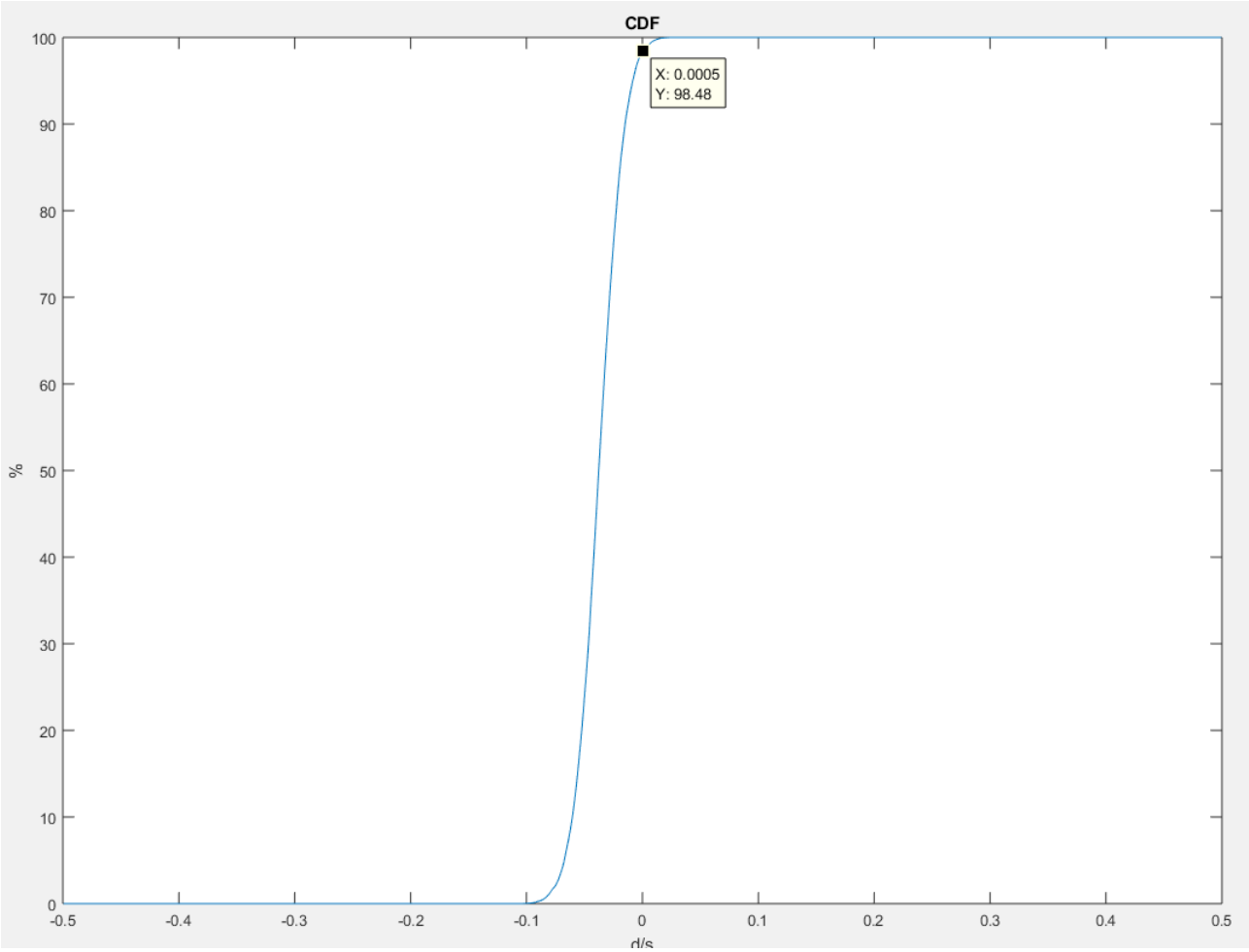

Figure 13: CDF for a SNR of 22 dB

Figure 13 shows the CDF for a SNR of 22 dB. At 22 dB SNR the resolution requirements are met with a probability of 98.48%. Hence 98.48% of the time the requirements are met.

There was no strict design decision made in terms of the necessary SNR and angular accuracy due to the fact that the noise figure of all low price point receivers are in the magnitude and their gain is variable. Hence a better noise figure can't be achieved at this low price point and the gain is still variable in software.

To further illustrate the idea behind the influence of the SNR in boresight accuracy let's take a look at the noisy linearised monopulse ratio illustrated in figure 14:

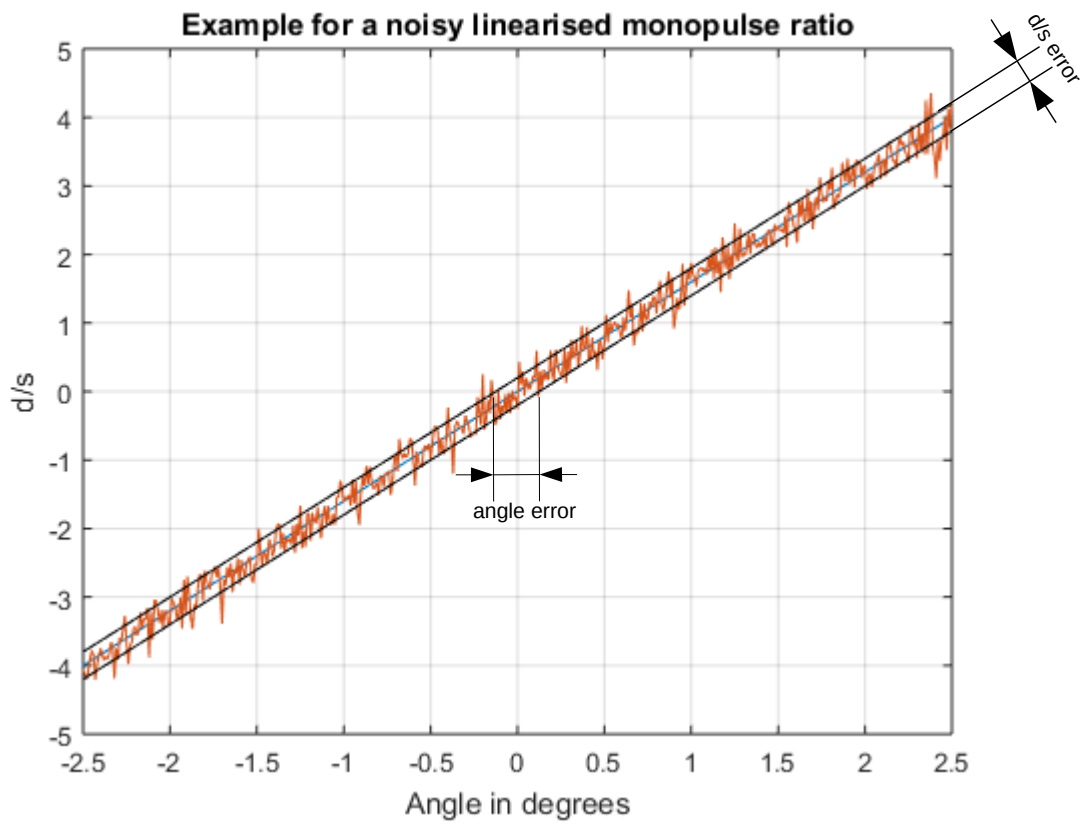

Figure 14: Noisy linearised Monopulse ratio

The noise is one of the main limiting factors for boresight angle estimations.

5.3 Phase sensitivity

[9]

The phase differences between the two antennas introduced by the off-boresight angle was derived in chapter 5.1:

$$\phi = \frac{2\pi}{\lambda} \cdot b \cdot \sin(\Theta) \tag{33}$$

The voltage phasors of the subarrays could be expressed with additional noise terms:

$$v_1 = a \cdot e^{-j\frac{\phi}{2}} + n_1 \tag{34}$$

$$v_2 = a \cdot e^{j\frac{\phi}{2}} + n_2 \tag{35}$$

where ***a*** is the amplitude.

The complex noise terms can be written as:

$$n_1 = n_{1I} + j\, n_{1Q} \tag{36}$$

$$n_2 = n_{2I} + j\, n_{2Q} \tag{37}$$

Where the variance of the distributions is:

$$\sigma_{iI}^2 = \sigma_{iQ}^2 = \frac{\sigma_{n1}^2}{2} = \frac{\sigma_{n2}^2}{2} \tag{38}$$

Assuming $\phi \ll 1\,rad$ the monopulse ratio can be roughly estimated by:

$$\frac{d}{s} = \frac{v_1 - v_2}{v_1 + v_2} = \frac{a \cdot e^{-j\frac{\phi}{2}} + n_1 - a \cdot e^{j\frac{\phi}{2}} - n_2}{a \cdot e^{-j\frac{\phi}{2}} + n_1 + a \cdot e^{j\frac{\phi}{2}} + n_2} = \frac{-2\,j \sin(\frac{\phi}{2})\, a + n_1 - n_2}{2\cos(\frac{\phi}{2})\, a + n_1 + n_2} = \frac{-j \sin(\frac{\phi}{2}) + \frac{n_1 - n_2}{2a}}{\cos(\frac{\phi}{2}) + \frac{n_1 + n_2}{2a}} \tag{39}$$

Hence, in an ideal phase monopulse system only the imaginary part of the monopulse ratio is necessary (depending on the combining network it could be a pure real part too).
Some further simplifications yield to the following result:

$$\Im\{\frac{d}{s}\} \approx -\tan(\frac{\phi}{2}) + \frac{\Im\{\frac{n_1 - n_2}{2}\}}{a \cdot \cos(\frac{\phi}{2})} \approx -\frac{\phi}{2} + \frac{\Im\{\frac{n_1 - n_2}{2}\}}{a} \tag{40}$$

Where the variance of the monopulse ratio can be obtained by:

$$\sigma_{d/s} = \frac{\sigma_{n_{\Delta Q}}}{2} = \frac{\sqrt{\sigma_{1Q}^2 + \sigma_{2Q}^2}}{2a} = \frac{\sigma_{nQ}}{a \cdot \sqrt{2}} \tag{41}$$

5.4 Free space path loss (FSPL) & Friis link budget

[4]

The FSPL is given by:

$$FSPL_{dB} = 20 \cdot \log_{10}(R) + 20 \cdot \log_{10}(f) + 20 \cdot \log_{10}\left(\frac{4\pi}{c}\right) \quad (42)$$

$$FSPL_{dB} = 20 \cdot \log_{10}(R) + 20 \cdot \log_{10}(f) - 147.55 \quad (43)$$

The frequency in the equation often causes the misconception that the free-space path loss would be frequency dependent. This is simply not the case. This equation assumes fixed antenna gain and this kind of compensates for the fact that the antenna gain increases according to the frequency.

The frequency is ***f=5,8 GHz*** and the two extreme cases are ***10 cm*** and ***300 meters***

$$FSPL_{300\,m} = 20 \cdot \log_{10}(300) + 20 \cdot \log_{10}(5.8 \cdot 10^9) - 147.55 = 97.26\,dB \quad (44)$$

For the other extreme case:

$$FSPL_{10\,cm} = 20 \cdot \log_{10}(0.1) + 20 \cdot \log_{10}(5.8 \cdot 10^9) - 147.55 = 27.72\,dB \quad (45)$$

5.4.1 Link Budget

From the FSPL, the power levels in the RF-frontend can be estimated.

Due to the regulations in the ***5,8 GHz ISM*** band the ***EIRP*** is limited by ***25 mW (14 dBm)***.

According to [4] the receive power is:

$$P_{Rx} = P_{EIRP} - FSPL + G_{Rx} \tag{46}$$

Putting in the calculated values, results in:

$$P_{Rx}(300\,m) = 14\,dBm - 97.26\,dB + G_{Rx} = -83.26\,dBm + G_{Rx} \tag{47}$$

$$P_{Rx}(0.1\,m) = 14\,dBm - 27.26\,dB + G_{Rx} = -13.72\,dBm + G_{Rx} \tag{48}$$

The gain of the receiving antenna can be approximated by:

$$G = \frac{25000}{\Delta\Phi[°]\cdot\Delta\phi[°]} = \frac{25000}{10°\cdot 90°} = 27{,}78 = 14{,}44\,dB \tag{49}$$

The power at the input of the receiver is:

$$P_{RX}(300\,m) = -83.26\,dBm + 15\,dB = -68.26\,dBm \tag{50}$$

$$P_{RX}(0.1\,m) = -13.72\,dBm + 15\,dB = 1.28\,dBm \tag{51}$$

This is the power which goes into the receiver depending on the distance of the transmitter.

6 Patch Antennas

The antenna is an important part in the monopulse systems. This chapter will cover the relevant patch antenna basics regarding monopulse systems.

6.1 Patch Antenna Principles

[8]

Patch antennas are often used if price, weight and size play an important role in the antenna design. One of the huge advantages is the possibility of integration on a PCB. So, everything can be packed onto the same substrate.

The disadvantage though is the low antenna gain caused by the relatively high losses in the substrate. [6]

A microstrip or patch antenna is basically just a very thin (compared to the wavelength) metallic strip (or patch) placed above a ground layer. The distance between the patch and ground (which is basically the thickness of the substrate) is typically in the magnitude of

$$0{,}003\lambda_0 \leq h \leq 0{,}05\lambda_0 \tag{52}$$

For 5,8 GHz the height of the substrate is generally between:

$$0{,}16\,mm \leq h \leq 2{,}58\,mm \tag{53}$$

6.2 Common Analysis Methods for Patch Antennas

[8]

There are 3 common models for analysing a patch antenna. These are the transmission line model, the cavity model and the full-wave analysis. From the three the transmission line model is by far the easiest, but lacks somewhat in accuracy.

The cavity model is kind of the middle ground of the three. It's more accurate but also more complex compared to the transmission line model.

But the transmission line and cavity model have issues in modelling coupling between elements. Of course, the full-wave analysis methods are the most accurate and can model coupling and arrays.

