



NEAR EAST UNIVERSITY

Faculty of Engineering

Department of Electrical and Electronic Engineering

MICROSTRIP ANTENNA

Graduation Project EE- 400

Student: Sami Aljundi (992298)

**Supervisor: Assoc. Prof. Dr.
Sameer Ikhdair**

Nicosia - 2006

ACKNOWLEDGMENTS

First I want to thank Assoc. Prof. Dr. Sameer Ikhdair to be my advisor. Under his guidance, I successfully overcome many difficulties and learn a lot about Microstrip Antenna. In each discussion, he explained my questions patiently, and I felt my quick progress from his advice. He always helped me a lot either in my study or in my life.

Thanks to Faculty of engineering for having such a good computational environment.

I also want to thank my friends in NEU: Samer Salah, Hazem AbuSamra, Mohammed Sabri, Hani Aljundi, Omar Touqan and all my friends who helped me in this project.

Finally, I want to thank my family, especially my parents. Without their endless support and love for me, I would never achieve my current position. I wish my mother and my father lives healthy and happily always.

ABSTRACT

This is given by the fact that microstrip antennas are manufactured exploiting dielectric substrates, and therefore, analysis methods have to respect fulfilling continuity conditions on the boundary of two media. When studying dynamic electromagnetic fields, we usually start with the propagation of waves in an infinite homogeneous medium, which is supposed being linear and isotropic. In this situation, a single wave is propagating, which can be attenuated only if a loss medium is assumed. Electromagnetic waves, which propagate in infinite homogeneous medium, can be classified according to their wave-surface as planar, cylindrical and spherical. Plane waves can be observed in a long distance from the transmitting antenna. Plane wave propagation can be mathematically described by a scalar differential equation in the Cartesian coordinate system. Cylindrical wave propagates from an infinitely long direct wire, which is flown by a high-frequency current. Cylindrical wave propagation is described by a scalar differential equation in the cylindrical coordinate system. Spherical wave propagates from a point source. Spherical wave propagation is described by a vector differential equation in the spherical coordinate system. Propagation of the above-described waves is relatively well understandable and well imaginable. Even the mathematical relations describing propagation of those waves are relatively simple, and moreover, the final equations are of a closed form (i.e., we get relatively simple formulae which can be easily used for practical calculations).

INTRODUCTION

Antennas can be considered as one of the most important parts of the radio-communication chain because antennas transform an electromagnetic wave propagating along the transmission line to the wave propagating in free-space and vice versa. Consequently, antenna parameters (directivity pattern, impedance, gain) significantly influence final properties of the radio link. Analysis of microstrip antennas is much more complicated compared to the analysis of wire antennas

Chapter 1 (ANTENNA FUNDAMENTAL) is equipped by the basic knowledge of antenna and who would like to improve this knowledge and to apply it in the solution of practical problems.

Chapter 2 (MICROSTRIP ANTENNA) about microstrip transmission lines, microstrip dipole, and microstrip antenna you can find the applets visualize discussed abstract electromagnetic phenomena. Since patch antennas play a significant role in today's wireless communication systems, we pay our attention to its analysis in this special part of this chapter. In this chapter, then we turn our attention to the computational details of the analysis of electromagnetic field distribution in a shielded microstrip transmission line and microstrip dipole

Chapter 3 (ANALYSIS OF FREQUENCY SELECTIVE SURFACE) we take a detailed look on basis functions, which can be used when numerically analyzing frequency selective surfaces (current distribution over an element is sought).

In chapter 4 (DESIGN OF MICROSTRIP ANTENNA FOR WLAN) theoretical investigations have been conducted to evaluate nearly square single polarized wireless local area networks (WLAN) applications. The accuracy of the theoretical model assessed by a comparison with previous working and the agreements were very good. Further optimization and improvement is required to ensure correct circular polarization and wider bandwidth, also more rigorous model can be used. Also antenna array can be investigated to improve the performance.

TABLE OF CONTENTS

ACKNOWLEDGMENT	i
ABSTRACT	ii
INTRODUCTION	iii
1. ANTENNA FUANDEMENTAL	1
1.1. Properties and definitions	2
1.2. Isotropic radiator	2
1.3. Radiation direction	4
1.4. Directivity	4
1.5. Gain	5
1.6. Polarization	7
1.7. Polar radiation patterns	8
1.8. Omni-Directional Antennas	8
1.9. Power density, field strength, and impedance	10
1.10. Radiation mechanism	12
1.11. Effective area	13
1.12. The near field, the far field, and Raleigh distance	14
1.13. Radiation resistance and antenna impedance	15
1.14. Reciprocity	17
1.15. Bandwidth and broad banding	17
1.16. Antenna calculation	18
1.17. Antenna source current synthesis	19
1.18. Types of antenna	20
1.19. Aperture antennas	23
1.20. Array of antenna elements	25
1.21. Directivity-enhancement in array antenna	26
1.22. Very long baseline interferometer (VLBE)	27
1.23. Measurements and simulations	28
2. MICROSTRIP ANTENNA	30
2.1. Microstrip transmission line	30
2.1.1 Simplex Coordinates	42
2.2. Microstrip dipole	46
2.3. Microstrip antenna	54
2.3.1. Introduction	54
2.3.2. Analysis of stand-alone antenna element	56
2.3.2.1. Transmission line method	57
2.3.2.2. Moment method	59
2.4. Patch antenna	60

3. ANALYSIS OF FREQUENCY SELECTIVE SURFACES	68
3.1. Numeric analysis of frequency selective surfaces	68
3.2. Metallic elements	69
3.3. Problem formulation	70
3.4. Problem solution	72
3.5. Harmonic basis functions	73
3.6. Combination of harmonic functions and Chebyshev polynomials	74
3.7. Slot elements	77
3.8. Conclusion	79
4. DESIGN OF MICROSTRIP ANTENNA FOR WLAN	80
4.1. Abstract	80
4.2. Introduction	80
4.3. Polarization types	81
4.4. Linear Polarization	81
4.5. Bandwidth	83
4.6. Feeding Techniques	83
4.7. Aperture Coupled Microstrip Antennas	84
4.8. Substrate Thickness	85
4.9. Excitation	86
4.10. Results and Discussion	86
CONCLUSION	90
REFERENCES	92

CHAPTER 1

ANTENNA FUNDAMENTAL

An antenna or aerial is an electronic component designed to transmit or receive radio signals (and other electromagnetic waves). The words "antenna" (plural: antennas) and "aerial" are used interchangeably throughout this article. Physically, an antenna is an arrangement of conductors designed to radiate (transmit) an electromagnetic field in response to an applied alternating voltage and the associated alternating electric current, or to be placed into an electromagnetic field so that the field will induce an alternating current in the antenna and a voltage between its terminals. The origin of the word antenna relative to wireless apparatus is due to Marconi. In 1895, while testing early radio apparatus in the Swiss Alps at Salvan, Switzerland, in the Mont Blanc region, Marconi experimented with early wireless equipment. A 2.5-meter long pole, along which carried a wire, was used as a radiating and receiving aerial element. In Italian, a tent pole is known as l'antenna centrale, and this pole with a wire alongside it used as an aerial was simply called l'antenna. Until this time, wireless radiating transmitting and receiving elements was known simply as aerials, but Marconi's use of the word antenna, Italian for pole, soon came to be the most popular term for what today is uniformly known as an antenna.

There are two fundamental types of antennas, which, with reference to a specific three dimensional (usually horizontal or vertical) plane, are either omni-directional (radiate equally in the plane) or directional (radiates more in one direction than in the other). All antennas radiate some energy in all directions but careful construction results in large directivity in certain directions and negligible energy radiated in other directions. By adding additional conducting rods or coils (called elements) and varying their length, spacing, and orientation, an antenna with specific desired properties can be created, such as a Yagi-Uda Antenna (often abbreviated to "Yagi"). Typically, antennas are designed to operate in a relatively narrow frequency range. The design criteria for receiving and transmitting antennas differ slightly, but generally an antenna can receive and transmit equally well. This property is called reciprocity. The vast majority of

antennas are simple vertical rods a quarter of a wavelength long. Such antennas are simple in construction, usually inexpensive, and both radiate in and receive from all horizontal directions (omni directional). One limitation of this antenna is that it does not radiate or receive in the direction in which the rod points. This region is called the antenna blind cone or null.

Antennas have practical use for the transmission and reception of radio frequency signals (radio, TV, etc.), which can travel over great distances at the speed of light, and pass through non-conducting walls (although often there is a variable signal reduction depending on the type of wall, and natural rock can be very reflective to radio signals).

1.1 Properties and Definitions

The purpose of a transmitting antenna is to radiate electromagnetic waves into "free space" (usually, but not necessarily, air). The power for this is supplied by a "feeder" which is often a length of transmission line or waveguide having a well-defined characteristic impedance. One can regard an antenna as a kind of "transducer" to turn generated electrical energy into radiating energy. The acoustic equivalent of an antenna is a loudspeaker, (or microphone in the case of a receive antenna).

Antennas are also used in "receive mode" to collect radiation from "free space" and deliver the energy contained in the propagating wave to the feeder and receiver. Usually, antennas are reciprocal devices; their essential properties do not depend on whether they are used as transmit or receive devices. So an efficient transmit antenna can also be used as an efficient receive antenna for the same kind of signal. The directional pattern also does not depend on the transmit or receive mode usage. These properties are collected together and called reciprocity.

1.2 Isotropic Radiator

An "isotropic radiator" has no preferred direction of radiation. It radiates uniformly in all directions over a sphere centred on the antenna. It is a reference radiator with which other antennas are compared. If the power supplied to the isotropic radiator is P watts, the energy density (watts per square metre) at a distance R metres from the centre of the

radiator is $P/(4\pi R^2)$. This is because we are spreading the power P uniformly across the area $(4\pi R^2)$ of a sphere of radius R . Since propagating electromagnetic waves are transverse, the electric and magnetic field vector directions are at right angles to each other and also at right angles to the direction of travel of the wave.

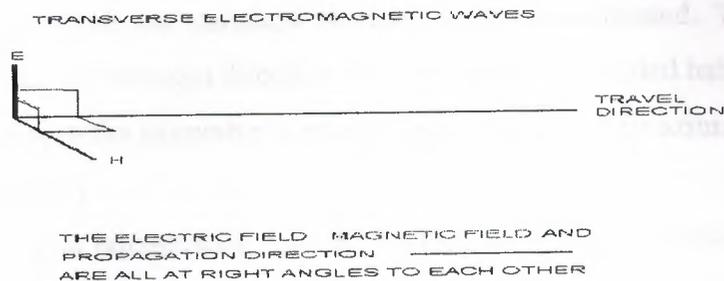


Figure 1.1 Transverse Electromagnetic waves.

Clearly then, we cannot realise an isotropic radiator in practice since there will be places on the unit sphere where we cannot specify a unique "polarisation direction" for the direction of the electric field. (For example, the lines of longitude on a sphere all meet at the poles, and the directions N and S are not defined at the poles). For this reason, it is impossible to construct, or even envisage, a perfect isotropic radiator. (If one allows more than one polarisation mode, one can approximate an isotropic radiator for some practical purposes. However, this can be argued to be a superposition of two antennas, and the polarisation properties are not isotropic even though the power summed over the two modes may be, approximately. To receive isotropic radiation from such an antenna over the sphere, one would have to construct a receive antenna whose polarisation properties depended on where it is located with respect to the source. An example of this dual-polarisation use is provided by the turnstile antenna, and tailored stacked turnstiles may be made a good approximation to an isotropic source in this sense.) It is however possible to have uniform radiation in all azimuth (see below) directions, or in all elevation directions at a particular azimuth plane.

Such an antenna, having uniform radiation azimuthally, is called "omnidirectional". This term is a misnomer, as the antenna is not isotropic and the radiation strength will

decrease if we increase the size of the elevation angle. Thus an "omni" antenna does not radiate equally in all directions.

1.3 Radiation Direction

Imagine you are standing upright on the ground. If you look straight ahead you are looking "along boresight". The boresight direction of an antenna is usually taken to be the direction along which the radiation is most highly concentrated. There can be, therefore, more than one boresight direction (eg, a vertically orientated half-wave dipole has uniform radiation in the azimuth plane, and any direction in this azimuth plane may be defined as boresight.)