6.3 Permittivity considerations

[8]

Considering the fringing fields cause the antenna to radiate, the influence of different permittivities in the substrate selection should be considered. Small permittivities cause the fringing component to be even larger which improves radiation. This is of course a desired behaviour for antennas but not for the RF-frontend.

6.4 Combining Network Considerations

The antenna pattern can be optimized by varying the amplitude weighting of the patches.

This has a major influence on the radiation pattern of the antenna. Weighting influences mainly the width of the main-lobe and the amplitude of the side-lobes. It's always a trade-off between those two.

You can think of these weighting factors like windowing prior computing a DFT. A rectangular window (which can be thought of like equal weighting of the patches) has the tightest main-lobe but also the highest side-lobes. Hence, equal weighting of the patches should also produce the tightest main-lobe in the antenna-pattern but also high side-lobe levels.

Regarding the combining network of the antenna, unequal weighting significantly

increases the implementation effort. It would be necessary to implement an unequal Wilkinson combiner or a rat-race with different impedances for every pair of antennas you want to combine.

Because we are only interested in the bore-sight (or monopulse) axis the high side-lobes caused by equal weighting should not matter too much. The d/s ratio is only zero if d is zero or s goes off to infinity. The sum approaching infinity is not possible therefore the only zero should be on the boresight axis. To prove this, the ratio is shown in Figure 15 and Figure 16:

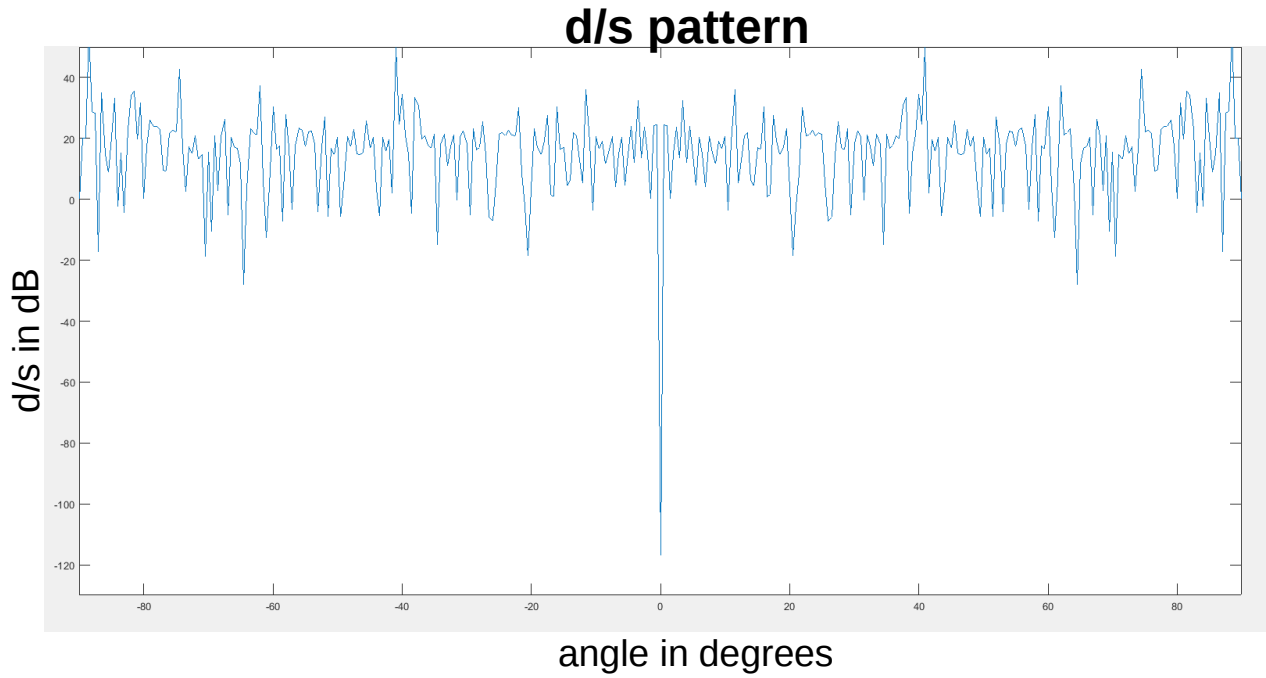

Figure 15: d/s pattern for angles from -90° to 90°

As you can clearly see in the plots there is only one significant minimum. Therefore, we should be fine even if we got an antenna with high side-lobes.

However a strong target near a local minimum of the d/s ratio with enough gain in the sum-pattern may eventually cause issues. This is highly dependent in the software implementation of the estimator.

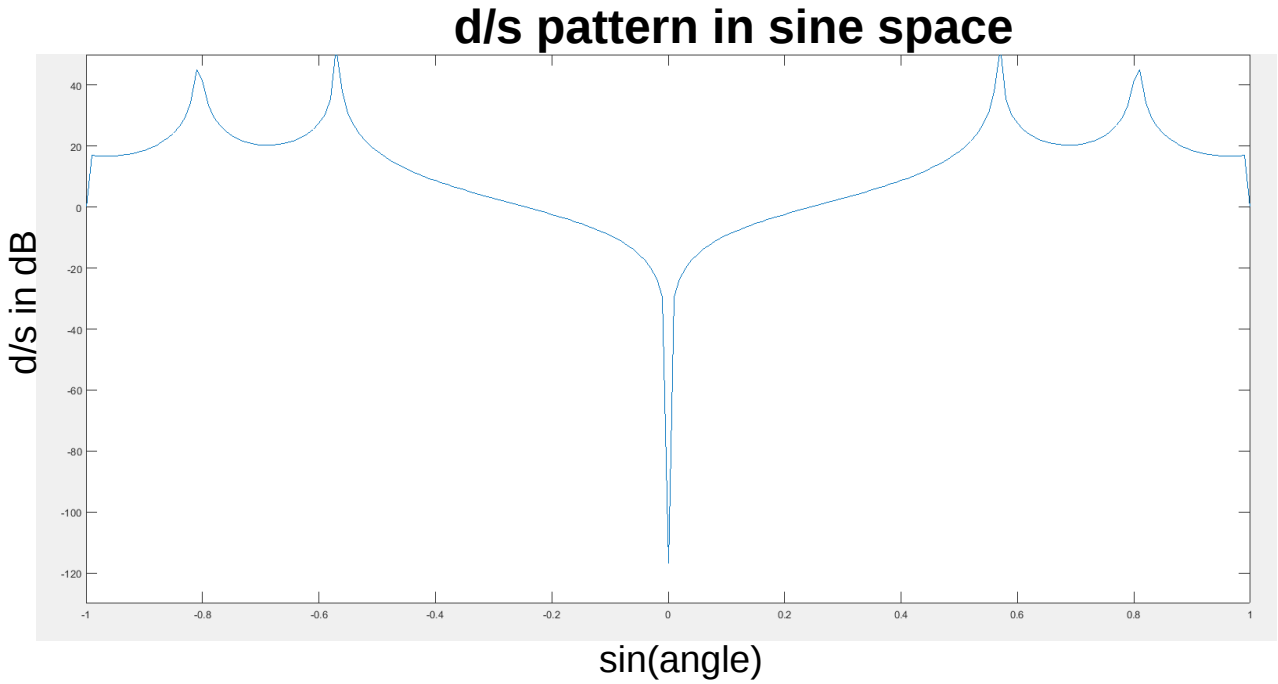

Figure 16: Illustration: d/s patter in sine space

From the desired main lobe width the overall array dimensions can be estimated. This was covered in chapter 5. With the maximum spacing without grating lobes the number of patches can be roughly estimated. The array is split into two subarrays to form the two antennas necessary for the phase monopulse principle. Hence the antenna design looks like this:

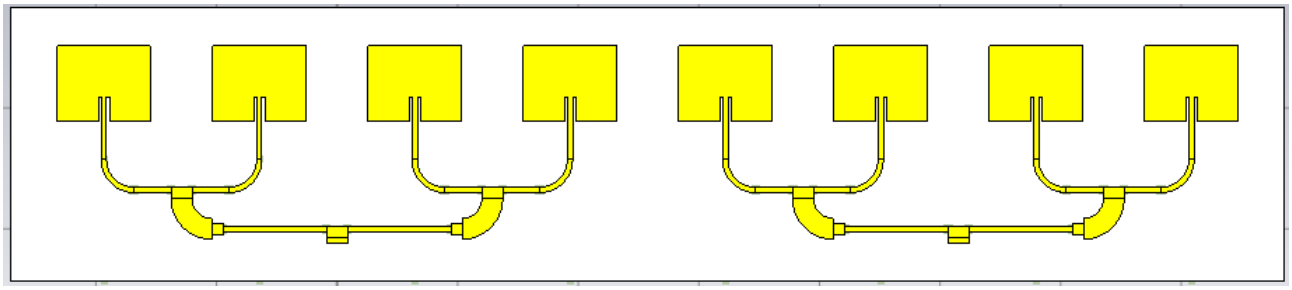

Figure 17: Antenna with feed network

To get an initial starting point for optimization, the transmission line model can be used to estimate the patch dimensions.

6.4.1 Transmission Line Model

6.4.1.1 Fringing Effect

[8]

To further discuss the transmission line model it is important to understand some fundamental principles like fringing.

The fringing effect is caused by the finite dimensions of the patch. This effect causes the field lines to bend at the edges of the patch. The effect is small but causes a shift in the resonant frequency. Therefore, this has to be taken into consideration. The microstrip line is a copper line between dielectric and air.

One of them is air and the other is your chosen substrate with a permittivity significantly larger than air. Hence, most of the field lines will reside in the substrate and just a few of them in air. Therefore the antenna appears to be larger than the physical dimensions suggest.

In order to take this fringing effect into account an effective relative permittivity $\epsilon_{r,eff}$ is introduced. The idea is to move the conductor (as it is without changing its dimensions) into the substrate. The effective relative permittivity is the constant of the uniform material (with the transmission line inside the material) that results in the same electrical behaviour as the real line on top of the substrate.

If the transmission line is between air and the substrate the effective relative permittivity is in the range of:

$$1 < \epsilon_{r,eff} < \epsilon_r \tag{54}$$

In case of ϵ_r being quite large, $\epsilon_{r,eff}$ approaches ϵ_r . Important to note is, that the effective relative permittivity is frequency dependent.
Higher frequencies cause the field lines to concentrate in the substrate even more. This causes $\epsilon_{r,eff}$ to approach ϵ_r .
In case of W/h > 1 the relative effective permittivity can be approximated by:

$$\epsilon_{r,eff} \approx \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2}\left[1 + 12\frac{h}{W}\right]^{-1/2} \tag{55}$$

Where W is the width of the patch and h is the thickness of the dielectric substrate.

This approximation doesn't take the frequency dependency into account.

6.5 Effective Length

As already noted, the length and width of the antenna appear to be larger than their physical dimensions suggest. The patch dimensions are show in figure 18.

A popular approximation for the fringing length is:

$$\Delta L \approx 0{,}412\, h \frac{(\epsilon_{r,eff} + 0{,}3)(\frac{W}{h} + 0{,}265)}{(\epsilon_{r,eff} - 0{,}258)(\frac{W}{h} + 0{,}8)} \tag{56}$$

This extension is caused on each side. Hence, the corrected patch length can be calculated by

$$L = L_{eff} - 2\Delta L \tag{57}$$

Where L_{eff} is the effective electrical length which is longer than the physical length L:

$$L_{eff} = \frac{c}{2 f_0 \sqrt{\epsilon_{r,eff}}} \tag{58}$$

W can be estimated by:

$$W \approx \frac{c}{2 f_0 \cdot \sqrt{\frac{\epsilon_r + 1}{2}}} \tag{59}$$

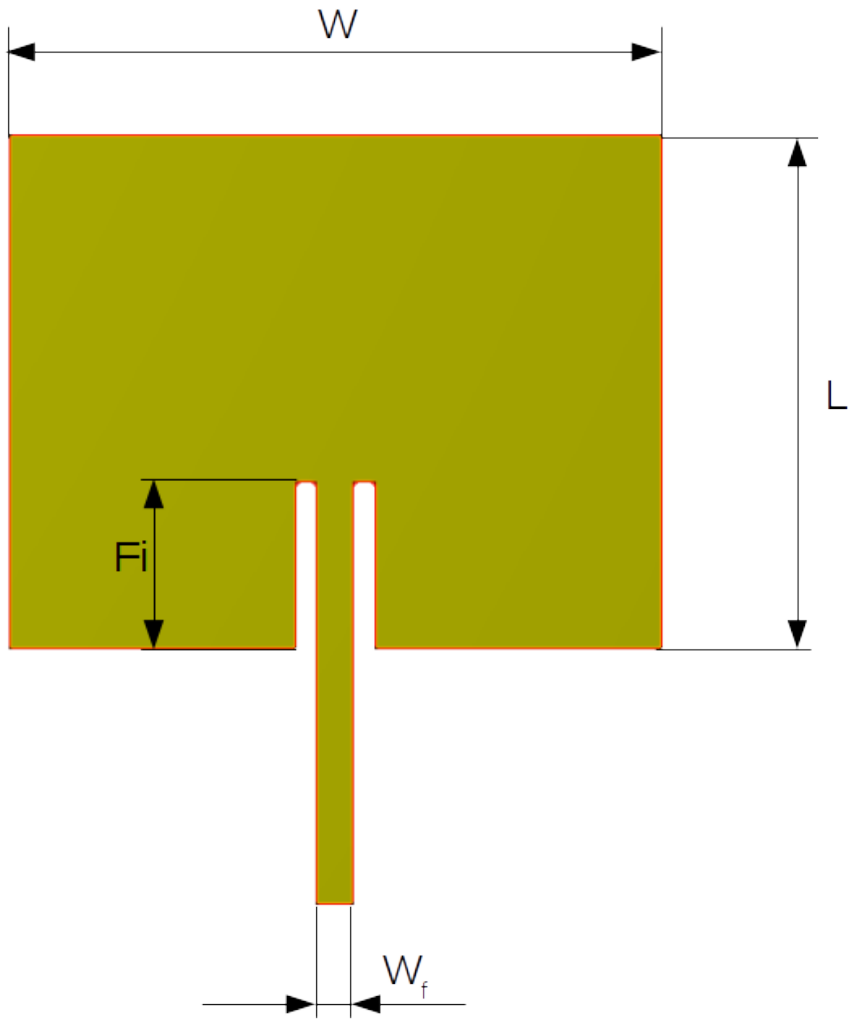

Figure 18: Patch dimensions

The input parameters for the transmission line model equations are:

center frequency	f_0	5,8 GHz
relative permittivity (Rogers 4350)	ϵ_r	3,48
substrate height	h	1,524 mm
conductor height	t	0,035 mm

Table 1: transmission line model input parameters

Using the transmission line model equations, results in the following approximated values:

patch width	W	17,27 mm
actual patch length	L	13,01 mm
length of inset	Fi	3,68 mm
width of feed	Wf	3,46 mm

Table 2: transmission line model results

Of course, the transmission line model is just a quick and simple approximation. In order to obtain accurate results and to consider coupling between the elements there is no way to get around full-wave analysis using finite elements method (FEM) or method of moments (MoM).

Using such software, results in:

patch width	W	15,47 mm
actual patch length	L	12,53 mm
length of inset	Fi	4,07 mm
width of feed	Wf	0,88 mm
spacing between patches	S	25,75 mm

Table 3: full-wave analysis results

The resonant length of the patch (L) suits the transmission line approximation quite nicely. The width of the patch differs quite significantly. Unlike the length the optimal width is not just some sort of resonant dimension. The ideal value highly depends on the optimization goals.

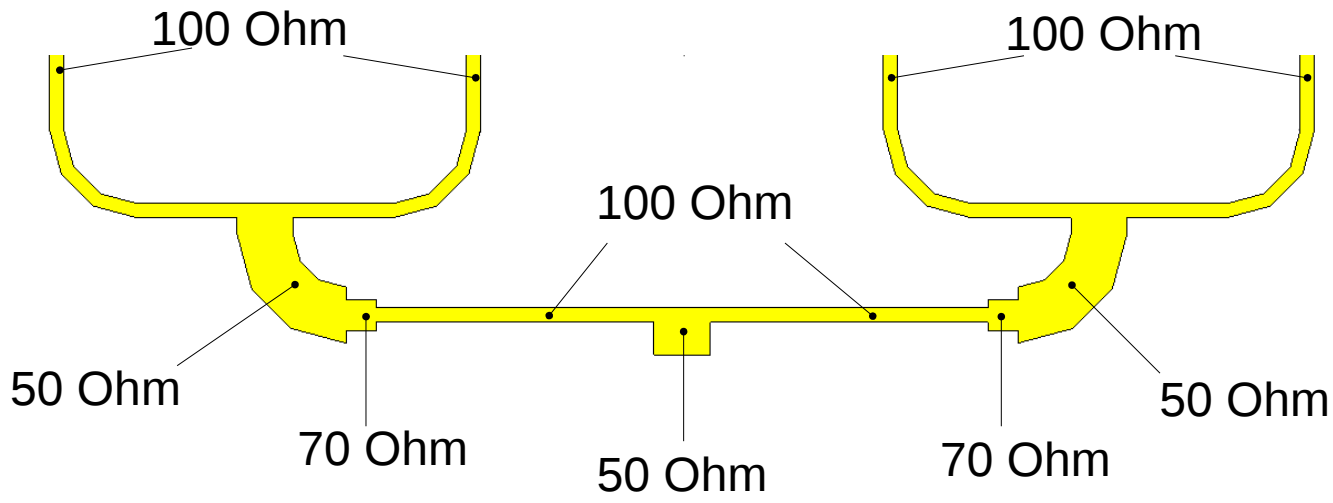

Figure 19: Feed network

Now let's take a closer look at the feed network in Figure 19. The ***100* Ω** lines feeding the patches are combined together to a ***50* Ω** line. This ***50* Ω** line is than matched to the ***100* Ω** line by a ***70* Ω *λ/4*** transformer. The two ***100* Ω** microstrip lines are then again combined to a single ***50* Ω** line.

6.6 Antenna Results

Using one of the less difficult to manufacture Rogers substrates (4350B), where the PCB can be manufactured using the standard FR-4 process, I was able to achieve considerably improved matching over the whole ISM band (5,725 GHz to 5,875 GHz) using different optimization algorithms and a relatively thick substrate which allows higher bandwidths at the expense of the antenna losses.

The results with the initial parameters were quite far away from desirable matching (e.g. > 10 dB return loss over the whole band). It is advantageous to start with a global optimization algorithm to prevent the algorithm from converging to a local minimum. Switching to a local algorithm if the results are reasonably close to the global minimum might be desirable.

Let's take a look at the reflection coefficients of the antenna in Figure 20. These parameters already include the combining network :

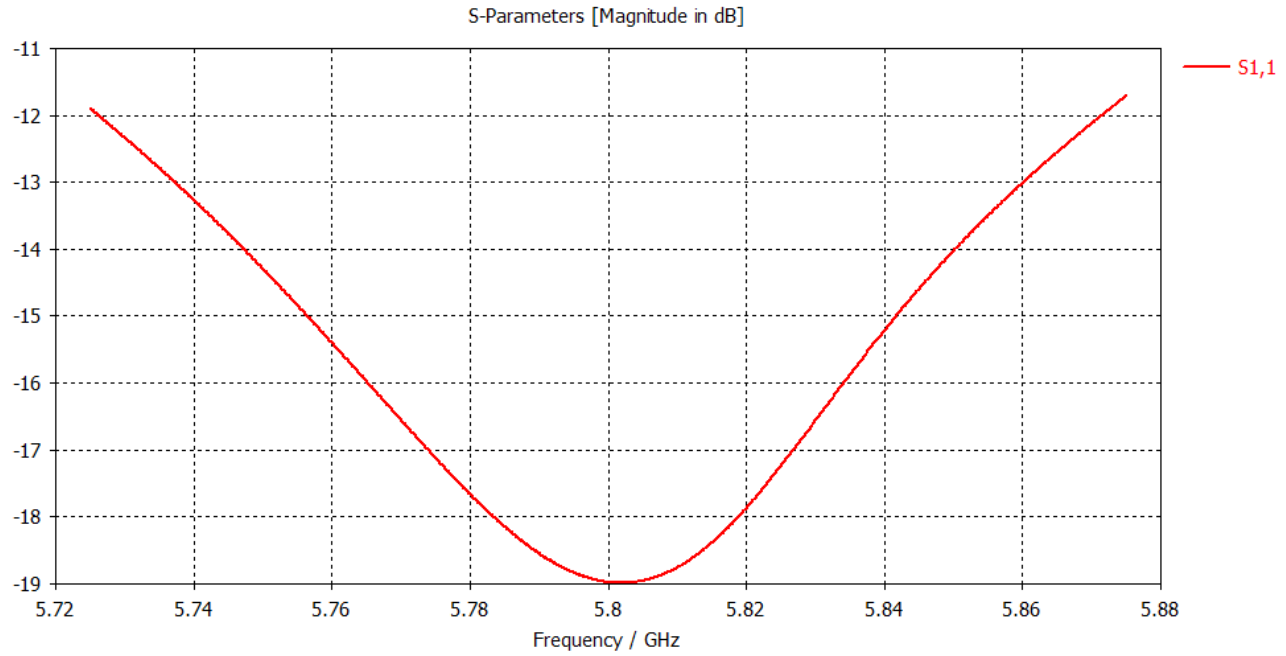

Figure 20: S11 of the combining network with antenna

As you can see from Figure 20, I was able to achieve a considerably good matching across the whole bandwidth. The return loss is always higher than 11 dB with a maximum of almost 19 dB return loss.

Let's take a look at the sum pattern of the patch array in Figure 21 and Figure 22. These 3D plots were obtained using the CST full wave analysis software.

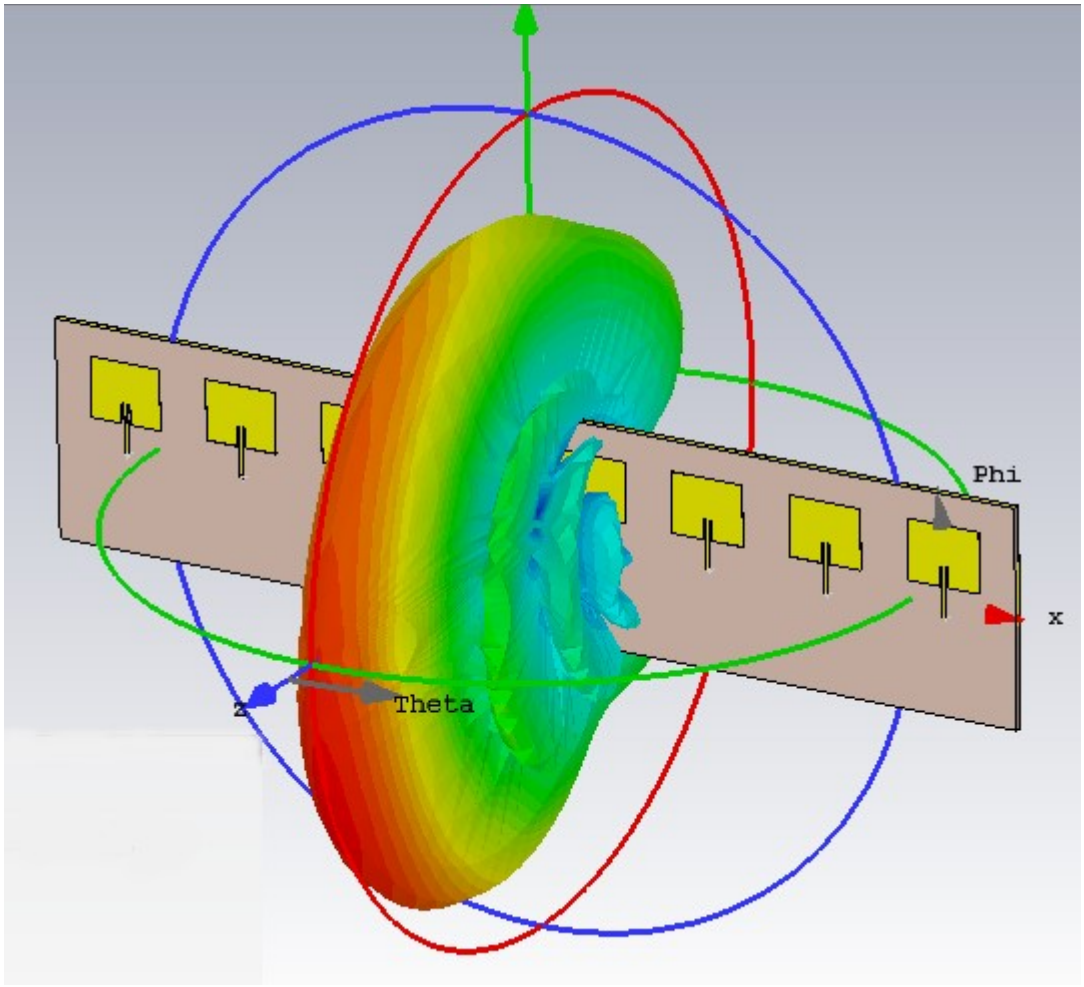

Figure 21: 3D-Farfield of sum pattern

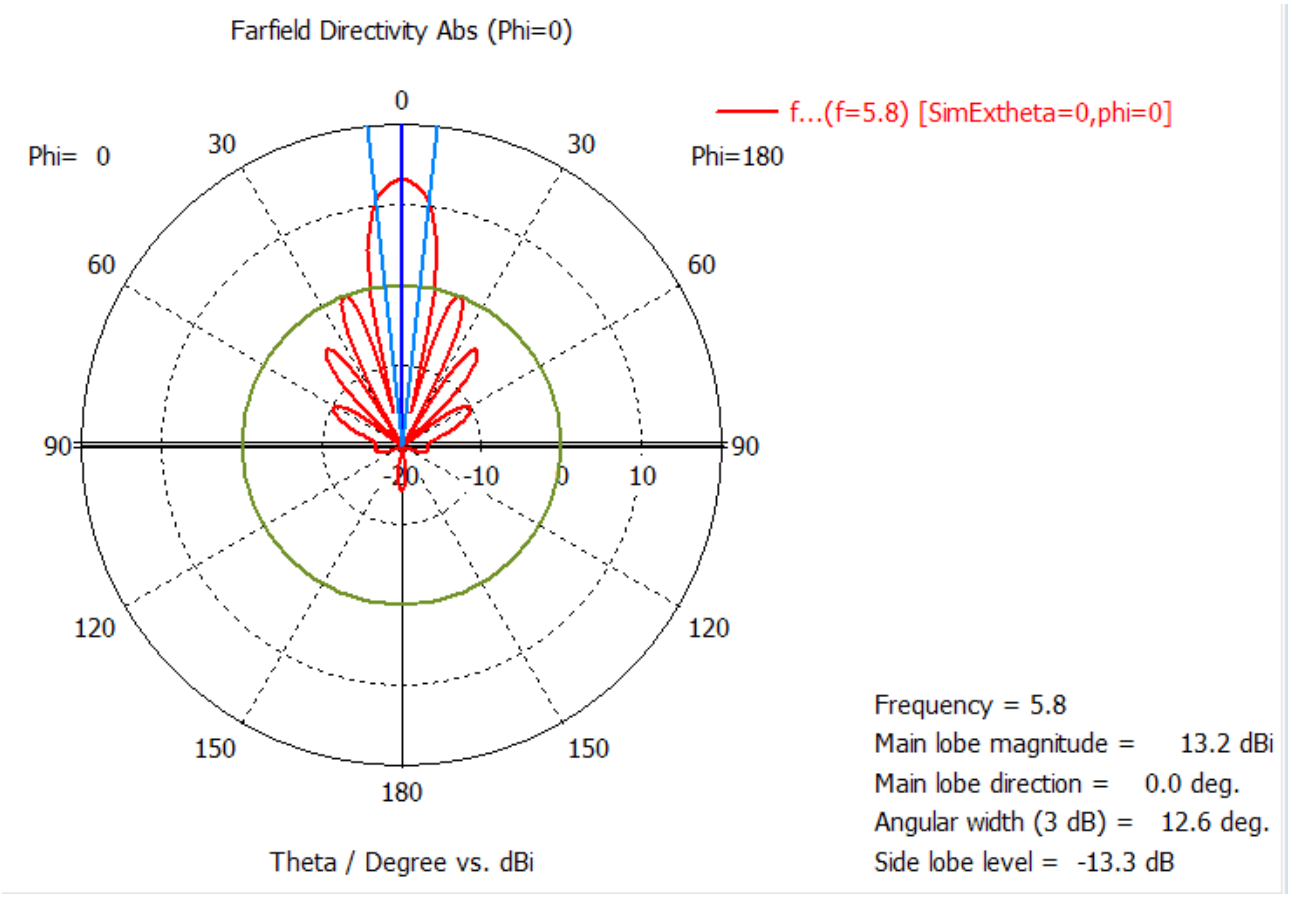

Figure 22: Horizontal sum pattern of Figure 21

Figure 22 shows a maximum gain of 13,2 dBi with a 3 dB-bandwidth of 12,6°.