If you look around you horizontally, you are looking in various "azimuth" directions. The azimuth angle varies from 0 to 360 degrees, allowing you to look in every horizontal direction. If you look up or down with respect to your local horizon, the angle of view up or down is termed the "elevation". The elevation angle varies from -90 degrees (straight down) to +90 degrees (overhead). We shall call the elevation angle "theta" and the azimuth angle "phi". The distance from the antenna is the radius "R". These are commonly termed "spherical polar co-ordinates". The set of angles phi (0 to 360 degrees) and theta(-90 to 90 degrees) allows us to specify any radiation direction uniquely. In the far field region (see below), the electric and magnetic fields fall off proportional to distance R; that is, they go as $1/R$. The power therefore falls off as $(1/R)^2$ and the total radiated power over the entire sphere surrounding the antenna is independent of distance R.

1.4 Directivity

Given a set of spherical polar co-ordinates (R, theta, phi) we can determine the power density in watts/(square metre) for both the antenna being investigated, and the isotropic reference antenna, which is radiating the same total power. The ratio of these power densities gives us the "directivity" of the unknown antenna in the direction (theta, phi) at a distance R from the antenna. If the direction (theta, phi) is not specified, the "directivity" is taken to be the maximum directivity of any of the directions of radiation. The quoted definition is the (directivity of an antenna is defined as the ratio of the

radiation intensity in a given direction from the antenna, to the radiation intensity averaged over all directions. This average radiation intensity is equal to the total power of the antenna divided by (4π) . If the direction is not specified, the directivity refers to the direction of maximum radiation intensity). We should caveat that this definition is for power radiated in the Far Field region. Here, the distance R from the antenna is taken out of the considerations by defining the radiation intensity as "the power radiated from an antenna per unit solid angle". For a directional antenna the radiation intensity will depend on direction from the antenna. However, in the far field it will not depend on the distance from the antenna. In both cases the power at large distances falls off with distance as $1/(R^2)$. This is called the "inverse square law" and is because the area of a sphere increases with distance as (R^2) . The "solid angle" subtended at the origin by an area A on a sphere is defined as $A/(R^2)$. The units of solid angle are called "steradians". There are 4π steradians in a complete surrounding surface of the antenna. The power radiated into any solid angle by an antenna is constant in the "far field region". Thus the directivity in the far field is not a function of the distance from the antenna.

1.5 Gain

The "gain" of an antenna takes account of the antenna efficiency as well as its directivity. The formal definition of gain in any direction (θ, ϕ) is "power density radiated in direction (θ, ϕ) divided by the power density which would have been radiated at (θ, ϕ) by a lossless (perfect) isotropic radiator having the same total accepted input power." If the direction is not specified, the value for gain is taken to mean the maximum value in any direction for that particular antenna, and the direction along which the gain is maximum is called the "antenna boresight". Allied to the concept of gain is the concept of efficiency. The efficiency of an antenna is the gain divided by the directivity, in any direction.

The efficiency and gain are limited by resistive losses in the antenna structure, and by resistive loss in objects, which may lie inside the "near field" region of the antenna.

The IEEE standards specifically exclude reductions in total transmitted signal arising from impedance mismatch (reflection loss) or polarisation mismatch. These reduce the

transmitted signal in any particular application by further amounts, and have to be considered in any link budget calculation. As an example of loss produced by objects close to the antenna radiating structure, there is a very substantial reduction in gain of the 10 element Yagi array antennas on the roof of BB building at about 4pm in the afternoon, when all the birds come and roost on the antennas. Birds have a high dielectric loss tangent.

Both Directivity and Gain are quantities relative to a reference antenna. Most often this is taken to be an isotropic antenna, which radiates equally in all (spherical) directions, and which cannot be made in practice. If the gain is expressed as dB it is usual to suffix the "isotropic reference" with the letter i; thus the gain with respect to an isotropic antenna in(dB) is called (dBi)

However, antenna designers often want to know how much better a given antenna (possibly a Yagi-Uda) is than a reference dipole antenna. A dipole antenna has directivity 2.2 (dBi). So it is frequently the case that we see the term (dBd), where the directivity is referenced (by the suffix letter d) to a perfect dipole. For the reference dipole, the gain and directivity are assumed to be the same, that is, the efficiency is assumed to be 100%. Thus the gain with respect to a perfect half-wave dipole antenna in dB is called (dBd)

Immediately, we see that there is the additive relationship

$$\text{Gain (dBd)} = \text{Gain (dBi)} - 2.2$$

Now since $10^{(2.2/10)} = 10^{(0.22)} = 1.6596$ we see that the numerical gain of a reference dipole is about 1.66 over isotropic, and that in numerical terms

$$\text{Gain (numerical over dipole)} \times 1.66 = \text{Gain (numerical over isotropic)}$$

These figures are good to a percent, which is adequate for practical antenna design and measurement purposes. Antenna gain is measured in either dBi or dBd.

It is important to note that antenna gain is different than amplifier gain. Antennas do not have a power source that allows the antenna to create additional energy to boost the signal. An antenna is similar to a reflective lens in principle - it takes the energy available from the source and focuses it over a wider or narrower area. Antenna gain is then a measure of the amount of focus that an antenna can apply to the incoming signal

relative to one of two reference dispersion patterns. dBi is the amount of focus applied by an antenna with respect to an "Isotropic Radiator" (a dispersion pattern that radiates the energy equally in all directions onto an imaginary sphere surrounding a point source). Thus an antenna with 2.1 dBi of gain focuses the energy so that some areas on an imaginary sphere surrounding the antenna will have 2.1 dB more signal strength than the strength of the strongest spot on the sphere around an Isotropic Radiator.

dBd refers to the antenna gain with respect to a reference dipole antenna. A reference dipole antenna is defined to have 2.15 dBi of gain. So converting between dBi and dBd is as simple as adding or subtracting 2.15 according to these formulas:

$$\text{dBi} = \text{dBd} + 2.15$$

$$\text{dBd} = \text{dBi} - 2.15$$

Specifying antenna gain in dBd means that the antenna in question has the ability to focus the energy x dB more than a dipole.

1.6 Polarization

The propagating wave has a transverse direction for the electric field called the "polarisation direction". This normally lies along the direction of electric field in the waveguide feed, or along the conducting driven rod element in a linear antenna.

It is of course possible to radiate from a conductor, which is not constructed in a straight line. However, there will still be a preferred polarisation direction.

The polarisation direction is necessarily at right angles to the line of sight joining the observer to the transmitting antenna. It is also at right angles to the magnetic field direction, which is also "transverse".

It is possible for the plane of polarisation, or the polarisation direction, to change with time, and to change with distance away from the source antenna. This is called "rotation of the plane of polarisation". If the antenna consists of a helix, or a crossed array of dipoles fed in quadrature, then the plane of polarisation can rotate one complete cycle every wavelength. The wave is then said to be "circularly polarised". It is possible to have right hand and left hand circular polarisation.

Clearly we can rotate the plane of polarisation in time and distance by spinning the antenna physically about an axis lying along the boresight direction.

1.7 Polar Radiation Patterns

The general dependence of directivity and gain on the angles (θ , ϕ) is called the "radiation pattern". In the case of a linear polarised antenna having fixed direction of polarisation, one can draw polar sectional plots in the "E-plane" and in the "H-plane". The E-plane contains the direction of propagation and the electric field vector. The H-plane contains the direction of propagation and the magnetic field vector. The E-plane is at right angles to the H-plane. E-plane and H-plane plots are normally regarded as sufficient to characterise an antenna. The radiated power density may fall into well-defined regions called "lobes", separated by regions of low intensity called "nulls". Strictly speaking the nulls can only be precisely zero intensity for particular directions (points from a continuous set). There is the "main lobe", which is usually where the wanted power from the antenna is directed, and "side lobes" where the antenna sends radiated energy which is regarded as "wasted" or may even interfere with other transmitting systems.

It is possible for there to be more than one main lobe having a given maximum value of gain. For example, a linear array of dipoles can have main lobes spaced 180 degrees apart, and both having the same gain.

1.8 Omni-Directional Antennas

Any radiating structure, which has rotational invariance around a vertical axis, will radiate equally in all directions in the horizontal plane, because there is nothing to define a preferred direction of (horizontal) radiation. Examples are a vertical whip antenna, or a vertical dipole, or a monopole over a ground plane. These antennas radiate with the electric field vertical, and the magnetic field horizontal.

OMNIDIRECTIONAL DIPOLE ROD ANTENNA

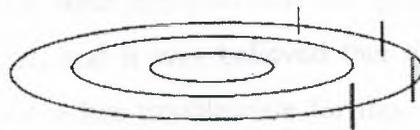


DIPOLE CURRENT ALONG VERTICAL RODS
 ELECTRIC FIELDS ALONG LONGITUDE
 PLANES, VERTICAL IN EQUATORIAL SECTION
 MAGNETIC FIELDS FORM HORIZONTAL LOOPS
 AND LIE IN LATITUDE PLANES
 RADIATION PATTERN IS DOUGHNUT SHAPE 

Figure 1.2 Omni-directional dipole rod antennas.

In the case of a horizontal loop or coil, the radiation is also omnidirectional but the magnetic field is vertical and the electric field is horizontal.

OMNIDIRECTIONAL HORIZONTAL LOOP ANTENNA



HORIZONTAL CURRENT LOOP
 ELECTRIC FIELDS FORM HORIZONTAL
 LOOPS IN LATITUDE PLANES
 MAGNETIC FIELDS LIE IN PLANES OF
 LONGITUDE, VERTICAL ON THE EQUATOR
 RADIATION PATTERN IS DOUGHNUT SHAPE 

Figure 1.3 Omni-directional horizontal loop antennas.

An alternative method of producing horizontally polarised (nearly) omnidirectional radiation is to use crossed horizontal dipoles fed in phase quadrature. Such an arrangement is called a turnstile antenna. Each dipole produces a characteristic figure-

of-eight radiation pattern in the horizontal plane; these are superposed in quadrature so the pattern, looked at from above, rotates about the axis once a cycle of radiation. Turnstile antennas also radiate circular polarisation vertically; the radiation may be concentrated in the horizontal plane by stacking turnstile antennas one above the other and feeding them in phase with each other. Turnstile antennas are commonly used as transmitting antennas when horizontal polarisation is required together with omnidirectional radiation.

In the early days of FM band II broadcasting, transmitters were horizontally polarised and the electric field was in the horizontal plane. This made reception on vertical whip antennas on motor vehicles unsatisfactory, as the polarisations were crossed. An ideal receiving antenna for this configuration would have been a horizontal loop above the top of the vehicle. Occasionally one can see such an arrangement on the streets. However, this has overheads of complexity and so the band II transmitters nowadays are slant polarised or elliptically polarised, so that there is a vertical component of electric field.

The horizontal polarization was adopted because it was found that the reception on antennas spaced around 10 metres from the ground could be maintained over a slightly greater service area than was the case when using vertical polarisation. Also it was found that the interference from unsuppressed car ignition systems was less for the horizontally polarised case, and it was believed that the multi-path reflections from aircraft flying overhead were less troublesome for this configuration. If one measures the signal strength of a band II transmission somewhere near a major airport, it is constantly fluctuating due to these multi-path effects from the moving reflecting surfaces of the aircraft.

1.9 Power Density, Field Strength, and Impedance

The power density in watts per square metre is numerically equal to the rms E field in the wave times the rms H field in the wave. We remember the rms values are 0.707 times the peak values, or $1/\sqrt{2}$ times the peak values.

We recall the SI unit of the electric field E is volts/metre, and the SI unit of the magnetic field H is amps/metre. Thus the product is (volts amps)/metre² or watts per square metre, as expected.

The characteristic impedance of free space (that is, vacuum or air) is 377 ohms or 120π ohms. This is the ratio of E field to H field. It is called Z_0 . Thus the power density in watts per square metre is $(H^2) \cdot Z_0$ or $(E^2)/Z_0$. The field strengths are therefore proportional to the square root of the power density, and they therefore fall off as $1/R$, or linearly with distance.

The field strengths therefore fall by a factor of 2 every time the distance from the antenna is doubled; the radiated power density falls by a factor of 4 (or 6dB) every time the distance from the antenna is doubled.

Exercise for the reader. Calculate the electric field strength on boresight, 100 kilometres away from a transmit antenna which has boresight gain 20dB and accepted input power 10 kilowatts. Compare the received voltage from a 1-metre length of wire antenna (assumed short compared to a wavelength) with the thermal noise voltage produced by a resistance of 75 ohms across a bandwidth of 10 kHz at a temperature of 300K. Hint. Boltzmann's constant $k = 1.38 \cdot 10^{-23}$ watts per degree K. Express the signal to noise ratio in dB.

The effective isotropic radiated power (e.i.r.p.) of the antenna is the power, which would have to be radiated by an isotropic source to give the same field strength as the real antenna under consideration, on, bore sight.