The difference pattern in figure 23 and figure 24 shows the two beams quite nicely.

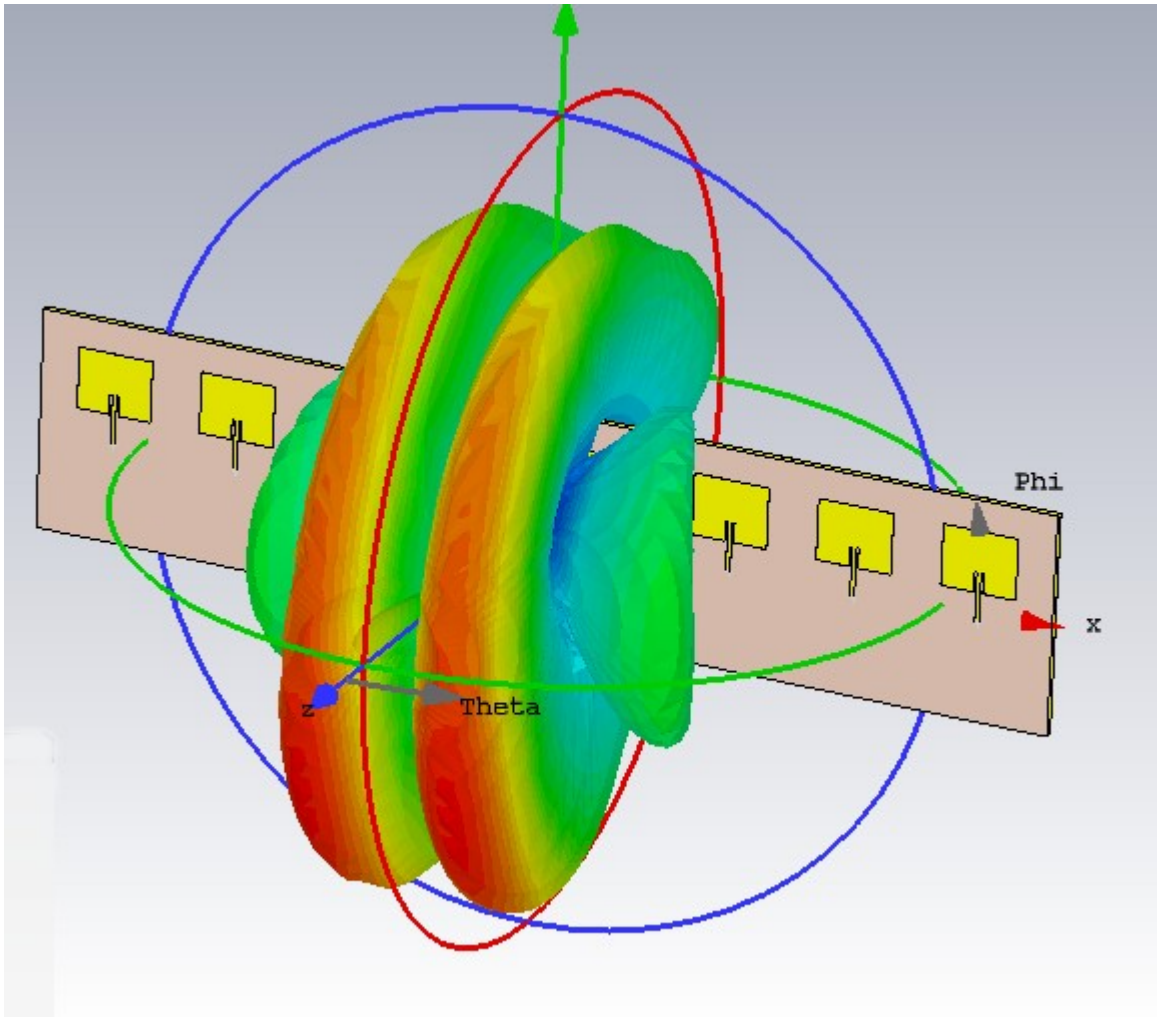

Figure 23: 3D-Farfield of difference pattern

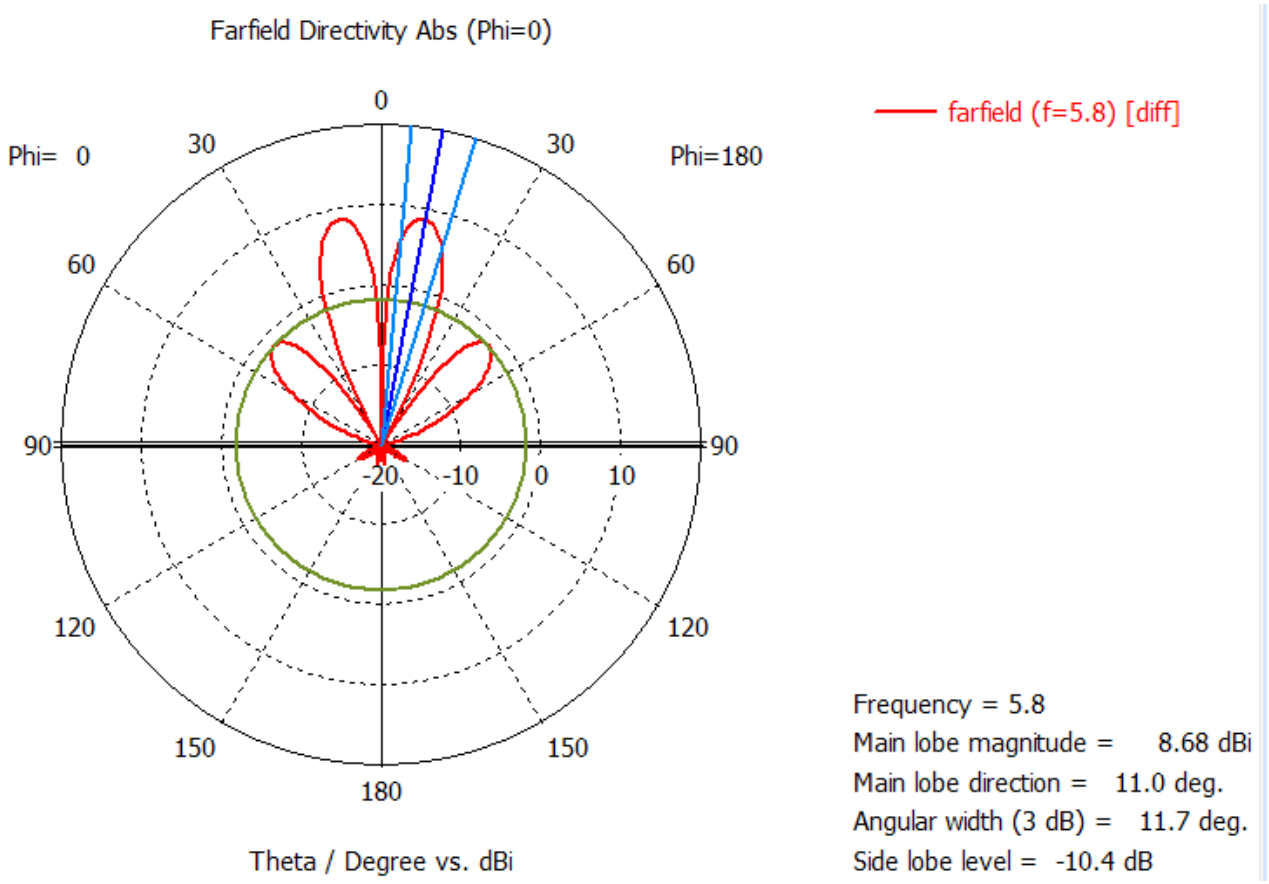

Figure 24: Polar plot of farfield difference pattern in the horizontal plane

7 Transmitter

There are various obvious reasons why an active transmitter in the model airplane compared to the radar approach is advantageous. I'll quickly list the most important points for my decision.

Without an amateur radio licence or a licence for a specific spectrum there are only the ISM bands with their corresponding limitations.

For this system, there are a few possible frequencies to take into consideration. The most relevant ones are 433 MHz, 800-900 MHz (SDR-band ...), 2,4 GHz, 5,8 GHz and 24 GHz and a few at even higher frequencies which are not feasible for us in terms of the price point.

In terms of usage/interference the 433 MHz and 2,4 GHz spectrum are horribly crowded. All cheap devices (the model airplane senders themselves, WLAN routers, garage door openers, temperature stations, microwave ovens …) use one of these two frequency bands. The list of devices in these two bands is endless. The 800-900 MHz band gets partially cut away by LTE and GSM and has further restrictions in terms of transmit duration.

24 GHz is probably a viable option but this band was not considered because of the significantly higher cost in terms of components, PCB material and manufacturing etc.

So, there's only one band left, which is the 5,8 GHz ISM band ranging from 5,725 GHz to 5,875 GHz with a maximum EIRP of 25 mW (14 dBm).

Due to the low maximum power of 14 dBm and the small radar cross-section of model airplanes a radar-based system is simply not feasible. The low cross-section and the significantly higher path loss (two way path loss) effect the SNR way to much especially at this low transmit power.

Due to these reasons an active radio system was chosen. The easiest way would be the transmission of a continuous sine wave, but this is not possible for legal reasons.

To meet the legal requirements, we have to send information (not just a single sine-carrier) or use the required duty cycle.

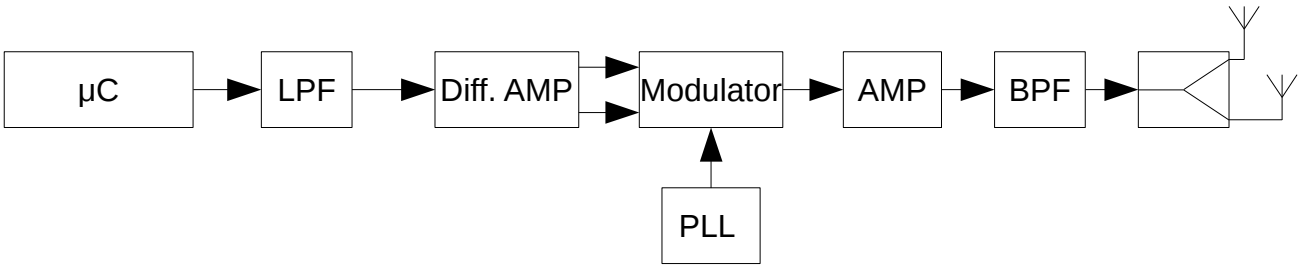

Figure 25: Transmitter Schematic

Let's take a look at the basic transmitter architecture illustrated in Figure 25.

In order to obtain the baseband waveform, a microcontroller (µC) generates the BPSK samples. The modulation scheme doesn't matter in this case. Hence simple BPSK is used.

The µC is followed by a lowpass filter to band-limit the rectangular baseband waveform. Followed by a differential amplifier to generate differential outputs needed for feeding the in-phase-part of the I/Q modulator. The quadrature-part is not necessary for simple BPSK modulation.

The output of the I/Q Modulator is fed into an amplifier. Any intermodulation products that may arise from the non-linearities of the amplifier get attenuated by the bandpass filter. A power divider is used to split up the signal into two orthogonal dipole antennas to cope with the varying orientation of the plane and prohibit polarisation mismatch. Additionally the orthogonal antennas also help in achieving a omni-directional pattern.

8 Receiver

The direct-conversion receiver concept, was only chosen because of the low budget for this project. There are significant drawbacks using a direct-conversion receiver which is the reason why this type of receiver gets hardly used nowadays.

Let me quickly sum up the main drawbacks of the direct conversion approach.

The first main disadvantage is related to the I/Q imbalance. The direct conversion receiver does the I/Q demodulation in the analog domain. The differences in phase and frequency of the local oscillator in the in-phase and quadrature phase part of the receiver yield to a constant (in case of phase deviations) or time-varying (in case of frequency deviations) phase difference between the two LO signals at the in-phase and quadrature phase mixer. Hence, the I and Q axis are not orthogonal anymore which yields to an error in the complex baseband signal.

This can simply be solved by using a superheterodyne architecture to sample at the intermediate frequency and shifting the I/Q demodulation process into the digital domain.

The second huge problem is the DC-offset caused by a self-mixing effect at the mixers of the I/Q Demodulator. The strong LO signal at the mixer leaks to the RF and IF port. Reflections cause the leaked LO signal to be reapplied at the RF port of the mixer which results in a DC-offset.

Building the sum and difference signal is done at an initial stage in order to achieve a more stable boresight/monopulse axis compared to doing it in the digital domain (radio interferometry) to avoid (time-varying) gain- and phase imbalances in the two receiver chains.

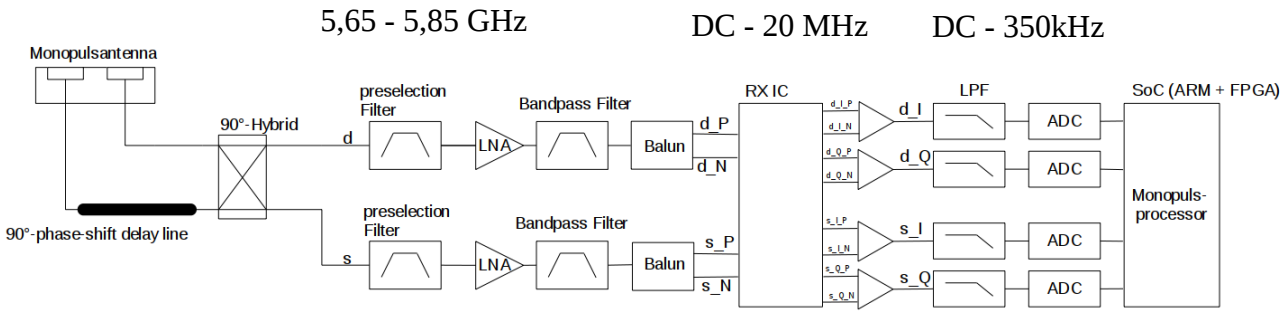

Figure 26: Receiver Design

The preselection filter removes the out of band components followed by an low noise amplifier (LNA) to achieve a better noise figure compared to amplifying the signal later in the RX-IC. The second filter is only used because the first filter does not achieve a reasonably well stop band attenuation. A balun provides the differential inputs needed for the direct conversion receiver IC. The lowpass filter attenuates noise and interference outside of the desired baseband and also acts as an anti aliasing filter (AAF) to prevent or attenuate aliasing components before the sampling process.

8.1 ADC Considerations

[16]

To select an appropriate analog-to-digital converter (ADC) it is necessary to get an estimate on the noise caused by the analog frontend in order to make an suitable decision on how much your ADC may contribute to the overall noise figure.

If the ADC should contribute at most ***0.8 dB*** (about 20 % or factor 1.2) at most to the overall noise, the ADC noise density has to be ***7 dB*** (factor 5) below the noise spectral density of the analog frontend (AFE). [16]

In order to compute the noise spectral density of the analog frontend the noise at the antenna output is assumed to be ***174 dBm/Hz*** which is the spectral noise density at an ambient temperature of 20° C. Hence, you can easily calculate the spectral noise density of the analog frontend from the noise at the antenna output and the overall gain and noise figure of the frontend:

$$NSD_{AFE} = -174\,dBm/Hz + G_{AFE} + NF_{AFE} \tag{60}$$

To obtain the overall full-scale range (FSR) of the ADC in ***dBm*** the following calculations can be done:

From the datasheet, the peak-to-peak FSR is:

$$FSR_{ADC} = P = \frac{\left(\frac{V_{peak}}{\sqrt{2}}\right)^2}{R} \tag{61}$$

Let's calculate the minimum necessary SNR from the noise spectral density (NSD) from the frontend and the ***0.8 dB*** which we want to contribute at most:

$$SNR_{ADC} = FSR_{ADC} - [\,NSD_{AFE} - guard + 10 \cdot \log\left(\frac{sampling\_rate}{2}\right)] \tag{62}$$

I did not specify any values here because everything depends on the settings in the RX-IC.

8.2 NF of the ADC

[16]

Although a ADC is a “voltage device” instead of a “power device” we can estimate a noise figure (NF) in case the ADC impedance is matched to the impedance of the rest of the circuit.

We can estimate the noise figure by:

$$NF_{ADC} = FSR_{ADC} - SNR_{ADC} - 10\log\left(\frac{sample\_rate}{2}\right) - \left(-174\,dBm/Hz\right) \tag{63}$$

9 RF Printed Circuit Board (PCB) Design

To cover all the important aspects of RX and TX design, I'd like to give you a brief overview about RF PCB design. This chapter is by no means extensive but covers the most relevant basics regarding high frequency PCBs.

9.1 PCB Traces and characteristic impedance

[10]

To obtain a defined characteristic impedance of all the lines in the PCB design it is very important to know the PCB stackup and relative permittivity of the dielectric layer between the RF-plane and the ground plane. So, in RF design it is necessary to order a custom stackup to get a guaranteed substrate thickness.

For really high frequencies a high-performance substrate like one of the Rogers substrates is necessary to achieve a low tolerance on the relative permittivity and therefore obtain a well-defined characteristic impedance and in addition to that a very low loss tangent to obtain considerably lower losses at high frequencies compared to FR-4, which is the standard PCB material at low frequencies.

Up to the lower GHz range (e.g. 5 GHz) using FR-4 cores as a dielectric between the RF-signal layer and the ground plane might still be fine. Prepregs can have significantly higher variation in permittivity (e.g. different thicknesses have different permittivities) compared to cores. Prepregs are basically pre-impregnated fibreglass with resin. On the other hand in cores the resin is already hardened. Hence the thickness and dielectric properties are comparably well-defined.

It is advisable to look at the datasheets of the FR-4 substrates used by the PCB manufacturer because the permittivity can vary greatly between different FR-4 substrates. Additionally the aging of the FR-4 can cause the substrate to drain and therefore to change its dielectric behaviour.

The current in the ground plane has its maximum value directly under the PCB-trace. This comes from the fact that the impedance is lowest directly below the trace.

In order to maintain a homogeneous nature of the current density it is advisable to use a solid return path and therefore solid ground-plane.

It is often adviceable to use short RF traces and avoid loops which cause a high inductance by its winding trace and the discontinuous ground plane.

Because the ground-plane is also used as a voltage reference, noise caused by high current densities can be a problem.

9.2 Copper Planes

[10]
Using a VCC plane results in a large RF bypass capacitor. Hence, this approach is often used as a complement to normal ceramic bypass capacitors. A small distance between the VCC and ground plane causes the capacity to increase.
In low frequency PCB design hatched ground-planes are often used. In RF, solid ground planes are used to minimize the inductance which is a desirable choice for high speed signals. This comes at the price of an increased stray capacitance for the signal traces caused by the larger area.
A positive side effect is the better cooling capabilities achieved by the larger area of a solid ground-plane. But this may also cause the heat to spread to temperature sensitive components.

9.3 Transmission lines

In this chapter i will quickly go over the commonly used types of transmission lines and their advantages and disadvantages.

9.3.1 Stripline

[11] [18] [19]

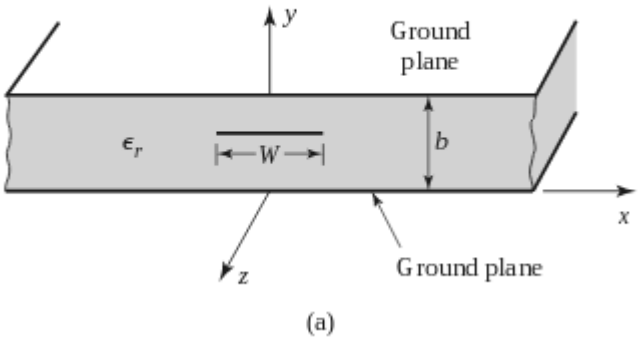

Figure 27: Stripline geometry (source:)

A stripline is basically a copper strip between two ground planes. Striplines are commonly used in the inner layers of a PCB design. Because of the homogeneous dielectric on each side it supports transversal electromagnetic mode (TEM) waves. TEM means that electric and magnetic field components are orthogonal to the propagation direction (in case of a lossless medium).

The advantages are:

- no dispersion
- Immunity to crosstalk
- electro magnetic interference (EMI) characteristic

The disadvantage is:

- connection to devices needs always vias

9.3.2 Microstrip

[11] [18] [19]

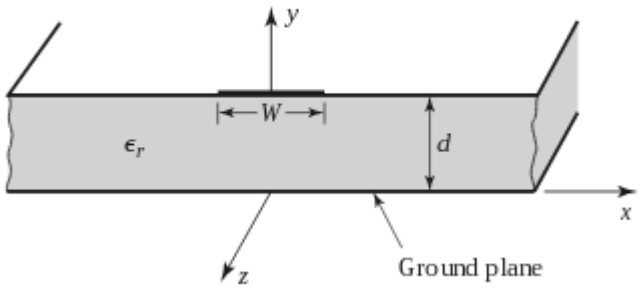

Figure 28: Microstrip geometry

In comparison to striplines the conductor is now between a dielectric and air. Hence the phase velocities in air and the dielectric are different and therefore the waves are no longer TEM but hybrid TM-TE. However if the substrate thickness is small (compared to the wavelength) the field is Quasi-TEM. This means that there is also a small component in the direction of propagation present. However in Quasi-TEM waves these components are very small.

In order to calculate the dimensions of a microstrip the actual permittivity of the substrate ε_r has to be replaced by a dispersive ε_{reff}. The permittivity is varying significantly with frequency. Hence microstrips should be avoided in broadband applications. At high frequencies field components stay mostly in the substrate. Hence ε_{reff} comes close to ε_r.

At low frequencies a significant part of the field propagates through air. In this case ε_{reff} is significantly smaller.

Upmost layers in PCBs are often microstrips.

The advantages are:

- lower loss tangent (compared to stripline)
- high packing density (compared to stripline)
- easy to combine with SMD parts (compared to stripline)
- low sensitivity to manufacturing tolerances
- easy to realize distributed components
- faster propagation time (compared to microstrip)

9.3.3 Coplanar Waveguide with Ground (CPWG)

[11] [18] [19] [20]

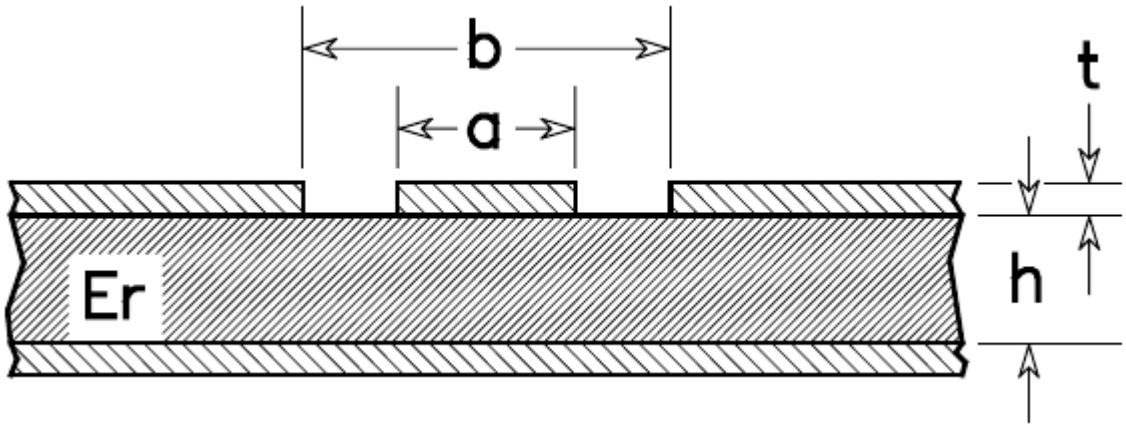

Figure 29: Coplanar Waveguide with Ground (source: [11])

Coplanar waveguides (CPW) have all conductors on the same side of the substrate and support a Quasi-TEM mode in terms of propagation.

Coplanar waveguides are often used at very high frequencies or if tight spacing between lines is required.

Advantages:

- lower loss tangent than microstrip (less radiation)
- can narrow 50 Ohm trace widths
- easy integration of shunt components
- overall flexibility to control the characteristic impedance
  - many degrees of freedom
- less crosstalk between adjacent lines
  - therefore tighter spacing between different lines is possible
  - for that reason often used in mircowave integrated circuits (MIC) and monolithic microwave integrated circuits (MMIC)

Disadvantages

- higher skin effect losses

9.3.4 Stray capacitance

[10]

Stray capacitance can be either a good or a bad thing. On one hand, the stray capacitance helps the CPWG to “build” the characteristic impedance. If a specific characteristic impedance is not necessary, it can slow down circuits.

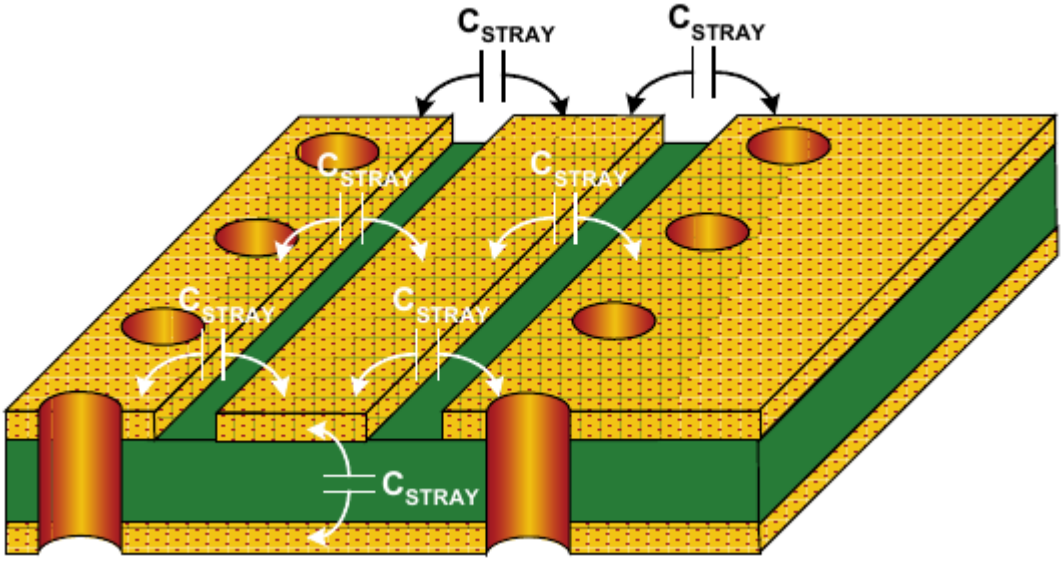

Figure 30: Stray capacitance of grounded coplanar waveguide (source: [10])

Figure 30 clearly shows the stray capacitance to the different ground planes. If the trace is a line which needs to meet the characteristic impedance, a high stray capacity is often acceptable. Otherwise consider increasing the clearance between the trace and ground plane and removing the ground plane under the trace.