In this case, 20dB antennas gain = an increase in power on bore sight of a factor of 100. So the e.i.r.p is 10,000 times 100 watts, or 1 megawatt.

At $R = 100 \text{ km} = 10^5 \text{ meters}$, the radiated power density is $1 \text{ megawatt}/(4\pi R^2) = 7.96 \text{ microwatts/square meter}$.

This must equal $(E^2)/Z_0$ with Z_0 120π ohms, from which we deduce the rms electric field $E = 54.8 \text{ milli volts per meter}$.

A 1-metre length of wire antenna therefore picks up about 55 mV of signal. It passes this to the receiver input stage, which has noise voltage in the 75 ohms input resistance of $\sqrt{4 K T B 75}$ where B is the bandwidth in Hz.

The rms noise voltage is $\sqrt{1.24 \cdot 10^{-14}}$ volts or $1.11 \cdot 10^{-7}$ volts and the signal to noise voltage ratio is $4.9 \cdot 10^5$ or 114dB, an enormous amount.

If we put down the transmitter power to 1 watt, and the range up to 10,000 km, the signal to noise ratio reduces by a factor of $\sqrt{10^4}$ times $10^4/100 = 10^4$ and we can still communicate comfortably.

Considering deep space communications, it is possible to communicate with 1 watt and a 20dB antenna over a distance of the order of 100 million km, or from here to the sun.

1.10 Radiation Mechanism

Apart from quantum electronic devices such as lasers and masers, radiation is always the result of accelerated charge. Since ions are so much heavier than electrons, in practice this means that radiation is almost entirely due to accelerated electrons. These do not have to be "free electrons" or even "conduction electrons"; it is possible for radiation to occur from the acceleration of electrons bound to positive ionic charge in a dielectric. We note especially that radiation does not occur from the displacement current term in Maxwell's equations if this term does not include dielectric effects. That is, normally the relative dielectric constant of a physical medium is greater than unity because of polarisation effects (displacement of physical charge). There can be a contribution to the radiated fields from this bound charge, which is represented by the displacement current contribution $(\epsilon_{\text{relative}} - 1)\epsilon_0 dE/dt$. However, in a vacuum, the displacement current $\epsilon_0 dE/dt$ does not contribute to the radiated far fields. The contribution to the E field at large distances is proportional to the [amount of charge being accelerated] times [its acceleration]. This has dimensions Coulombs-metres/sec/sec, which are the same dimensions as the rate of changes of the quantity (IL) for a current I in a little length L of conductor. Thus we see, in order to increase the radiation from a short length of antenna, we either

- (1) Make the current change faster with time (i.e. raise the frequency), or
- (2) Increase the size of the current (i.e. deliver more power to the antenna), or
- (3) Increase the length of the antenna (i.e. make it longer up to $\lambda/2$ when the current direction reverses and starts to subtract from the radiated field.) Or more than one of these adjustments.

Thus we see that antennas become (on the whole) physically smaller as the frequency increases (wavelength decreases), and it becomes progressively harder to get them to radiate large total amounts of power.

1.11 Effective Area

The effective area multiplied by the wave incident power density in watts per square metre gives the total power delivered to the antenna's feeder. This is for a receive antenna.

The effective area A of an antenna is related to the boresight gain G and the free space wavelength λ of the radiation by the formula $G = (4 \pi A)/(\lambda^2)$. This is a most important formula.

A half-wave dipole has effective area of $0.13 \lambda^2$, which is roughly an area $\lambda/2$ by $\lambda/4$. The directivity of a half wave dipole, in the azimuth direction or H-plane, is about 1.67 or about 2.23 dB. Within elevation angles of size about 32.6 degrees the dipole has higher directivity than an isotropic source; outside this range it has lower directivity.

Considering an antenna as a transmitter, if it is fed with power P (accepted power) then the power density on boresight is $G P/(4 \pi R^2)$ at distance R . Here, G is a straight number calculated from the directivity and the efficiency. It is also possible to give the gain G in decibels; remember G is a power gain so in dB a gain G of 10 is 10dB, a gain of 100 is 20dB, a gain of 1000 is 30dB and so on.

If we transmit between two antennas each of gain G , spaced by a distance R , the field strength at the second due to the first is $G P/(4 \pi R^2)$ watts per square metre, and the effective area of the second is $A=G*(\lambda^2)/(4 \pi)$ so the total power transferred from transmitter to receiver is the product of these factors. The received power is therefore $P*(G \lambda)^2/(4 \pi R)^2$. This can be factorised into three parts as follows; the gain of the transmitting antenna times the gain of the receiving antenna times a "divergence factor" because not all of the power transmitted is picked up by the receiver. This latter factor is $[\lambda/(4 \pi R)]^2$ and the reciprocal of this, namely $[(4 \pi R)/\lambda]^2$ is often referred to as the "free space loss". We note that it is not really a "loss" as free space itself is a lossless propagating medium.

These antenna transmission formulae only apply in the far field region, so we need to know when we are in the far field.

1.12 The Near Field, the Far Field, and the Raleigh Distance

In the near field region, the polar radiation pattern depends on distance from the antenna and there is reactive power flow in and out of the region. One can imagine that the energy, instead of propagating uniformly and steadily away from the antenna, has an oscillatory longitudinal component. Energy is transferred to and from the near field region which represents the reactive part of the antenna driving point impedance. As one moves further away, this oscillatory energy flow reduces leaving just the regular power flow in the resistive characteristic impedance (377 ohms or 120 pi ohms) of free space.

In the far field the polar radiation pattern is completely independent of distance from the radiating source.

The transition from near to far field happens at the "Rayleigh distance", sometimes called the "far field distance". An estimate for this distance may be made from the formula $(2d^2)/\lambda$ where d is the maximum dimension of the radiating structure. In the case of a circular dish this is just the diameter; but in the case of a rectangular horn it is the diagonal distance across the mouth. This is only an estimate, and nothing suddenly happens at the far field distance thus estimated.

As an example, for a 10GHz antenna having dish diameter 30 cm, the wavelength is 3 cm and $2d^2/\lambda = 2*30*30/3$ cm or 6 metres. This is a sizeable distance compared to the dish dimensions.

If we consider the Tidbinbilla dish at 5GHz, shown elsewhere in this collection of notes, the wavelength is .05 metres and the diameter 75 metres so the Rayleigh distance is $2*75*75/0.05 = 225$ kilometres. Thus when the dish is pointing upwards we need to be above the atmosphere before we arrive at the far field region.

For this reason, it is impossible to measure the far field radiation pattern of a deep space antenna on a terrestrial antenna range. One has to resort to complete measurements of the near field response, and computer calculation to turn them into a far field pattern.

Alternatively one can measure the beamwidth by scanning across a small radio star. However it is often difficult to obtain reliable measurements of the sidelobe responses.

1.13 Radiation Resistance and Antenna Impedance

Part of the function of an antenna is to match the impedance of the feeder, or driving transmission line, to the impedance of free space.

A half wave dipole presents a resistive impedance of 73 ohms to a transmission line. It also has a small inductive reactance, of about 11 ohms. (The size of the reactive part depends on the length/diameter ratio of the rods of the antenna). The impedance close to resonance varies in a similar manner to a series tuned circuit. If the dipole is shortened from $\lambda/2$ there is additional series capacitive impedance and if it is cut too long there is additional series inductive impedance. Thus to make a dipole which has entirely resistive impedance it must be cut a few percent shorter than $\lambda/2$. The precise amount of shortening needed depends on the diameter of the rod elements. In general, the amounts of reactive impedance depend on the ratio of diameter to length of the antenna rods. Now we see why coaxial cable is often designed to have 75 ohms characteristic impedance. As the dipole is shortened, the radiation resistance falls sharply and it becomes a very inefficient radiator. For example, Elliott (op. cit. p304) has a calculation which indicates that for a rod-radius/wavelength ratio of 0.2% a dipole has to be cut about 5% shorter than a half-wavelength long to have a vanishing radiation reactance (that is, to present an entirely resistive impedance to the feed) and that then, its radiation resistance has fallen to about 63 ohms. To put this into perspective, at a band II frequency of 100 MHz, the wavelength is 3 metres and a rod-radius of 0.2% of a wavelength is 0.6 cm so the rod diameter is about 1.2 cm, which is typical for such an antenna. Incidentally, those of you who try to model dipoles using NEC (Numerical Electromagnetic Code) software may find that the returned real part of the driving point impedance stays close to 72 ohms, whatever the value of the diameter/length ratio of the rods. There is an argument that this is due to the stray capacitance between the closely-spaced ends of the thick rods. The combination of shunt capacitance with the radiation resistance in series with the residual inductance provides an impedance transformer, as is found in RF power amplifiers for example. This transformer steps up the actual

radiation resistance to a higher driving point resistance; at the same time the shunt capacitance resonates with the residual inductance. There is an argument that the radiation resistance which "matters" is the driving point resistance; however, we then find that this is critically dependent on the gap capacitance and varies with the spacing of the rods, and whether they are made from solid metal or tubes. Thus we see that these notions about the impedance of a half-wave dipole are only a guide to what we would measure in a practical installation. Indeed, a balanced feeder of characteristic impedance about 70 ohms is impracticable; so we have to incorporate some kind of balance-imbalance ("balun") transition between feed and antenna. The separation of the antenna rods also affects the total antenna length and the feed characteristics, and the physical feed structure and balun affect the near-field distribution of the dipole. It is thus possible to prefer cut-and-try methods for matching practical dipole antennas over the carefully calculated nostrums of antenna theorists. By the time the dipole length has reduced to $\lambda/10$ the radiation resistance has decreased to about 2 ohms and the reactance has increased to between 1 and 5 kilohms depending on the diameter of the rods. An infinitesimally short dipole is called a "Hertzian" dipole and is important theoretically since in practice all its properties may be calculated analytically. However, it is never used in practice because of its vanishingly low radiation resistance. For many purposes, calculations on a Hertzian dipole give a useful guide to the behaviour of longer dipoles. For practical reasons, particularly in mobile applications, it is necessary to cut dipoles short or to use monopoles loaded with inductance over a ground plane. The radiation resistance of a short dipole is given by the formula $R_{rad} = 20 \cdot (\pi \cdot L / \lambda)^2$ and for a $\lambda/8$ dipole is only 3 ohms. The series capacitive impedance for this length antenna may be as much as 1000 ohms; most of the transmission line voltage is lost across this capacitive reactance unless it is tuned out. One often sees short monopoles with a coil at the foot, to provide inductive tuning for this capacitive reactance. However, this is poor policy as it puts up the Q factor and reduces the bandwidth of the antenna. The tuning can be quite critical, especially in the presence of variable near-field obstacles.

1.14 Reciprocity

All the above properties of a linear passive antenna are identical whether it is used in transmit or receive mode. There is only one exception to this rule called "reciprocity", and that is when the antenna contains magnetically biased magnetic materials such as ferrites with resonantly rotating electron spin systems. The physical reason for reciprocity is that the only difference between outgoing and incoming waves lies in the arrow of time. Since the electromagnetic equations are invariant except for the signs of magnetic fields and currents, under time reversal, there can be no difference between transmit and receive mode in the physical current and field distributions. However, if we have a magnet providing a steady bias field, under time reversed conditions we would have to reverse the direction of this bias field. But for incoming and outgoing waves, the bias field direction remains the same. Thus it is possible for the system to be non-reciprocal. Of course, antennas containing amplifiers, or diodes, or spark gaps, may well not be reciprocal for obvious reasons. Also, practical antenna installations having metal-oxide-metal contacts, "rusty bolts", dry soldered joints and other electrical contact imperfections are also likely to behave differently under transmit and receive modes of operation.

1.15 Bandwidth and Broad-Banding

If we recall the definition of the Quality factor or Q factor as being the ratio of the stored energy to the energy dissipated per radian of oscillation, it is clear that in an antenna the part of dissipation is taken chiefly by the radiated energy. The stored energy is held in the near field region of the antenna structure. Since the fractional bandwidth $(\Delta f)/f$ is just the reciprocal of the Q factor, for a given radiated energy the Q will be smaller and the bandwidth larger if we minimise the amount of energy stored in the near field region of the antenna structure. One way of doing this is to make the antenna elements fatter in relation to their length. For a very fine wire antenna, the magnetic field for a given current rises as we approach the axis of the conductor, as $1/r$, where r is the radial distance out from the conductor. Thus making dipole antennas out of thick rods rather than thin wires is a good method of broad banding, up to a point. The biconical antenna, and its derivatives, the broad-banded Yagi, the bow-tie antenna and

the phantom conical antenna (which doesn't have a complete conical surface, but just conically disposed rod elements), is a good method of broad-banding a dipole type of antenna. There is a degenerate form of biconical antenna where the rods are arranged as an X with the upper \vee and lower \wedge fed as opposing arms of the dipole. This was very common in the early days of TV broadcasting, and was also relatively broadband compared to a simple dipole. A variant of this kind of X antenna had the upper and lower arms of the $>$ as a dipole, and the $<$ as a reflector. The radiation pattern had a maximum in the direction away from the reflector, but again the antenna structure was more broadband than a simple H antenna. It was also easier to construct. A rule of thumb is that a typical half wave dipole with sensible diameter rods has a fractional bandwidth of about 15%.

1.16 Antenna Calculations

Given a geometrical current distribution on the antenna structure, it is relatively straightforward to calculate the radiation integrals (see any good em theory textbook) to determine the radiation patterns. Often this has to be done numerically. . The difficulty with most antenna theory lies in determining the current distributions on the antenna conductors, given an arrangement of feeds, and the terminal voltages at the antenna ends of the feeds. The Hertzian dipole is artificial in that it assumes there is a uniform current density along the arms of the dipole. There is thus an unphysical current discontinuity at the ends of the arms, which cannot be realised in practice. However, given a uniform current distribution, the properties of the antenna are reasonably straightforward to calculate. An equivalent problem pertains in simple loop antennas. Here it is often assumed that the current is constant around the loop. That is only a reasonable assumption if the loop perimeter is short compared to a wavelength. There is capacitance between the opposite sides of the loop, which can carry displacement current, which results from a build up of charge due to the voltage drop around the loop, which has inductive impedance. One can easily see that although there is only a single continuous conductor, the current does not have to be the same everywhere around the loop, as some of it goes to charge the stray capacitance. Normally, self-consistent calculations are used to calculate together the current distributions and the radiated

fields. The "method of moments" is popular. In this method the antenna structure is split up into a number of regions, on each of which the current distribution is assumed uniform. The integral equations for the antenna then reduce to solving (what may be a quite large) matrix equation. This is well adapted to computer solution. There are issues of accuracy, and sensitivity to the model framework assumed. This method is also used to work out radar cross sections of complicated objects such as helicopters and aircraft. In a reflector-aperture antenna fed from the front by a sub-reflector and/or a feed, the far field radiation pattern can be calculated from the Fourier Transform of the field distribution across the aperture, accounting as well for phase variations across the illuminated area. The side lobe behaviour of a reflector antenna is particularly well suited to this calculation method.

We recall from Fourier Transform theory that sudden changes in a function give rise to the presence of high frequencies in the Fourier Transform. In this particular case the sudden change in spatial illumination gives rise to high spatial frequencies in the transform, which directs the energy well away from boresight as the "spatial frequency" translates into the deviation of the propagation direction from boresight. In array antennas, consisting of a number of identical or similar elements driven in synchronism, the problems are very similar to those encountered in filter design situations. Again, one needs to find the individual currents on each element. Often it is assumed that if the elements are identical, the currents must be also. However, consider the situation where there are four dipoles arranged in a straight line. Two of these will be end elements, and two will be interior elements. The coupling impedances between the adjacent elements will thus be different, and the currents necessarily different also.

1.17 Antenna Source Current Synthesis

Since the Fourier Transform method mentioned in the previous sections has an inverse, it is in principle possible to start from a knowledge of the required far-field radiation pattern, and also knowledge of the desired geometry of the source structure supporting the generating currents, and use inverse transform methods to derive the geometrical distribution of current amplitudes and phases on the source structure which would give rise to the desired far-field pattern. Of course, whilst this is in principle no more

difficult than is calculating the forward Fourier Transform to derive the far field pattern from the source currents, the difficulty, as always, from a practical point of view, comes in setting up the calculated amplitudes and phases of current elements on the specified source structure from an appropriate feed structure.

1.18 Types of Antenna

Here we list some of the common types of antenna. Apart from exotic applications, such as the banana tree, most antennas consist of a juxtaposition of conductor and insulator, which may be dielectric or it may be air or free space. This is not necessary; any structure, which will support a current on its surface, or guide or modify the direction of propagation of an electromagnetic wave, may be pressed into service as a kind of antenna.

Wire antennas. The wire need not be straight.

Loop antennas. The loop need not be circular. There can be more than one turn.

Rod antennas. The diameter of the rod is significant. These antennas include whip antennas and dipoles of all descriptions.

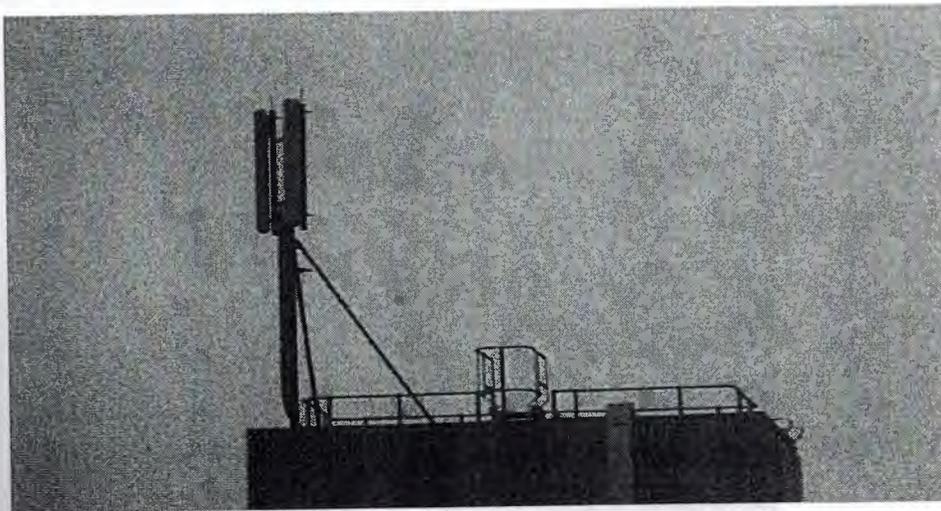


Figure 1.4 Mobile phone mast containing.

Aperture antennas. Examples are waveguide horns.

Slot antennas. These are holes in waveguide or cavities.

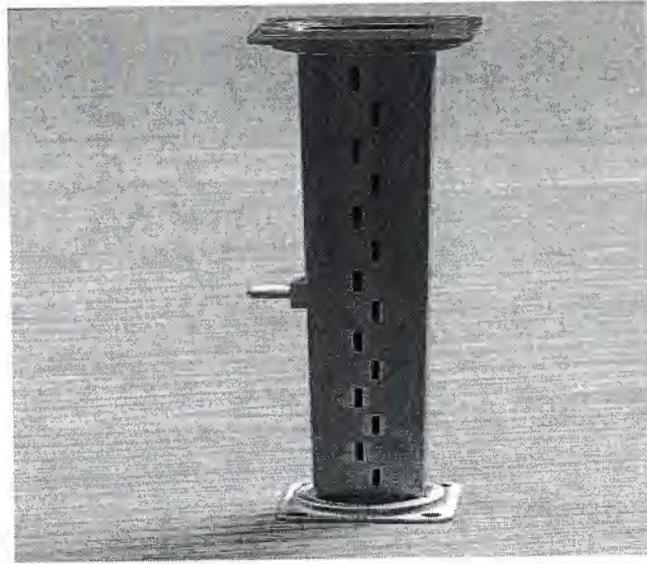


Figure 1.5 Slotted X-band waveguide array antenna.

Reflector antennas. These are used in combination with a "feed" formed from one of the other types.

Dielectrics are material structures containing bound electrons. Under the influence of an incident electromagnetic field, the electrons can move, accelerate, and therefore re-radiate. Thus, the presence of dielectrics close to an antenna conductor structure can profoundly modify the performance of the antenna.



Figure 1.6 Dielectric-rod ("polyrod") X-band antenna.

A dielectric lens may be used in front of a horn feed in a similar manner to a physical optical glass lens; but its action is then to modify the wave velocity, and therefore the curvature of the wavefront across the antenna. These are becoming increasingly popular for microwave applications, as they are small and easily fabricated. An area (almost any shape is possible) of conductor is excited on the surface of a dielectric substrate having a back plane conductor. The excitation can be by means of Microstrip transmission line, from either the front, or from the back through an aperture in the back plane. It can also be by means of front illumination from a horn feed; the patches are of different sizes and can mimic the phase profile of a parabolic reflector dish even though they are deposited on a flat plane surface. Microstrip antennas have substrate dielectric constant in the range 3 upwards. That means that there is more energy stored in the reactive near field region, so the antennas are narrow band high Q devices compared to other types of antenna. This is not so much of a problem at the higher microwave frequencies, where narrow fractional bandwidth still gives useful signal handling capacity.

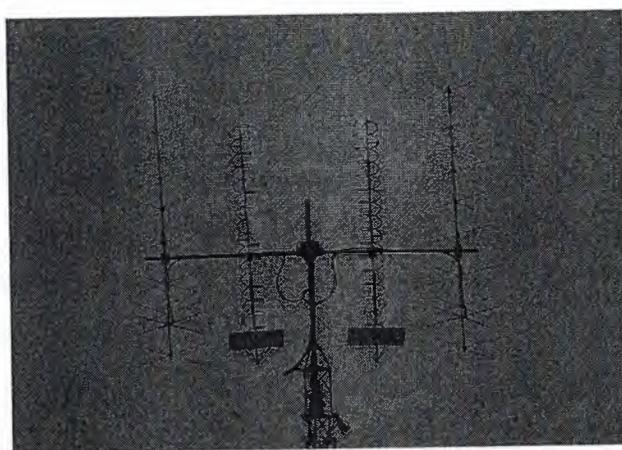


Figure 1.7 Crossed Yagi antennas for circular polarization and right-handed and left-handed helical antennas.

Log-periodic antennas. These are wideband antennas consisting of dipoles of successively diminishing length connected in parallel across the feed. Only that dipole which is close to a half wavelength long loads the feed; the dipoles behind and in front act as reflector and director to give the array a little gain.

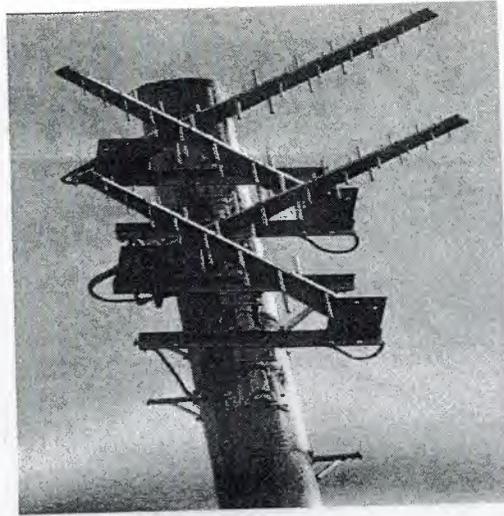


Figure 1.8 Log periodic antennas in Wales, 2005, for digital TV coverage.

These are leaky transmission lines whose wave velocity is close to that of waves in free space. The resulting "phase matching" condition allows resonant transfer from the transmission line to the free space wave. They can also be used in wideband applications if the transmission line is reasonably non-dispersive.

Active antennas. The individual transmitter modules form part of the radiating structure. This method is proposed for arrays. By altering the phase shift between successive elements in an array antenna, the bore sight direction may be steered electronically without physically moving the antenna structure.

1.19 Aperture Antennas

This class of antenna contains important technology for satcoms applications.

The simplest kind of aperture antenna consists of a tapered waveguide transition in the form of a "pyramidal horn".



Figure 1.9 18-dBi X-band pyramidal horn antenna on swivel mount.

The TE₁₀ mode in a rectangular waveguide has longitudinal components of magnetic field. As the waveguide is flared to form the horn pyramid, the longitudinal magnetic field components become less and, as the notes on wave-guides explain the characteristic impedance of the TE mode approaches that of free space, 377 ohms.

This kind of pyramidal horn aperture antenna is very important in the laboratory, as it is one of the few types of antenna whose boresight gain may be very accurately calculated (to within 0.1 dB). Consequently, it is used for producing reference field strengths and for calibrating the gains of other antennas.

A variant on the rectangular pyramidal horn is the circular horn feed. Such an aperture antenna is commonly used with a circularly symmetric waveguide mode (not the lowest mode in circular guide, NB) to produce uniform illumination of a Cassegrain antenna, which has a circular reflector dish of much larger diameter than the feed. The large reflector dish produces higher gain. The circular waveguide feed can also be used to produce circular polarisation.