I went with coplanar waveguides with ground mainly because of the size requirements in the transmitter. The tight spacing of the lines would results in too much crosstalk between adjacent lines in case of a microstrip.

10 Issues regarding monopulse systems - Monopulse Angle Errors

We already digged through the theory about the monopulse principle, the monopulse ratio and several design decisions regarding the design of such systems. Now it is time to talk about several issues regarding the monopulse technique and the basic approaches how to deal with them.

A monopulse system is subject to various errors which cause errors in angle estimations. This chapter will deal with errors due to multiple targets or interferers, multipath and other errors (e.g. due to noise).

10.1 Multiple Targets

[3]

One of the most obvious issues is the response of the system to other targets or interferers.

For radar systems we would have to deal with unresolved targets. Targets are unresolved if all angle estimations can't be done without significant errors.

In case of a radar system the targets are often called unresolved if they are in the same resolution cell.

Hence were not dealing with a radar-based approach we are only concerned with in-band interferers.

In case of an interferer or unresolved target the indicated angle might not be the actual angle of the desired target. Other targets or interferes cause a change in phase in the difference and sum signal which causes the estimated angle to deviate from the real one. Hence, the estimated angle does not correspond perfectly with the real angle even at perfect SNR conditions. It does not even correspond to the angle of any target or interferer or scattering point anymore.

The other signals present cause a relative phase shift between difference and sum signal which causes the monopulse ratio d/s to become a complex quantity. In general, for a single target the monopulse ratio is purely real or imaginary depending on the principle (amplitude vs. phase monopulse) and/or combining network to form the sum and difference patterns.

In case of multiple targets, it is probably more useful to treat all amplitudes and phases of the targets as random variables.

Anyway, this chapter is about one of the fundamental problems of all electromagnetic measurement systems. The issue is basically that any distortion of your electromagnetic wave front will additionally decrease your measurement accuracy.

The real and imaginary parts of the monopulse ratio correspond to the gradients of phase and amplitude of the arriving plane wave. One way you could use the information of the real and imaginary part would be in detection of other targets, interferers or scattering components.

In order to simplify the analysis on these topics we have to introduce some sort of simplifications.

The assumptions are:

- The sum and difference signals from a single target are 0° (or 180°) in relative phase.
  - the monopulse ratio is therefore real
  - depending on amplitude/phase monopulse and combining network the phase shift might be 90° or 270°
    - this can be easily solved by changing real parts to imaginary parts
- The monopulse processor is an ideal processor and considers only the real part of the d/s ratio
- The d/s ratio is proportional to the angle in case of a single target being present.
  - this is usually only true for small deviations from the boresight angle
- if two angular coordinates are measured they are assumed to be independent
- Noise and all other errors are neglected

10.1.1 Point targets in monopulse

[3]

First let me describe the definition of a point target for our context. In monopulse measurements a point target is concentrated into a single point in angle space or sine space. However, this doesn't have to be true for range.

10.1.2 Two-Target problem

We have already briefly talked about the d/s ratio being purely real (or imaginary) in the single target case. Multiple targets cause a relative phase-shift between the sum (s) and difference (d) signal which causes the ratio to become a complex quantity.

To confirm this statement, consider two unresolved targets at different angles. To illustrate this further let's look at Figure 31:

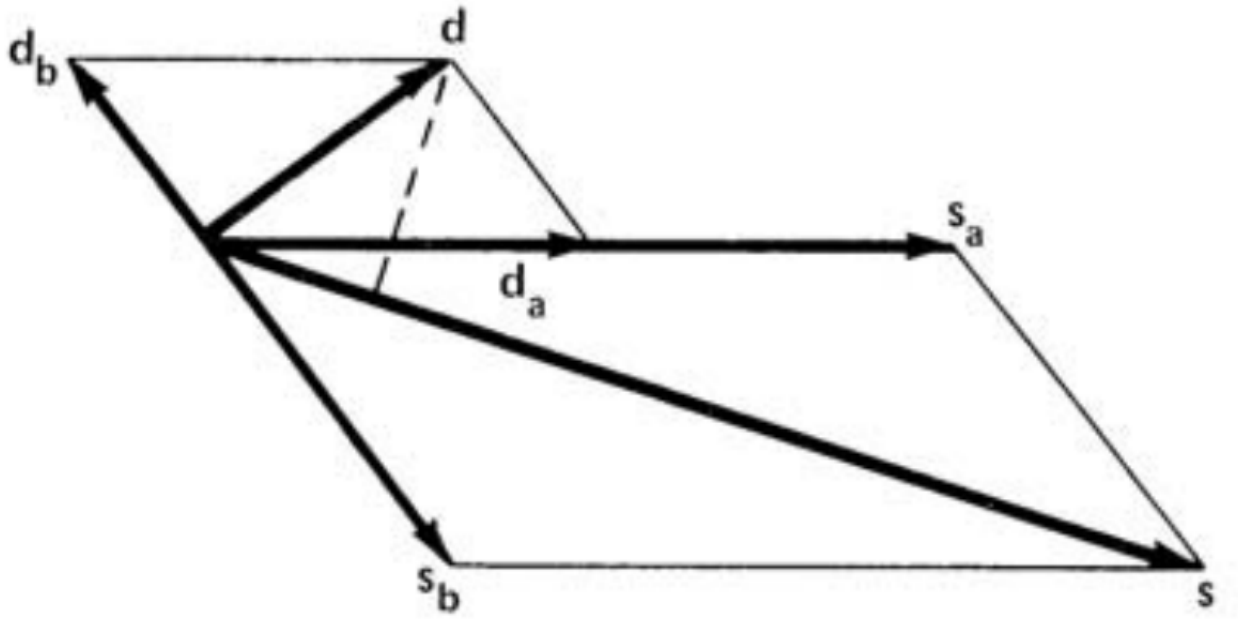

Figure 31: Sum and difference phasors of two unresolved targets (source: [3])

To clarify this: s_a and d_a are the sum and difference signals from target a and s_b and d_b are the sum and difference signal of target b. Where d and s are the combined difference and sum signals. For this figure the targets are assumed to be on the opposite site of the beam axis which causes sum and difference signal to be in phase for one target and to be opposite phase for the other target.

The normalization of d by s causes the resulting phase to be in relation to the phase of

the sum signal.

Figure 31 shows us that this causes the d/s ratio to be complex because the difference signal has a component in phase with the sum signal and also a component which is orthogonal to the sum signal.

It is quite obvious that if s_a and s_b are 180° out of phase this causes the sum signal to become very small which results in a small denominator of the monopulse ratio and therefore the ratio to become really large.

To confirm the statements made let's express this in a mathematical fashion.

$$\frac{d_a}{s_a} = k_m \theta_a \tag{64}$$

$$\frac{d_b}{s_b} = k_m \theta_b \tag{65}$$

Where s_a and d_a are the sum and difference signals of target a, s_b and d_b the sum and difference signal of target b, k_m the monopulse slope and θ_a and θ_b are the the angular displacements of target a and b from the boresight axis.
The indicated angle can be calculated from the monopulse slope and the monopulse ratio:

$$\theta_i = \frac{1}{k_m}\frac{d}{s} = \frac{1}{k_m}\frac{d_a + d_b}{s_a + s_b} = \frac{\theta_a s_a + \theta_b s_b}{s_a + s_b} \tag{66}$$

The indicated angle is a weighted average over the actual angles of the two targets with the weighting proportional to their corresponding sum signals and therefore the corresponding power in the sum pattern.
It is important to note that the weighting is complex which might not be obvious at first. This comes from the fact that s_a and s_b may have different in phase. Therefore, the result cannot be obtained that easily.
Let's assume:

$$\frac{s_b}{s_a} = p\, e^{j\Phi} \tag{67}$$

Where p is the amplitude ratio of the sum signal returns of the two targets and Φ being the relative phase of the sum signals of the two targets.
The definition of amplitude and phase ratio being the ratios from target b in relation to target a are arbitrary.

By dividing the right-hand side of equation (66) by s_a and by using the relation in equation (67) results in:

$$\theta_i = \frac{\theta_a + p e^{j\Phi} \theta_b}{1 + p e^{j\Phi}} \tag{68}$$

The result can also be expressed by the angular midpoint θ_{mid} and angular separation $\Delta\theta$ of the two targets:

$$\theta_i = \theta_{mid} - \frac{\Delta\theta}{2} \frac{1 - p e^{j\Phi}}{1 + p e^{j\Phi}} \tag{69}$$

Where the angular midpoint is:

$$\theta_{mid} = \frac{\theta_a + \theta_b}{2} \tag{70}$$

And the angular separation:

$$\Delta\theta = \theta_a - \theta_b \tag{71}$$

Another way to express the equation would be in relation to the angle of the first target θ_a and an error term which might be more convenient in error approximations:

$$\theta_i = \theta_a + \Delta\theta \frac{p e^{j\Phi}}{1 + p e^{j\Phi}} \tag{72}$$

10.1.3 The complex indicated angle

[3]

We already quite exhaustively discussed the indicated angle θ_i . In general, this angle is called complex indicated angle or simply complex angle.

To further discuss the complex angle, some special cases can be assumed. At first let's assume target b is not present which means p=0 and therefore the indicated angle is the angle of target a. On the other hand, assuming target a is not present means p approaching infinity which results in the complex indicated angle approaching the angle of target b. It should be obvious that if both targets are at the same angle the result of the indicated angle is also correct.

Typical monopulse processors in general deal mostly with the real (or imaginary) part of the monopulse ratio. Hence, the indicated angle which is calculated from the monopulse ratio and the slope of the monopulse curve is also purely real or imaginary.

In order to further analyse the behaviour of these processors on two target problems we have to extend equations (68), (69) and (72) by the conjugate complex of the denominator. Hence, it gives:

$$\Re(\theta_i)=\frac{\theta_a+p\cdot\cos(\Phi)\cdot(\theta_a+\theta_b)+p^2\theta_b}{1+2\cdot p\cdot\cos(\Phi)+p^2} \tag{73}$$

$$\Re(\theta_i)=\theta_{mid}-\frac{\Delta\theta}{2}\frac{1-p^2}{1+2\cdot p\cdot\cos(\Phi)+p^2} \tag{74}$$

$$\Re(\theta_i)=\theta_a+\Delta\theta\frac{p\cdot\cos(\Phi)\cdot+p^2}{1+2\cdot p\cdot\cos(\Phi)+p^2} \tag{75}$$

In general, the imaginary part is ignored. But in principle it is available at least if your monopulse processor deals with I/Q samples. It would look like this:

$$\Im(\theta_i)=\Delta\theta\frac{p\cdot\sin(\Phi)}{1+2\cdot p\cdot\cos(\Phi)+p^2} \tag{76}$$

From the already obtained real and imaginary parts the relative phase can be obtained:

$$\tan\delta = \frac{\Im(\theta_i)}{\Re(\theta_i)} = \frac{p\cdot\sin(\Phi)}{\theta_a + p\cdot\cos(\Phi)\cdot(\theta_a+\theta_b)+p^2\theta_b} \tag{77}$$

For a point target in the farfield it is sufficient to assume that the gradient of the phase, which is basically the normal on the planar phase front aligns well with the target angle. Nevertheless, unresolved targets will cause a significant distortion of the phase front which causes amplitude variations and the phase-front becomes non-planar.

10.1.4 Detection of unresolved targets

[3]

If the architecture of our monopulse processor provides us with I/Q samples it might be favourable to make use of the quadrature samples.

As denoted several times a single point target should not cause a complex monopulse ratio. Hence, the most obvious way for detecting unresolved targets would be thresholding the imaginary part. However, this threshold should be chosen high enough to avoid false alarms caused by noise, clutter or multipath.

It might be necessary to average the imaginary part over several samples in order to make the detection more accurate.

10.1.5 Statistical properties of the complex indicated angle

[3]

To start the analysis, the equation of the complex indicated angle can be used as a starting point. To obtain the mean of the indicated angle, path integration can be applied:

$$\bar{\theta}_i = \theta_a + \frac{\Delta\theta}{2\pi}\int_0^{2\pi}\frac{p e^{j\Phi}}{1+p e^{j\Phi}}d\Phi \tag{78}$$

Φ is assumed to be uniformly distributed.
Substituting and taking the derivative results in:

$$w = p\,e^{j\Phi} \tag{79}$$

$$dw = jpe^{j\Phi}\,d\Phi \tag{80}$$

Using this substitution results in:

$$\bar{\theta}_i = \theta_a + \frac{\Delta\theta}{2\pi}\oint_C \frac{1}{j}\frac{dw}{1+w} \tag{81}$$

Where C represents the integration path. This complex contour is not restricted to some circle but can be any path depending on p.
If one target remains stronger over the whole integration period the indicated angle corresponds to the angle of the stronger target. In case that one target is not always stronger the indicated angle is the weighted average where the weighting is the timeframe one target is stronger.
In case of the amplitudes following a Rayleigh distribution and the phases are uniformly distributed the average angle would be the angles weighted by the power levels of the targets. [3]
The beam axis has to remain fixed during the averaging. Hence, this is only true for open-loop measurements or really slow tracking systems.

The variance has been found to be the same for real and imaginary parts in constant amplitude ratio scenarios:

$$Var[\Re(\theta_i)] = Var[\Im(\theta_i)] = \frac{p^2(\Delta\theta)^2}{2(1-p^2)} \tag{82}$$

In this equation p is assumed to be the amplitude ratio of the weaker target to the stronger one.
For fluctuating targets, the variance has to be averaged over the power ratio with respect to the probability density function (PDF). If none of the targets is always stronger the variance is computed for each of the targets again weighted by the corresponding timeframe where the individual targets are stronger.
In practice, results obtained by using averaging won't be the same result as above. We're working with quantities which are discrete in time and amplitude. Hence, the integral changes to a sum and the analytic results differ.
The averaging interval might be significantly shorter in practice.

10.1.6 Weighted mean of complex indicated angle

The results presented in the preceding section were obtained with equal weighting of the measurements. To analyse the topic even further the measurements can be weighted by the power of the sum signal. This should weight the “good” estimations (obtained with a higher power) with higher weighting factors than the “bad” ones.

Integrating over the numerator and denominator of equation (73) results in:

$$\Re\left(\bar{\theta}_i\right)_{\mathrm{wp}} = \frac{\int_0^{2\pi} \left[\theta_a + p\cdot\cos(\Phi)\cdot(\theta_a+\theta_b) + p^2\theta_b\right] d\Phi}{\int_0^{2\pi} \left[1+2\cdot p\cdot\cos(\Phi)+p^2\right] d\Phi} \tag{83}$$

The subscript wp denotes the weighting for power. The denominator corresponds to the sum-signal power normalized by the power of the first target. [3]
It is assumed that the target angles do not change over the averaging interval and that p remains constant.
Integrating over a period of a cosine function is zero. Hence, this can be simplified to:

$$\Re\left(\bar{\theta}_i\right)_{\mathrm{wp}} = \frac{\theta_a + p^2\theta_b}{1+p^2} \tag{84}$$

Calculating the mean of the imaginary part in equation (76) is straightforward. Because the numerator only consists of a sine term, integrating over an integer multiple of the period is always zero. Hence, the mean of the imaginary is zero.

10.1.7 Estimating the angles of unresolved targets

[3]

We have already derived the equations for the real (equation (73), (74) or (75)) and imaginary part (eq. (76)) of the complex indicated angle. Let's see if this information can be used to calculate the angles of both targets. Looking at the equations indicates that there are two equations (real and imaginary part) and 4 unknowns $(\theta_a, \theta_b, p \text{ and } \Phi)$. Hence, this is an underdetermined equation system which results in an infinite number of possible solutions for the equation system.

So, if a solution with a single set of measurements is not possible, additional insight might be gained by using several measurements to solve the given equation system.

However, parameter between measurements have to change to achieve independent measurements in order to solve the given equation system.

Under the assumption that a relative motion between the two targets cause a change in the relative phase Φ the measurements are independent.

Now this results in 5 instead of 4 unknowns, the second relative phase being the 5^{th} unknown quantity. The power ratio and the real (x_1 x_2) and imaginary (y_1 y_2) parts of both measurement sets are the known quantities.

However, several assumptions have to be made. Hence, the usefulness is quite limited.

In [17] a maximum likelihood (ML) estimator is described for the case that the relative SNR between two targets is known and also several expectation maximization (EM) joint angle estimators depending on different prior knowledge assumptions.

10.2 Multipath

[3]

I already talked about interferences causing angle errors affecting monopulse angle estimation. Now it's time to dig into a really important topic called multipath.

Depending on the environment, wave propagation may happen in a direct path and additional reflected paths. The surface causing the reflection can lead either to a specular or diffuse reflection. Hence, the transmitted waveform arrives at different paths on the monopulse receiver. A special case of the multipath effect can be treated in the same way as unresolved targets. Although the effects caused by multipath are more general, they are also more deterministic in a sense that the additional contributions are related to geometry and surface of the reflection points. However, this would only be valid for specular multipath. At the same time, diffuse multipath contributions have to be treated statistically by means of a random variable (RV). Specular reflection at grass or other vegetation would also introduce some sort of random surface behaviour caused by wind and other weather conditions.

10.2.1 Flat-Earth Specular Model

[3]

In order to describe the effects of specular multipath I will start with the flat-earth specular model. This model assumes a flat and horizontal surface and the rays of the direct and reflected path are assumed to be parallel. Hence it is assumed that the transmitter is in the farfield. This is illustrated by Figure 32:

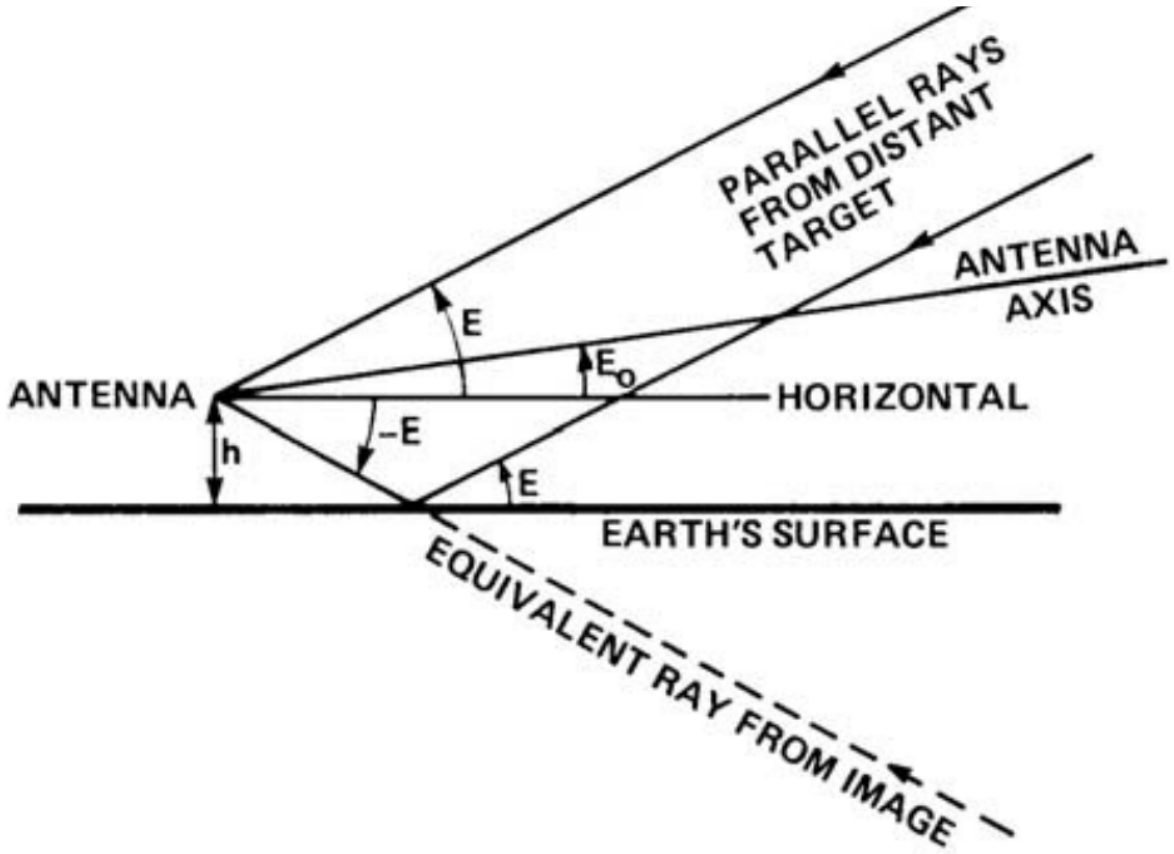

Figure 32: Flat-Earth specular model geometry (source: [3])

As also denoted in Figure 32 the reflected ray can be expressed by an image ray which is symmetric among the horizontal axis. E and E_0 represent the elevation angle from the target and antenna axis.

Now the target angles can be expressed by:

$$\theta_a = E - E_0 \tag{85}$$

and

$$\theta_b = -E - E_0 \tag{86}$$

where θ_a is the elevation angle of the target and θ_b is the elevation angle of the reflection. Substituting these two equations into equation (69) gives:

$$\theta_i = -E_0 + E\frac{1 - p e^{j\Phi}}{1 + p e^{j\Phi}} \tag{87}$$

Splitting into real- and imaginary parts gives:

$$x = \Re(\theta_i) = -E_0 + E\frac{1 - p^2}{1 + 2\,p\cos\Phi + p^2} \tag{88}$$

$$y = \Im(\theta_i) = -E\frac{2\,p\sin\Phi}{1 + 2\,p\cos\Phi + p^2} \tag{89}$$

Remember p denoted the amplitude ratio between the two targets or in this case the direct and reflected path.

The relative phase can be split into two components:

$$\Phi = \Phi_s + \Phi_p \tag{90}$$

where Φ_s is the phase change introduced by the reflection on the surface and Φ_p is the phase difference introduced by the different travelling distances of the direct and reflected ray.

The phase-shift introduced by the path differences can be calculated by using Figure 32:

$$\Phi_p = -4\pi(h_a/\lambda)\sin(E) \tag{91}$$

where h_a represents height of the antennas phase center.

It should be obvious that 180-degree phase-shifts in the reflected path can cause several minima in the elevation sum and difference patterns introduced by the interference with the reflected ray. In a worst-case scenario, it can affect detection and also introduce angle errors in elevation but it should not affect azimuth estimations (in case of only ground reflections) as long as the target can be detected in the sum pattern.

In our case ground reflections are less of a problem because this does not effect azimuth angle estimations.

10.2.2 Specular Multipath Countermeasures

[3]

Now let's dive into the different methods of reducing the effect of specular multipath on monopulse systems. In a pulsed radar system, there would be the possibility to differentiate between direct and reflected components by the different arrival times of their corresponding pulses and therefore their range.

10.2.2.1 Beam pattern designs

[3]

It should be obvious that low sidelobes and a narrow elevation beamwidth are beneficial in terms of multipath. For this system elevation angle estimations do not matter. Hence, a 1-dimensional array to obtain a narrow azimuth beamwidth is sufficient. However low sidelobes may not help at all if the multipath component is in the main-lobe. On the other hand, if the multipath component is in the direction of one of the sidelobes the benefit in terms of specular and also diffuse multipath behaviour can be significant. However, low sidelobes cause lower maximum gain in the main lobe. In addition to that the beamwidth of the main lobe increases which is undesired.

10.2.2.2 Making use of the complex indicated angle

[3]

As already denoted several times unresolved targets cause the monopulse ratio to become complex. The same goes for multipath propagation. Hence, the indicated angle can be seen as a complex quantity. Let's see if this complex angle can be used to counter multipath effects.

In order to take advantage of the complex angle you can gather a calibration curve. Let's take a look at an estimation of the complex angle with such a calibration curve:

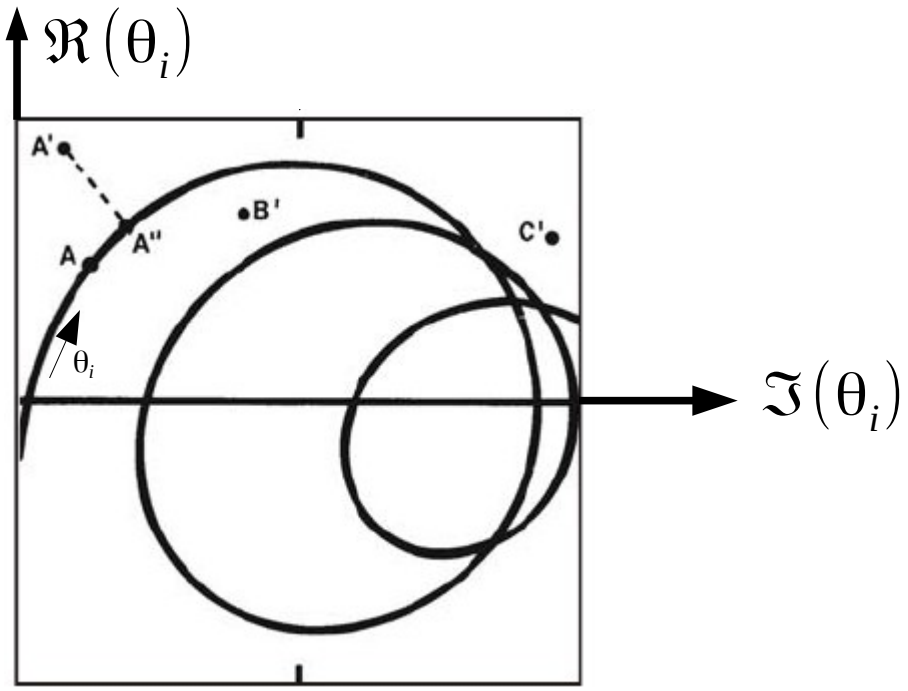

Figure 33: Estimating the complex indicated angle using a calibration curve

Figure 33 shows the actual complex angle A, the measured complex angle A' and the estimated angle A''. Where the x-axis represents the real and the y-axis the imaginary part. Using the calibration curve, better angle estimates can be achieved. However, as you can see from the curve, there are some angle ambiguities caused by the intersecting circles. In order to get rid of them smoothing, low antenna heights and some sort of diversity can be used. With frequency or boresight diversity slightly different multipath behaviour can be obtained

10.3 Monopulse angle errors

[3]

Errors caused by multiple targets were covered in Chapter 10.1, multipath effects in chapter 10.2 and this chapter deals with everything else.

10.3.1 Errors caused by noise

[3]

In case of low SNR, noise is one of the dominant sources of angle errors. Hence, additive noise is assumed in sum and difference channel:

$$s' = s + n_s \tag{92}$$

$$d' = d + n_d \tag{93}$$

where s' and d' are the noisy measurements of the sum and difference signal, d and s the noise free signals and n_s and n_d the noise contributions in the sum and difference channel.