Most ground based small broadcast satellite receiver dishes have a small horn feed of low gain placed at the focus of a dish between 0.5 and 1 metre diameter. Often the feed is offset from the boresight direction of the reflector dish; this "offset feed" arrangement

directs the main beam away from the feed, and this results in less blockage and improved sidelobe performance.

If the main beam in a cassegrain antenna hits the feed, or the sub-reflector, it will be diffracted around the obstacles and radiation will be scattered or diffracted into the sidelobe directions. The effective area of the dish is reduced, and the interference with other satellite systems from the sidelobes will be increased. This is not so important in deep space antennas. If we look at the Tidbinbilla deep space tracking antenna, we see there are two reflectors between the main feeds and the main beam. The sub-reflector at the focus of the large 75-metre dish is convex. The feeds are pointing along boresight, and are arranged to have a beam divergence angle, which is just sufficient completely to illuminate the sub-reflector. The sub-reflector returns the energy to the main reflector, and again the reflections are arranged so that there is minimal spillover at the edges of the dish, although maintaining uniform illumination as far as is possible.

The feeds are also conveniently located at the centre of the main dish, which moves little as the dish is steered. This has mechanical advantages, and makes the final HPA and LNA electronics more accessible for servicing.

Aperture antennas such as this are used in "very long baseline interferometry" methods. Here, two or more high gain large antennas, having large collecting areas, are separated by many hundreds of kilometres, and used to synthesise an aperture array having the diameter of the baseline separation of the dishes. Radio astronomers use these systems to pinpoint the location of radio sources to great accuracy in elevation and azimuth.

1.20 Arrays of Antenna Elements

If we want to increase the gain of a dipole antenna we can add another dipole antenna alongside it. This is the simplest form of array antenna.

First, we assume the antennas are fed in phase with each other and spaced $\lambda/2$ apart. Considering the radiation in a direction, which is normal to the plane containing the dipoles, the contribution from each element arrives in phase with the other. The field strength in this direction is double that for one element, so the radiated power density, which is the square of the field strength, is four times that for one element. However,

the two elements together are fed with twice the power of a single element. The increase in gain is therefore a factor $4/2 = 2$.

This calculation scales with the number of elements. If we use a 10 by 10 array, the boresight power gain is increased by a factor of 100, which is the number of elements. The field strength is 100 times more along boresight than for a single element, so the power density is 10,000 times greater. But 100 times the power is being fed to the array compared with a single element, so the gain increase is a factor of 100 as stated.

This gain increase is over and above any boresight gain of the individual elements. If we start off with an array of 100 horn feeds at 10GHz, of size 14 cm by 14 cm each, their intrinsic gain is about 20dB and the array factor gives an additional power gain of 100 which is 20dB so the combined structure has a boresight gain of 40dB or so.

Now consider, are we "getting something for nothing" or does this increased gain along the boresight come at the expense of gain elsewhere in the radiation pattern? The answer is clearly that the array concentrates the total radiated power along certain directions at the expense of others. If we go back to our 2 element dipole array, spaced $\lambda/2$, there can be no radiation along a line joining the centres of the two dipoles as their contributions are in anti-phase in this direction, there being a $\lambda/2$ path difference to get from one to the other.

In general then, the element pattern times the array pattern equals the total radiation pattern of the arrangement. What is the array pattern? It is the pattern you would observe for a set of isotropic radiators spaced as the array elements are actually spaced, and fed with the same amplitudes and phases of signals that the actual array elements receive.

1.21 Directivity-Enhancement in Array Antennas

The argument presented in the section above just happens to be correct when the elements are spaced by $\lambda/2$. However, for other spacings of the elements, we run into difficulties. Really what we should do is to integrate the array pattern over the sphere surrounding the antenna, to find the averaged isotropic radiated intensity, and then compare that with the intensity on boresight. This is the approach taken by

Constantine Balanis in his book on Antenna Theory and Design. He finds that the gain enhancement depends not only on the number of isotropes, but also on their spacing.

Clearly, if we have N isotropes collected together in a very small region much less than a wavelength across, then the radiation pattern (given uniform and equal excitations of each element) will be a good approximation to a sphere, and the maximum directivity must be just 1, as it is for a single isotropic element, and not N as predicted by the arguments above. What Balanis says is that a good approximation to the maximum directivity D_{max} is given by $D_{max} = 2N(d/\lambda)$ for N elements spaced d apart. We see that this reduces to N , quoted above, where d is equal to $(\lambda/2)$ and the spacing is a half-wavelength.

1.22 Very Long Baseline Interferometer (VLBI)

If we use two aperture antennas, spaced by a great many wavelengths, as an interferometer, the fringe spacing will be of the order of the angle subtended by an object of diameter one wavelength at a distance equal to the separation of the aperture antennas. For example, at 10GHz the free space wavelength is 3cm or 0.03m, so if we separate the antennas by 3000km or $1E8$ wavelengths, we can resolve radio sources about $1E-8$ radians across, or about 2 milliseconds of arc. By comparison, the beam width of one of the aperture antennas will be of the order of the angle subtended by a wavelength of radiation at a distance equal to the diameter of the reflector. Thus, if we considered a system where there were two 30 metre diameter antennas separated by 3000km, there would be $(3E6)/30 = 100,000$ interference fringes within the main beam of one of the apertures. Of course, the sensitivity of the interferometer is still governed by the total capture area of the two dishes; but the resolution is now comparable with that of a dish of diameter 3000km. The interference fringes from these two circular dishes will form parallel straight lines across the circular beam, as can be seen in the pictures below:

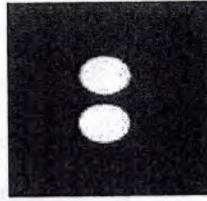


Figure 1.10 Two circular apertures spaced a distance apart.

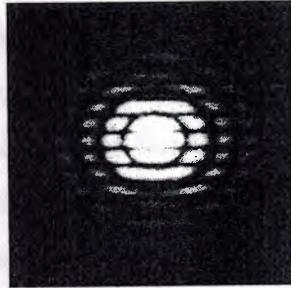


Figure 1.11 The interference bands from the apertures above.

In order to get resolution at right angles to these bands it is necessary to add a third aperture at the vertex of a triangle, whose base is delimited by the first two apertures.

1.23 Measurements and Simulations

Measurements on antennas are difficult. The behaviour of an antenna is best seen by monitoring the reflection coefficient with a network analyser over a band of frequencies, and for convenience a frequency about 1 GHz is appropriate. At 1 GHz a wavelength is 30 cms; the antenna is a reasonable size and it is possible to investigate the effects of adjacent objects, and different feed lengths, without too much difficult physical manipulation. The results may safely be transferred to other frequency bands by thought and analogy.

When this is done, one rapidly appreciates that an antenna can not be considered as a closed, isolated component having well-defined properties. Nearly every electronic measurement on an antenna is grossly affected by its environment and physical mounting. One might well ask the question, "What is an antenna?", or equivalently, "Where does the antenna stop and the outside world begin?". A sensible answer to this question is to consider all objects inside the near field as contributing to the radiation.

A helpful example is a Yagi-Uda antenna. We might regard this as a simple dipole with lots of resonant rods placed in the near field. But if we just consider the properties of the driven "antenna", namely the driven dipole, we know we will be grossly in error in assessing the performance of the installation. So why should we stop considering the effects of metallic structures at the end of the boom? We should add in the scattering from the mast, guys, feed (outer coaxial shields can carry induced current) and even dielectric objects (like the chimney stack or adjacent building) in the near field.

Many people now have access to software, which accurately simulates antenna behaviour. To do this it is necessary to construct a model. The process of "modelling" is critical to this enterprise as the simulation has limitations of accuracy depending on the kind of model chosen. In itself, the software is essentially accurate and useful. However, the results it returns, for simulation of real antennas, depends critically on what is built into the model. It is not usually possible, in the NEC2 and miniNEC and NEC4 software, to add in all the local effects, which will affect the results. This is not just because it is too difficult; there are difficulties in principle, knowing the correct dielectric and conductivity parameters to put in for a real-world installation. Details of the feed arrangement are also difficult to get right. So it is often difficult to know if the results from the simulation of the model represent the real behaviour of the antenna it was intended to investigate. The process of running the software always returns a result, and the internal checks on validity, while possible, are subtle. Belief in the results often dissolves into a matter of opinion or faith. This can be the subject of strongly-held views, which can only be resolved by recourse to measurements. Thus, simulation should be regarded (taking the most cautious view) as merely a rough guide to an antenna's behavior in a real installation. Any modelling process needs careful validation by measurements. One is then presented with the choice of which to believe, if there is disagreement.

CHAPTER 2

MICROSTRIP ANTENNA

2.1 Microstrip Transmission Lines

Shielded microstrip transmission line is a member of the family of planar microwave transmission lines. Open microstrip line (Fig. 2.1a), shielded slot-line (Fig. 2.1b) shielded fin-line (Fig. 2.1c) or coplanar waveguide (Fig. 2.1d) belong to the most common planar transmission lines.

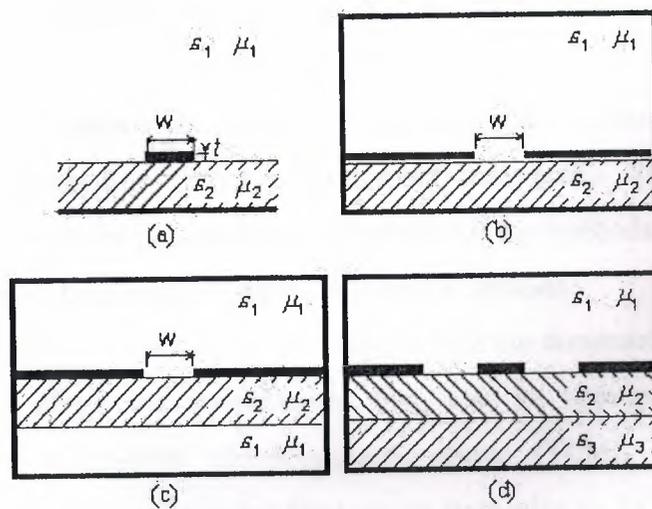


Figure 2.1 Selected types of planar microwave transmission lines: (a) open microstrip lines, (b) shielded slot-line, (c) fin-line, (d) shielded coplanar line.

Using segments of planar transmission lines, microwave circuits can be composed. Planar conductors can serve for connecting circuit components or can be used for creating passive circuit elements (capacitors and inductors, especially). Moreover, planar circuits can be relatively simply completed by active elements (transistors and diodes, e.g.). Therefore, the planar technology is of wide use these days and analysis of planar structures is of rising importance.

In this paragraph, we deal with the computation of the electromagnetic field distribution in shielded microstrip transmission line (Fig. 2.1), which parameters are assumed to be

constant in the longitudinal direction. Then, only a two-dimensional structure has to be analyzed (cross-section of the transmission line)

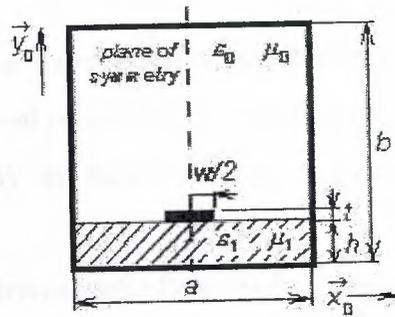


Figure 2.2 Shielded microstrip transmission line (longitudinally homogeneous).

Even if the above-described simplification is done, field distribution in the shielded microstrip transmission line cannot be computed analytically. Therefore, numerical methods or approximate ones have to be exploited. All the methods can be divided into two groups, to quasi-static methods and to full-wave methods.

Quasi-static methods are based on the assumption that the dominant mode of the wave, which propagates along the transmission line, can be approximated (with good accuracy) by the transversal electromagnetic wave (TEM). Unfortunately, this assumption is valid on low microwave frequencies (typically by 5 GHz). If frequency is increased, then the value of longitudinal components of electromagnetic field rises, and hence, it cannot be neglected.

Working on higher microwave frequencies, full-wave methods have to be exploited for the analysis. Full-wave methods are based on the direct solution of Maxwell equations. Since the transmission line is going to be analyzed in the harmonic steady state, Maxwell equations are of the following form:

$$\nabla \times E = -j\omega\mu_0\mu_r H \quad (2.1.1.a)$$

$$\nabla \times H = +j\omega\varepsilon_0\varepsilon_r E + \sigma E + j_s \quad (2.1.1.b)$$

$$\nabla \cdot (\varepsilon_0\varepsilon_r E) = \rho \quad (2.1.1.c)$$

$$\nabla \cdot (\mu_0\mu_r H) = 0 \quad (2.1.1.d)$$

Here, E denotes electric-field intensity vector, H is magnetic-field intensity vector, and operator nabla is in Cartesian coordinate system of the form

$$\nabla = x_0 \frac{\partial}{\partial x} + y_0 \frac{\partial}{\partial y} + z_0 \frac{\partial}{\partial z}$$

J_s denotes density of currents, which are imposed to the analyzed system by sources, ω is angular frequency ϵ_0 and μ_0 are permittivity and permeability of vacuum, ϵ_r and μ_r denote relative permittivity and permeability of dielectrics inside the analyzed structure, ρ is volume charge density in dielectrics and σ denotes electric conductivity of dielectrics.