The receiver itself should only produce independent noise contributions in the channels. Hence, the noise contributions can be treated as independent random variables (RVs).
However, leakage, coupling or interference can cause correlated noise contributions. Therefore, we should take them also into account.
This can be done by splitting the noise contribution into a correlated and uncorrelated part:

$$n_s = n_{su} + n_c \tag{94}$$

$$n_d = n_{du} + c\, n_c \tag{95}$$

where n_{su} and n_{du} are the uncorrelated contributions in the sum and difference channel and n_c is the correlated contribution. The factor c indicates that the correlated contribution might not be the same in both channels.
The distribution of noise in I/Q systems is usually described as circularly-symmetric

complex Gaussian (CSCG) with equal variance in the real and imaginary parts. The CSCG distribution is often called complex Gaussian (CG) or complex normal (CN) for short, especially in the signal processing and mobile communication literature.
Now a noise corrupted monopulse ratio can be obtained:

$$\frac{d'}{s'}=\frac{d+n_{du}+cn_c}{s+n_{su}+n_c} \tag{96}$$

Calculating the difference to the ideal ratio gives us the error:

$$\epsilon_{d/s}=\frac{d'}{s'}-\frac{d}{s} \tag{97}$$

$$\epsilon_{d/s}=\frac{n_{du}-(d/s)n_{su}+(c-d/s)n_c}{s+n_{su}+n_c} \tag{98}$$

From the monopulse ratio the corresponding angle error can be obtained. If the target is close to the boresight axis the nonlinear monopulse function is relatively linear. Therefore, the function can be linearized around the target angle (for angles near boresight). Hence, the error in monopulse ratio can be converted to the angle error by using the 3 dB-bandwidth θ_{3dB} and the monopulse slope k_m:

$$\frac{\epsilon_\theta}{\theta_{3dB}}=\frac{\epsilon_{d/s}}{k_m} \tag{99}$$

$$\epsilon_\theta=\frac{\theta_{3dB}\epsilon_{d/s}}{k_m} \tag{100}$$

ϵ_θ and $\epsilon_{d/s}$ represent the errors in angle and monopulse ratio, k_m the slope of the monopulse ratio and θ_{3dB} the 3 dB-bandwidth of the sum pattern.
After deriving the angle errors it is time to look at the bias and variance of this angle error.

10.3.2 Monopulse bias

[3]

The influences of noise in monopulse systems are covered in the preceding chapter. The noise variables follow a CSCG distribution with zero mean. Hence, there won't be any bias. However correlated noise components can introduce a “pulling” effect which causes the monopulse axis to pull away from the ideal axis. Therefore this correlated noise will introduce a bias.

To obtain the bias of a monopulse system the average of $\epsilon_{d/s}$ over the various distributions of the noise RVs have to be computed.

Let's take the uncorrelated noise component in the difference signal n_{du} as an example: the error contribution can be expressed by $n_{du}/(s + n_{su} + n_c)$. Taking the average over the noise density n_{du} while other noise variables are held constant gives zero. Other contributions behave the same. Hence, the contribution of n_{du} to the bias is zero.

Since the presence or absence of the uncorrelated difference noise component does not matter in terms of overall bias, the contribution can be treated separately.

By picking out the correlated component of (98) (setting other contributions to zero) the error in the monopulse ratio caused by n_c is:

$$\epsilon_{d/s} = \frac{(c - d/s)n_c}{s + n_c} = \frac{(c - d/s)(n_c/s)}{1 + n_c/s} \tag{101}$$

As already denoted the noise components follow a CSCG distribution. Hence, the magnitude of n_c follows a Rayleigh distribution with following PDF:

$$\mathrm{p}(|n_c|) = \frac{|n_c|}{N_c} \exp\left(-\frac{|n_c|^2}{2N_c}\right) \tag{102}$$

The complementary cumulative distribution function (CCDF) for $|n_c|$ exceeding $|s|$ can be obtained by integrating over the PDF:

$$P(|n_c| > |s|) = \int_{|s|}^{\infty} p(|n_c|)\, d|n_c| = \exp\left(-\frac{|s|^2}{2N_c}\right) \tag{103}$$

where N_c represents the average power of n_c.

Hence, the bias caused by n_c is the contribution in equation (101) multiplied by the fraction of time that $|n_c|$ exceeds $|s|$ (see CCDF in equation (103)):

$$bias_{n_c} = (c - \frac{d}{s}) \exp(-\frac{S}{N_c}) \tag{104}$$

For boresight contributions this results in:

$$bias_{n_c}(\frac{d}{s} = 0) = c \exp(-\frac{S}{N_c}) \tag{105}$$

The same procedure can be done for n_{su}:

$$bias_{n_{su}} = -(\frac{d}{s}) \exp(-\frac{S}{N_{su}}) \tag{106}$$

where again N_{su} is the average power of n_{su}.
In our case where we are only interested in the boresight angle the d/s ratio is zero (at boresight). Hence, the bias is also zero:

$$bias_{n_{su}}(\frac{d}{s} = 0) = 0 \tag{107}$$

At angles near boresight at moderate SNR the bias is negligible.

If none of the two derived contributions are negligible, the total bias can't be obtained by simply adding the two.

The total bias has been obtained by [12] and [3]:

$$Total\,bias = (\frac{c\,N_c}{N_s} - \frac{d}{s}) \exp(-\frac{S}{N_s}) = (\rho \sqrt{\frac{N_d}{N_s}} - \frac{d}{s}) \exp(-\frac{S}{N_s}) \tag{108}$$

where N_s is the overall noise in the sum channel, ρ the correlation coefficient between sum and difference channel.

10.3.3 PDF of d/s

[3]

We have already talked about the bias in this kind of systems. The next logical thing would be to talk about the variance. This causes some problems. The variance does not exist. In fact, the standard deviation and variance would be infinite. [3] [12]

I won't go into the details of these derivations. They are described in detail in [12].

One way to get around this would be to define a standard deviation for a Gaussian distribution which approximates the real pdf which is described in [12].

However, in practical systems the standard deviation won't be infinite. Finite dynamic range in hardware and software prevents it.

10.3.3.1 First order approximation for SNR >> 1

[3]

The derivation starts from equation (98). By dividing the numerator and denominator by s and expressing the noise by its normalized counterparts (e.g. $n_{du} \rightarrow n_{du}/s$) the equation becomes:

$$\epsilon_{d/s} = \frac{\frac{n_{du}}{s} - \frac{d}{s}\frac{n_{su}}{s} + \left(c - \frac{d}{s}\right)\frac{n_c}{s}}{1 + \frac{n_{su}}{s} + \frac{n_c}{s}} \tag{109}$$

The first order approximation assumes that the SNR in the sum channel is high enough so that noise components present in the sum channel can be neglected. Hence, the denominator noise terms can also be neglected:

$$\epsilon_{d/s} = \frac{n_{du}}{s} - \frac{d}{s}\frac{n_{su}}{s} + \left(c - \frac{d}{s}\right)\frac{n_c}{s} \tag{110}$$

Here are some thoughts about the equation above:

The components n_{du}, n_{su} and n_c are CSCG distributed with zero mean. Hence, $\epsilon_{d/s}$ is also zero mean. This means that this approximation does not consider any bias.

Since each contribution follows the CSCG distribution, $\epsilon_{d/s}$ also follows this distribution.

This equation can be converted to a more useable format by using several

relationships. The approximated variance of the d/s ratio is:

$$\sigma_{d/s}^2 = \frac{N_d}{2S} + \left(\frac{d}{s}\right)^2 \frac{N_s}{2S} - \frac{d}{s}\frac{c\,N_c}{S} \tag{111}$$

with 3 terms regarding the different noise contributions.

Hence, the standard deviation is:

$$\sigma_{d/s} = \frac{1}{\sqrt{2S/N_d}}\left[1 + \left(\frac{d}{s}\right)^2 \frac{N_s}{N_d} - 2\frac{d}{s}\frac{c\,N_c}{N_d}\right]^{1/2} \tag{112}$$

In case of that the correlated noise component is only caused by the noise of the local oscillator, the contribution of the correlated noise in the sum and difference channel is equal. The contributions of the uncorrelated components are also equal. Hence, the correlation coefficient $c=1$ and $N_s=N_d=N$:

$$\sigma_{d/s} = \frac{1}{\sqrt{2S/N}}\left[1 + \left(\frac{d}{s}\right)^2 - 2\frac{d}{s}\frac{N_c}{N}\right]^{1/2} \tag{113}$$

As you can see from the equation above the last two terms are only present in case of off-axis targets. The d/s factor in these terms causes them to vanish at boresight. Hence, the last two terms can be omitted near boresight for the sake of simplicity.

Converting the standard deviation to angle error results in:

$$\sigma_\theta = \frac{\theta_{bw}}{k_m\sqrt{2S/N}}\left[1 + \left(\bar{k_m}\frac{\theta}{\theta_{bw}}\right)^2 - 2\bar{k_m}\frac{\theta}{\theta_{bw}}\frac{N_c}{N}\right]^{1/2} \tag{114}$$

where $\bar{k_m}$ denotates the average of the monopulse slope beginning from boresight to target angle. S/N expresses the sum signal power to difference channel noise.

This equation can be split into on-axis and off-axis contributions. The on-axis component is:

$$\sigma_{\theta\,a} = \frac{\theta_{bw}}{k_m\sqrt{2S/N_d}} \tag{115}$$

The off-axis component is:

$$\sigma_{\theta b}=\frac{\theta}{\sqrt{2S/N_s}}\sqrt{1-\frac{s}{d}\frac{2cN_c}{N_s}} \tag{116}$$

Important to note is that for off-axis measurements both components are present. Hence, the standard deviation for off-axis measurements is:

$$\sigma_{\theta}=\sqrt{\sigma_{\theta a}^2+\sigma_{\theta b}^2} \tag{117}$$

However, averaging would decrease these figures.

11 Alternatives

Monopulse is not the only option. There are several alternatives which may or may not be viable. Each of these options has its own advantages and drawbacks.

11.1 Multilateration (MLAT)

[13] [14]

The principle behind MLAT is measuring the time differences (TDOA) with several receivers at known positions. Each receiver provides one hyperbola with possible positions. [15] Detecting the two turns in the model airplane competition requires the estimation of one coordinate. Hence, one nonlinear differential equation and therefore two receivers to obtain one time difference is necessary. However obtaining a least squares solution based on an overdetermined equation system can significantly improve performance. Solving such overdetermined non-linear least-squares problem can be achieved by the Gauss-Newton algorithm. Averaging and nonlinear Kalman Filters (extended Kalman Filter (EKF) or unscented Kalman Filter (UKF)) can also significantly improve such nonlinear estimation.

11.1.1 MLAT issues

[13] [14]

There are some drawbacks to this approach which are also dependent on the design decisions. Estimating one coordinate with two receivers can yield to ambiguous solutions. Considering multipath propagation this will get worse. Solving nonlinear equation systems is not feasible in embedded processors. However for higher number of receivers you can turn the nonlinear equation system into a linear one.

The accuracy of the timing and therefore the receiver clocks have a huge influence on measurement accuracy. Synchronisation between the transmitter and receiver clocks is also a big deal in terms of accuracy. The clocks of all receivers have to be synchronized. Even very small offsets in timing cause significant errors. Hence a very accurate timing source for permanent synchronization between all receivers is needed.

In addition to that all receivers positions have to be known in advance.

11.2 Other alternatives

There are several other viable alternatives. However Monopulse and MLAT are probably the most promising approaches. There are also other alternatives like triangulation (based on angle measurements) or optical approaches (camera or LIDAR based). However if high cost and high development effort don't matter, a viable option might be to combine several measurement systems using the sensor fusion capabilities of a Kalman filter.

Bibliography

[1] *“Sporting Code Section 4: F5 Electric”*, Internet: http://www.fai.org/downloads/ciam/SC4_F5_2017 [26.08.2017]
[2] Jürgen Burkhardt. *“FCD - FxxTrack”*, Internet: http://rc-network.de/forum/showthread.php/616342-FCD-FxxTrack-Reloaded [26.08.2017]
[3] Samuel M. Sherman, David K. Barton. *Monopulse Principles and Techniques*, 2011
[4] Brian D. Woerner, Theodore S. Rappaport, Jeffrey H. Reed. *Wireless Personal Communications: Research Development*
[6]Klaus W. Kark. *Antennen und Strahlungsfelder*, 2013
[8]Constantine A. Balanis. *Antenna Theory Analysis and Design*, 2016
[9]E. Jeff Holder. *Angle-of-Arrival Estimation Using Radar Interferometry*, 2014
[10]Texas Instruments. *High Speed Analog Design and Application Seminar*
[11]Rick Hartley. *RF / Microwave PCB Design and Layout*
[12]I. Kanter. *The Probability Density Function of the Monopulse Ratio for N Looks at a Combination of Constant and Rayleigh Targets*, 1977
[13]future airport 2/2016. *The future of Multilateration*
[14]W.H.L. Neven, T.J. Quilter, R. Weedon, R.A. Hogendoorn. *Wide Area Multilateration Report*
[15] Junichi Naganawa, Hiromi Miyazaki, and Hirohisa Tajima. *Detection Probability Estimation Model for Wide Area Multilateration*
[16]Brad Brannon. *Selecting ADCs for digital receiver applications*, 2002
[17]Zhen Wang, Abhijit Sinha Peter Willett, Yaakov Bar-Shalom. *Angle Estimation for Two Unresolved Targets with Monopulse Radar*
[18]Holger Heuermann. *Hochfrequenztechnik: Komponenten für High-Speed- und Hochfrequenzschaltungen*
[19]Pozar David. *Microwave Engineering*
[20]Rainee N. Simons. *Coplanar Waveguide Circuits, Components and Systems*