Assume that sources of electromagnetic field are in a large distance from the area where computations are performed (then, imposed currents J_s are zero in this area). Next, zero charge density is supposed in the structure and the media are expected to be linear and isotropic (permittivity and permeability are scalar quantities, which do not depend on the value of respective field intensities). Further, dielectrics are assumed to be lossy (represented by electric conductivity σ) and all metallic parts (shielding waveguide, microstrip) are expected to be perfect electric conductors.

All the assumptions are substituted to (2.1.1). Then, both sides of (2.1.1.a) are multiplied by the operator nabla from left (in vector way). That way, curl of the magnetic-field intensity vector is obtained on the right-hand side of (2.1.1.a), which is replaced by the right-hand side of (2.1.1.b). Finally, the following equation is obtained:

$$\nabla \times (\nabla \times E) = -j\omega\mu_0\mu_r(\sigma + j\omega\epsilon_0\epsilon_r)E \quad (2.1.2)$$

If the square of wave number in vacuum is denoted as:

$$\kappa_0^2 = -j\omega\mu_0 j\omega\epsilon_0 \quad (2.1.3)$$

and complex relative permittivity of dielectrics as:

$$\tilde{\epsilon}_r = \frac{\sigma}{j\omega\epsilon_0} + \epsilon_r \quad (2.1.4)$$

Equation (2.1.2) can be rewritten to:

$$\nabla \times (\nabla \times E) - \tilde{\epsilon}_r\mu_r\kappa_0^2 E = 0 \quad (2.1.5)$$

Analyzed microstrip transmission line is situated to the Cartesian coordinate system (coordinates x and y in transversal directions, coordinate z longitudinal). Hence,

electromagnetic wave propagates along the axis z (along microstrip) and electric-field intensity vector depends on the longitudinal coordinate according to

$$E(x, y, z) = E(x, y) \exp(-\gamma z) \quad (2.1.6)$$

Where γ is propagation constant

$$\gamma = \beta + j\alpha \quad (2.1.7)$$

β is attenuation constant and α denotes phase constant.

Mathematical description of electromagnetic wave propagating in the direction z (2.1.6) is substituted to vector equation (2.1.5). Then, all the partial derivatives according to z can be evaluated (respective terms are multiplied by the propagation constant $-\gamma$). Rewriting all vectors in (2.1.5) as a sum of the transversal component and the longitudinal one

$$E = E_t + z_0 E_z \quad (2.1.8.a)$$

$$\nabla = \nabla_t + z_0 \frac{\partial}{\partial z} = \nabla_t - \gamma z_0 \quad (2.1.8.b)$$

We get

$$\nabla_t + (\nabla_t \times E_t) - \gamma(\nabla_t E_z + \gamma E_t) = \kappa_0^2 \mu_r \tilde{\epsilon}_r E_t \quad (2.1.9.a)$$

$$\nabla_t \times [(\nabla_t E_z + \gamma E_t) \times z_0] = \kappa_0^2 \mu_r \tilde{\epsilon}_r E_z z_0 \quad (2.1.9.b)$$

Where as (2.1.9.a) is vector equation of transversal components, (2.1.9.b) is scalar equation of longitudinal components.

The set of differential equations (2.1.9) has to be completed by boundary conditions, which have to be met by the solution of (2.1.9)

$$\begin{aligned} n_0 \times E_t &= 0 \\ E_z &= 0 \end{aligned} \quad |_{\Gamma_1}$$

(2.1.10.a)

$$\begin{bmatrix} \nabla_t E_z + \gamma E_t \\ \nabla_t \times E_t = 0 \end{bmatrix} \cdot n_0 = 0 \quad |_{\Gamma_1} \quad (2.1.10.b)$$

Equation (2.1.10.a) describes the fact that components of electric-field intensity vector, which are tangential to the perfectly electrical-conductive surfaces Γ_1 , have to be zero on those surfaces. Equation (2.1.10.b) expresses the fact that variation of components of

electric-field intensity vector in the normal direction with respect to perfectly electric-conductive surface Γ_2 has to be zero on this surface.

Equations (2.1.9) completed by boundary conditions (2.1.10) are the initial relations to the full-wave analysis of microstrip transmission line. We have to keep in mind, that equations (2.1.9), (2.1.10) cover the first Maxwell equation and the second one only, and therefore, the solution have to be checked to meet the third Maxwell equation and the fourth one. In the opposite case, the solution would be physically non-existing (spurious one).

In the following paragraphs, we describe the way of implementing the full-wave analysis of a shielded microstrip line, which is based on (2.1.9) and (2.1.10), by finite-element method.

First, we briefly comment physical phenomena in the investigated structure because success of the analysis is conditioned by their correct modeling.

First, consider a hollow rectangular waveguide (inside, vacuum is assumed). In such waveguide, a transversal electric wave (longitudinal component of magnetic intensity is present) or a transversal magnetic one (longitudinal component of electric intensity is present) can propagate. If a dielectric substrate is inserted into the waveguide (filling becomes inhomogeneous), longitudinal components of both the electric-field intensity and the magnetic-field one appear in the structure at the same time. In the opposite case, boundary conditions could not be met at the interface between dielectric and vacuum.

Wave propagation in a waveguide of inhomogeneous filling can be described by hybrid modes LSM (Longitudinal Section Magnetic) and LSE (Longitudinal Section Electric). Moreover, in inhomogeneous filling is completed by a metallic microstrip, transversal and longitudinal currents are induced there, which mutually couple LSE and LSM modes.

Numerical analysis of wave propagation in a metallic waveguide (vacuum inside) is relatively simple because homogeneous wave equation is needed to be solved for the longitudinal component of electric-field intensity (modes TM) or magnetic-field one (modes TE). If microstrip transmission line is going to be analyzed, attention has to be paid to two potential sources of spurious solutions.

First, production of spurious solutions can be caused by incorrect modeling of electromagnetic conditions on the boundary between the substrate and vacuum. Second, incorrectly modeled influence of metallic strips can cause appearance spurious solutions. Exploiting hybrid finite elements can eliminate both the causes.

Analyzing a shielded microstrip line by hybrid finite elements, all the components of electric-field intensity or magnetic-field one have to be included into the computations. Equations (2.1.9) are the initial relation of the analysis.

Matter of hybrid finite elements consists in modeling a longitudinal intensity component using nodal approximation, and transversal intensity components using an approximation, which is based edge vectors.

Turn our attention to the general solution of (2.1.9) under conditions (2.1.10) by finite-element method. At every step of the general approach, we discuss the operation of the method in our situation.

First, an analyzed structure is subdivided into sub-spaces (finite elements), which do not overlap and which cover all the points of the structure. In the space of a finite element, parameters of the analyzed structure (permittivity, permeability, conductivity) have to be constant. There are no restrictions to size and shape of finite elements. Finite-element mesh can be denser in areas where details of the solution are of our interest, and can be relatively sparse in areas where details are out of our interest. Finite elements can be curvilinear, and therefore, an arbitrary geometry can be modeled.

Analyzing longitudinally homogeneous microstrip transmission line, two-dimensional problem is solved. Therefore, we turn our attention to two-dimensional finite elements. Leaving piecewise constant approximation, a linear function is the simplest approximation function. In two dimensions, the linear function is a plane over a finite element (the computed quantity can be imagined to be drawn above the element).

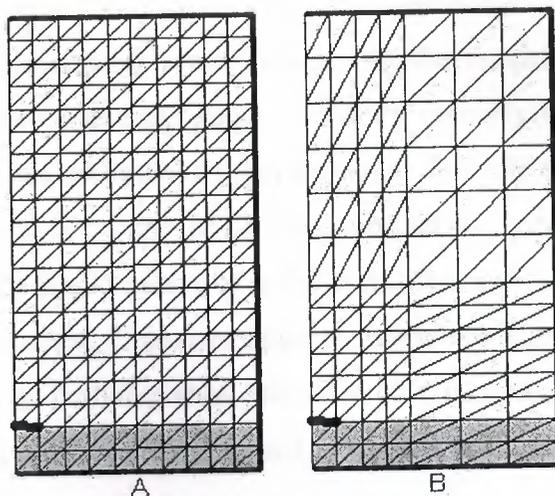


Figure 2.3 Mesh examples of rectangular bi-elements for the analysis of a shielded microstrip transmission line.

If three points uniquely determine the approximation plane, the analyzed structure is subdivided to triangular finite elements. Coordinates of vertexes of triangular finite elements are independent variables of the two-dimensional linear approximation. Samples of sought field distribution are dependent functional values. Interlaying the approximation plane through the dependent functional values in triangle vertexes, a unique approximation plane is obtained.

The analyzed structure has to be divided to finite elements very carefully because both the final error and CPU-time demands of the analysis strongly depend on the mesh geometry.

In (Fig. 2.3) two examples of triangular finite-element meshes for the analysis of shielded microstrip line are depicted. For simplicity, two triangular finite elements create a rectangular bi-element together, which is the basic building block of our mesh. Mesh A is homogeneous (consists of identical finite elements). Since electromagnetic field can be expected to quickly decrease with the distance from the microstrip, a sparser mesh B can cover distant areas. As shown later, the mesh B negligibly decreases accuracy of the analysis, but significantly reduces CPU-time demands (lower number of finite elements corresponds to lower number of mathematical operations).

In the second step of the solution, an approximation of a sought function over each finite element is expressed in a formal way. Usually, the unknown solution is

approximated by a linear combination of elected approximation functions and unknown approximation coefficients. A linear function is the simplest approximation. As an example, we consider a general plane, which is uniquely determined by functional values of the approximated function in three points of a finite element.

This general plane can be composed of three sub-planes (a linear combination), each of which is unitary in a single points of a finite element and which is zero in other two points of a finite element. Unknown approximation coefficients, which appear in the linear combination at partial planes, play the role of spatial samples of the spatial distribution of computed quantity in points, where partial functions are of unitary value (other functions are zero in this point, and therefore, they do not participate in sampling). For the longitudinal component of electric-field intensity, we get therefore

$$E_z^{(n)}(x, y) \approx \sum_{m=1}^{M_n} c_m N_m^{(n)}(x, y) \quad (2.1.11)$$

Where $E_z^{(n)}$ is an approximation of a sought function over n^{th} finite element, c_m are unknown approximation coefficients and $N_m^{(n)}$ are elected partial approximation functions on n^{th} finite element, M_n denotes the number of partial approximation functions exploited for composition of the total approximation of sought function over n^{th} finite element.

Collecting approximations over all N finite elements, a global approximation over the whole analyzed structure is obtained

$$\tilde{E}_z(x, y) \approx \sum_{n=1}^N E_z^{(n)}(x, y) \quad (2.1.12)$$

Equation (2.12) is a single equation for M unknown approximation coefficients c_m . If these coefficients are found, approximation of the sought quantity is obtained.

In the third step, the formal approximation of the solution \tilde{E}_z is substituted to the solved partial differential equation. Since the approximation differs from the exact solution, the initial equation is not met perfectly. This fact is respected by introducing a residual function, which equals to the difference between the exact solution and the approximation

$$R(x, y) = \tilde{E}_z(x, y) - E(x, y) \quad (2.1.13)$$

The approximation is as accurate as low values of the residual function are. Therefore, we are going to minimize the residual function over the whole analyzed structure. The residual function is minimized exploiting the method of weighted residuals.

The method of weighted residuals consists in multiplying the residual function $R(x, y)$ by a weighting function $W(x, y)$. The product is integrated over the whole analyzed space S and the result is set to equal zero

$$\iint_S R(x, y)W(x, y)dS = 0 \quad (2.1.14)$$

where $dS = dx dy$.

If as many properly elected weighting functions W_n are used as many unknown approximation coefficients are computed, a set of linear equations (as many equations as unknown approximation coefficients) is obtained. Solving this set of equations, unknown approximation coefficients are obtained. If partial approximation functions of nodes, where nodal quantity is unknown, are elected as weighting functions the method of weighted residuals comes to Galerkin method.

For n^{th} finite element, the final matrix equation is of the form

where

$$\begin{bmatrix} S_t^{(n)} - k_0^{(2)} \mu_r^{(n)} \tilde{\epsilon}_r^{(n)} T_t^{(n)} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_t^{(n)} \\ E_s^{(n)} \end{bmatrix} = \gamma^2 \begin{bmatrix} T_t^{(n)} & G^{(n)} \\ G^{(n)T} & S_z^{(n)} - k_0^{(2)} \mu_r^{(n)} \tilde{\epsilon}_r^{(n)} T_z^{(n)} \end{bmatrix} \begin{bmatrix} E_t^{(n)} \\ E_z^{(n)} \end{bmatrix} \quad (2.1.15.a)$$

$$T_t^{(n)} E_t^{(n)} = \sum_{i,j} \left\{ e_{ij}^{(n)} \iint_{S^{(n)}} [N_{t,rs}^{(n)} \cdot N_{t,ij}^{(n)}] dS \right\} \quad (2.1.15.a)$$

$$G^{(n)} E_z^{(n)} = \sum_m \left\{ e_{zm}^{(n)} \iint_{S^{(n)}} [N_{t,rs}^{(n)} \cdot (\nabla_t N_{z,m}^{(n)})] dS \right\} \quad (2.1.15.b)$$

$$S_z^{(n)} E_z^{(n)} = \sum_m \left\{ e_{zm}^{(n)} \iint_{S^{(n)}} [(\nabla_t N_{z,q}^{(n)}) \cdot (\nabla_t N_{z,m}^{(n)})] dS \right\} \quad (2.1.15.c)$$

$$T_z^{(n)} E_z^{(n)} = \sum_m \left\{ e_{z,m}^{(n)} \iint_{S^{(n)}} [N_{z,q}^{(n)} N_{z,m}^{(n)}] dS \right\} \quad (2.1.15.d)$$

$$S_t^{(n)} E_t^{(n)} = \sum_{i,j} \left\{ e_{t,ij}^{(n)} \iint_{S^{(n)}} [(\nabla_t \times N_{t,rs}^{(n)}) \cdot (\nabla_t \times N_{t,ij}^{(n)})] dS \right\} \quad (2.1.15.e)$$

In the above relations, $E_t^{(n)}$ is unknown column vector of edge approximation coefficients (approximation of transversal components of electric-field intensity) over n^{th} finite element and $E_z^{(n)}$ denotes column vector of three unknown nodal approximation coefficients (approximation of longitudinal component of electric-field intensity) over n^{th} finite element. Next, γ denotes complex propagation constant, k_0 is wave number in vacuum, $\mu_r^{(n)}$ is relative permeability of n^{th} finite element and $\epsilon_r^{(n)}$ is complex relative permittivity of the same element. Symbol dS denotes an elementary facet for the integration over n^{th} finite element and symbol $S^{(n)}$ give the total surface of n^{th} finite element. Summation including index m symbolizes addition over all nodes of a finite element (i.e., $m = 0, 1, 2$) a summation including indexes i, j symbolized addition over all edged of an element (i.e., $i, j = 0-1, 1-2, 2-0$). Symbols $e_{t, ij}^{(n)}$ are edge approximation coefficients, symbols $e_{z, m}^{(n)}$ are nodal approximation coefficients.

Matrices $T_t^{(n)}$, $G^{(n)}$, $S_z^{(n)}$, $T_z^{(n)}$ and $S_t^{(n)}$ are matrices of coefficients of n^{th} finite element of the size 3×3 . Elements of the above matrices are computed by integration of the product of partial approximation functions and weighting functions (or their derivatives) over n^{th} finite element. Matrices can be evaluated using the following relations

$$S_t^{(n)} = \frac{1}{A^{(n)}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (2.1.16.a)$$

$$T_t^{(n)} = \frac{1}{12} \sum_{i=0}^2 Q_i \cot g[\delta_i^{(n)}] \quad (2.1.16.b)$$

$$G^{(n)} = \frac{1}{6} \sum_{i=0}^2 C_i \cot g[\delta_i^{(n)}] \quad (2.1.16.c)$$

$$S_z^{(n)} = \frac{1}{2} \sum_{i=0}^2 D_i \cot g[\delta_i^{(n)}] \quad (2.1.16.d)$$

$$T_z^{(n)} = \frac{A^{(n)}}{12} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad (2.1.16.e)$$

where

$$Q_0 = \begin{bmatrix} 3 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix}, Q_1 = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & -1 & 1 \end{bmatrix}, Q_2 = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

$$C_0 = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{bmatrix}, C_1 = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ -2 & 2 & 0 \end{bmatrix}$$

$$D_0 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}, D_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}, D_2 = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$A^{(n)}$ is surface of n^{th} finite element and $\theta_i^{(n)}$ is angle at i^{th} vertex of n^{th} finite element.

Given relations are valid for the following organization of nodes and edges

$$E^{(n)} = [e_{z,0}^{(n)} e_{z,1}^{(n)} e_{z,2}^{(n)} e_{t,12}^{(n)} e_{t,20}^{(n)} e_{t,01}^{(n)}]^T \quad (2.1.16.f)$$

If the matrix equation is solved out for the vector of unknown approximation coefficients, then solution of the problem is obtained. Substituting approximation coefficients to the formal approximation, a real approximation of a sought function in each point of n^{th} finite element is obtained. Joining approximations over all finite elements, the global solution is found.

Local approximation of the longitudinal component of electric-field intensity vector over a finite element is expressed as a linear combination of elected partial approximation functions and unknown approximation coefficients. Considering linear approximation, the approximation plane over a finite element is composed of three partial approximation planes. Each partial approximation plane is unitary in a unique vertex of the triangle and is zero in the other two vertexes (Fig. 2.4). Coefficients c_n at partial functions in the linear combination play the role of spatial samples of the computed function in vertexes of the finite element (Fig. 2.4). Vertexes of the finite element are called nodes and respective functional values are called nodal values.

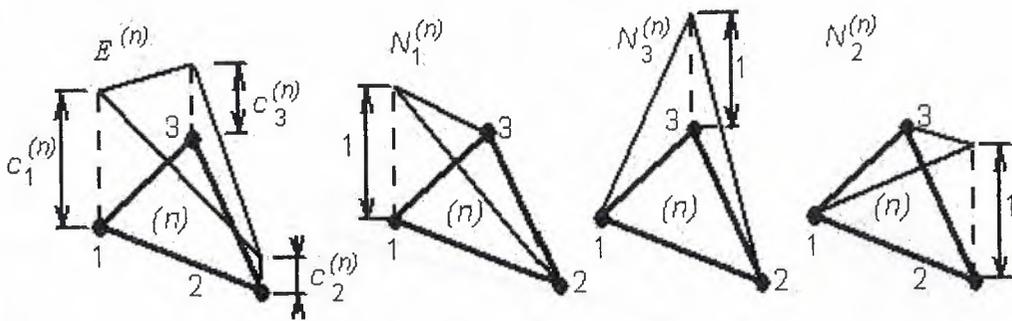


Figure 2.4 Linear approximation of (E) over finite element composed of three linear shape functions.

Partial approximation functions are called shape functions. All the shape functions, which are unitary in the same node (Fig. 2.5), compose together a basis function.

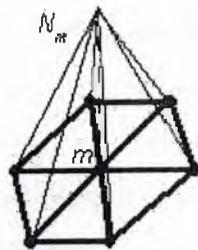


Figure 2.5 Linear function related to m -th node.

In many cases, approximation functions of higher order are more suitable than the linear function. Although the triangular element has to contain more nodes (6 for quadratic approximation, 10 for cubic one, etc.), the same error comparing to linear approximation is reached even if significantly lower number of finite elements is exploited. Approximation functions of higher order are smoother, and therefore, they represent better natural quantities.

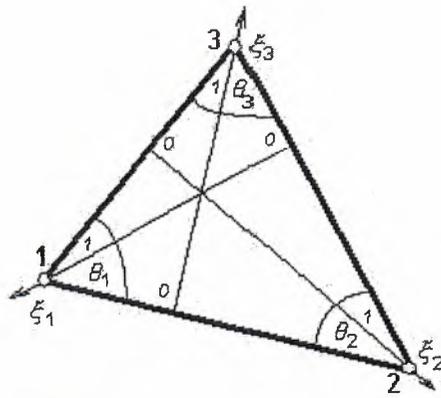


Figure 2.6 Two-dimensional simplex coordinates.

Now, we are familiar with shape functions, and therefore, we can find their proper mathematical representation. For this purpose, Lagrange polynomials expressed in simplex coordinates are usually used.

2.1.1 Simplex Coordinates

Considering triangular finite elements, simplex coordinate axes are of the direction of heights of the triangle. Simplex coordinates are unitary in the vertex and are zero on the opposite edge. Simplex coordinates do not depend neither on the shape nor on the dimensions of the finite element, and therefore, all the computations are sufficient to be performed once for a single finite element, and the results are recomputed for the other elements only.

Dealing with physical matter of simplex coordinates, a general point P inside a triangular finite element divides its surface to three partial triangles (Fig. 2.7). The ratio of the surface of a triangle, which is positioned in front of the first node, to the surface of the whole finite element equals to the simplex coordinate of P on the first simplex axis

$$\xi_1 = \frac{\sigma(S_1)}{\sigma(S)} \quad (2.1.17)$$

For other simplex coordinate axes, the situation is similar. In equation (2.1.17), $\sigma(S_1)$ denotes surface of the partial triangle, which is positioned in front of the first node, and

$\sigma(S)$ is surface of the whole finite element. Obviously, addition of all three simplex coordinates in an arbitrary point is unitary

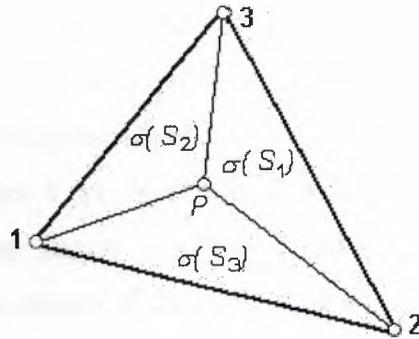


Figure 2.1.7 Matter of simplex coordinates.

$$\xi_1 + \xi_2 + \xi_3 = 1 \tag{2.1.18}$$

Now, we turn our attention to Lagrange interpolation polynomials. Lagrange polynomial of n^{th} can be expressed (using simplex coordinate ξ) as

$$R_m(n, \xi) = \frac{1}{m!} \prod_{k=0}^{m-1} (n\xi - k) \quad m \geq 1 \quad R_0(n, \xi) = 1 \tag{2.1.19}$$

Here, n is order of the approximation polynomial. Equation (2.1.19) describes the whole family of polynomials: family members differ in the index m , which can vary from zero to the order of polynomial n .

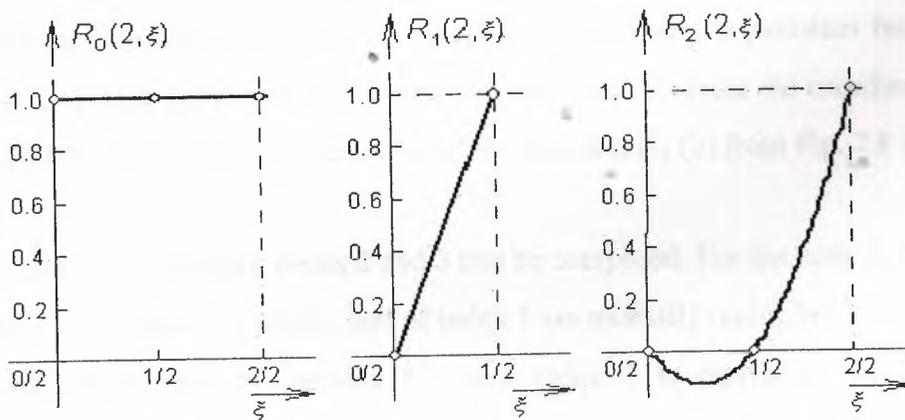


Figure 2.8 Family of Lagrange polynomials of 2nd order.

Nulls of polynomials $R_m(n)$ are equidistantly placed on coordinates $\xi = 0, 1/n$ till $(m-1)/n$, the polynomial is of unitary value in $\xi = m/n$. Hence, $R_m(n)$ is of m equidistantly placed nulls at the left from the coordinate $\xi = m/n$ and of zero nulls at the right.

(Fig. 2. 8) illustrates the above-given statement, where all the members of the family of quadratic polynomials $R(2)$ are depicted. Figure demonstrates the above-described equidistant distribution of nulls. The family member of index 0, i.e. $R_0(2)$, does not have any null at the left from the coordinate 0 and is of unitary value at the coordinate 0. The family member of index 1, i.e. $R_1(2)$, is of single null at the coordinate 0 and is unitary at $1/2$. Finally, the family member of index 2, i.e. $R_2(2)$, is of nulls at coordinates 0 and $1/2$ and is unitary at the coordinate 1.

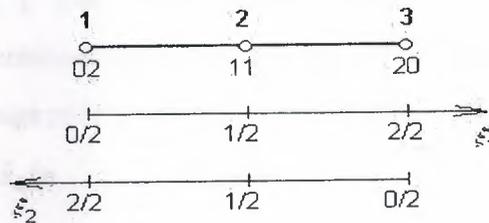


Figure 2.9 Shape function

Using Lagrange polynomials, we compose quadratic shape functions for one-dimensional finite element. ξ_1 is oriented from left to right on this element; the coordinate ξ_2 goes from the right to the left (Fig. 2.9). The shape function related to the node 1 (unitary value in the node 1, zero value in the nodes 2 and 3) is then composed by multiplying Lagrange polynomial of variable ξ_1 and index 0 (constant function of value 1) by Lagrange polynomial of variable ξ_2 and index 2 (since the coordinate ξ_2 is oriented from the right to the left, the course of the function $R_2(2)$ from Fig. 2.8 has to be reverted).

Similarly, shape functions for nodes 2 and 3 can be composed. For the node 2, Lagrange polynomials of variables ξ_1 and ξ_2 and of index 1 are mutually multiplied. For the node 3, Lagrange polynomial of variable ξ_1 and index 2 is multiplied by Lagrange polynomial of variable ξ_2 and index 0. Indexes of Lagrange polynomials, which form

shape functions of respective nodes, are written at these nodes (Fig. 2.9) in the form of a fraction; numerator is an index of Lagrange polynomial of the coordinate ξ_1 and denominator is an index of Lagrange polynomial of the coordinate ξ_2 . Adding numerator and denominator, order of approximation polynomial n has to be obtained. In general, the shape function of the node (i, j) of a one-dimensional finite element can be expressed as

$$\alpha_{ij} = R_i(n, \xi_1) R_j(n, \xi_2) \quad i + j = n \quad (2.1.20)$$

Here, n is order of an approximation polynomial, R denotes Lagrange polynomials defined by equation (2.1.19) and ξ are simplex coordinates.

In the next step, we turn our attention to a two-dimensional finite element. The only change, which has to be done, is adding a new simplex coordinate ξ_3 to two existing coordinates ξ_1 and ξ_2 . Two multiplicands, which appear in relations for shape functions of a one-dimensional finite element, are completed by the third multiplicand, corresponding to Lagrange polynomial of a new simplex coordinate ξ_3

$$\alpha_{ij} = R_i(n, \xi_1) R_j(n, \xi_2) R_k(n, \xi_3) \quad i + j + k = n \quad (2.1.21)$$

Here, ξ_1 , ξ_2 and ξ_3 denote simplex coordinates of a two-dimensional finite element, n is order of an approximation polynomial and R is Lagrange polynomials.

Substituting to (2.1.20) a (2.1.21), we get for linear approximation the following shape functions

$$\alpha_{100} = N_1^{(n)} = \xi_1 \quad \alpha_{010} = N_2^{(n)} = \xi_2 \quad \alpha_{001} = N_3^{(n)} = \xi_3 \quad (2.1.22)$$

Now, we are familiar with basis functions for the approximation of a scalar function E_z . Therefore, we can turn our attention to the approximation of vector function E_t . The approximation of a vector function formally corresponds to the approximation of a scalar function; only basis functions are of vector nature

$$E_t^{(n)} = N_{t,01}^{(n)} e_{t,01}^{(n)} + N_{t,12}^{(n)} e_{t,12}^{(n)} + N_{t,20}^{(n)} e_{t,20}^{(n)} = \sum_{i,j} N_{t,ij}^{(n)} e_{t,ij}^{(n)} \quad (2.1.23)$$

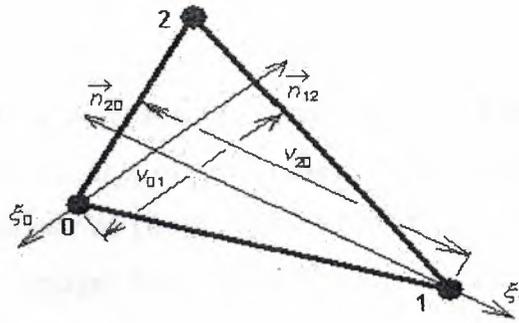


Figure 2.10 Explanation of behavior of vector shape function: for simplicity, superscripts (n) are missed.

2.2 Microstrip Dipole

As a microstrip dipole, we denote an antenna consisting of two narrow microstrip arms, which are fed by a symmetric source in the center. The antenna is placed on an upper side of the dielectric substrate. The bottom side of the substrate is fully covered by a metallic layer and I of zero potential (Fig. 2.11).

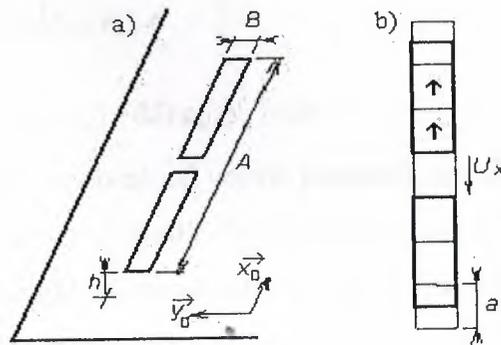


Figure 2.11 Microstrip dipole puls reflector: (a) global view, (b) discretization net for x component of current density

Effects of currents flowing on the antenna can be described by vector potential

$$A(r) = \iint_S \{G_A(r | r_0) \cdot J(r_0)\} dS_0 \quad (2.2.1.a)$$

Effects of charges on the antenna can be described by scalar potential

$$V(r) = \iint_S \{G_v(r | r_0) \rho(r_0)\} dS_0 \quad (2.2.1.b)$$

In the above-given equations, $A(r)$ denotes vector potential in the point r , J is current density vector in the point r_0 , G_A is dyadic Green function and G_v denoted scalar Green function. The parameter of Green functions $r|r_0$ tells us that we compute the contribution of a current (charge) from the point r_0 to potentials in the point r . The symbol ρ denotes charge density.

Current density and charge one are mutually associated by the continuity equation

$$-j\omega\rho = \nabla \cdot J \quad (2.2.1.c)$$

If both the vector potential and the scalar one are expressed on the surface of the microstrip dipole, we can evaluate electric intensity of a wave, which is radiated by the antenna.

$$E^S = -j\omega A - \nabla V \quad (2.2.1.d)$$

Applying (2.2.1) to the microstrip dipole from (Fig. 2.2.1) and substituting from continuity equation to (2.2.1.b), we obtain the following relations

$$A_x(x_m, y_n) = \iint_S \{G_A^{xx}(x_m, y_n | x', y') J_x(x', y')\} dx' dy' \quad (2.2.2.a)$$

$$V(x_m, y_n) = -\frac{1}{j\omega} \iint_S \{G_v(x_m, y_n | x', y') [\partial J_x(x', y') / \partial x]\} dx' dy' \quad (2.2.2.b)$$

$$E_x^S(x_m, y_n) = -j\omega A_x(x_m, y_n) - \partial J(x_m, y_n) / \partial x \quad (2.2.2.c)$$

Here, A_x denotes x -component of vector potential and V is scalar potential, G_A^{xx} denotes x diagonal element of dyadic Green function and G_v is scalar Green function, J_x is x -component of sought vector of current distribution and E_x is x -component of the radiated electric field intensity.

Substituting vector potential (2.2.2.a) and scalar one (2.2.2.b) to the relation (2.2.2.c), we obtain the initial equation for moment analysis of the dipole

$$E_x^S(x_m, y_n) = -j\omega \iint_S \{G_A^{xx}(x_m, y_n | x', y') J_x(x', y')\} dx' dy' + \frac{1}{j\omega} \iint_S \frac{\partial}{\partial x} \{G_v(x_m, y_n | x', y') \partial J_x(x', y') / \partial x\} dx' dy' \quad (2.2.3)$$

Magnitude of electric field intensity on the surface of the microstrip dipole can be computed using a boundary condition if perfect conductivity of metallic parts is assumed (except of the excitation gap, the intensity is zero). Therefore, x -component of current density J_x is the only unknown in (2.2.3).

In the first step, surface of the dipole is divided into discretization elements. Center of the first discretization element is denoted by 1, center of the second element by 2, etc. The upper bound of the discretization element is denoted by the same number completed by superscript "+" (plus), the lower bound by superscript "-" (minus); (Fig. 2.11).

Next, points in the center of discretization elements are used for computing x -component of electric field intensity. Since contribution of the current to the magnitude of electric field intensity is described by vector potential without presence of derivatives (equations 2.2.2), vector potential is computed in the center of elements.

On the contrary, contribution of charges to electric field intensity is described by scalar potential performing two derivations according to x . In numerical computations, derivations are replaced by central differences. Values of scalar potential V , which derivatives are used for determining contributions of charges to the electric intensity of the radiated wave, have to be known at the borders of discretization elements so that the result of central differentiating appears in the center of the element. Values of charge density are computed from continuity equation deriving components of current density in centers of discretization elements (derivatives are replaced by central differences again). In order to obtain values of charge density on the border of elements, we have to differentiate components of current distribution in the center of elements. This fact suits us very well because the computed values of current density are valid just in these points.

Finally, values of current density components are computed in the center of elements and values of charge density have to be evaluated on borders of the elements. Therefore, value of components of vector potential and value of components of electric field intensity are valid for the center of elements, and value of scalar potential for border of elements.

In the next step, we substitute piecewise constant approximation of current distribution to initial relations and we replace partial derivatives by central differences. Exploiting

continuity equation, charge density on the upper edge and on the low one of the discretization element is expressed

$$\rho(m_x^+, n_x) = \frac{-1}{j\omega} \frac{J_x(m_x + 1, n_x) - J_x(m_x, n_x)}{a} \quad (2.2.4.a)$$

$$\rho(m_x^-, n_x) = \frac{-1}{j\omega} \frac{J_x(m_x, n_x) - J_x(m_x, -1, n_x)}{a} \quad (2.2.4.b)$$

Symbol a height of the discretization element (Fig. 2.11), $J_x(m_x, n_x)$ corresponds to a constant value of x -component of current density vector on the surface of the element containing the central point (m_x, n_x) and ω is angular frequency.

On the basis of equations (2.2.4.a) and (2.2.4.b), we compute contribution of charges, which are represented by charge density ρ , to the x -component of electric field intensity vector. Considering charge densities of the upper edge $\rho(m_x^+, n_x)$ and on the low one $\rho(m_x^-, n_x)$, scalar potential on those edges can be computed. Moreover, substituting partial derivatives according to x by central differences, we obtain contribution of charges to x -component of electric field intensity.

Now, charge densities on borders of discretization elements are known and we assume that those values are valid not only on the borders but too over the whole surface of charge elements (they are of the same size as discretization elements but they are shifted so that borders of discretization elements can be in the center of charge ones as depicted in (Fig. 2.11). Then, charge densities can be described by the following piecewise constant functions

$$\rho(x, y) = \rho(x_m^+, y_n) \Pi(x_m^+, y_n | x, y) \quad x \in (x_m, x_{m+1}), y \in (y_n^-, y_n^+) \quad (2.2.5.a)$$

$$\rho(x, y) = \rho(x_m^-, y_n) \Pi(x_m^-, y_n | x, y) \quad x \in (x_{m-1}, x_m), y \in (y_n^-, y_n^+) \quad (2.2.5.b)$$

In these relations, $\Pi(x_m^+, y_n | x, y)$ denotes a function, which is unitary in the rectangular region with the center in (x_m^+, y_n) , a is height of the element and B is its width. The similar situation is for points (x_m^-, y_n) . Values of charge density $\rho(x_m^+, y_n)$ and $\rho(x_m^-, y_n)$ in the center of this rectangular region are given by relations (2.2.5).

If charge density on the microstrip dipole is known, we can substitute the distribution to

$$V(x, y) = -\frac{1}{j\omega} \iint_S \{G_v(x, y | x', y') \rho(x', y')\} dx' dy'$$

and we can compute scalar potential on respective charge elements. For the charge element, which center lies on the upper edge of the element (m_x, n_x) , we get

$$\begin{aligned}
 V(m_x^+, n_x) &= -\frac{1}{j\omega} \iint_S \left\{ G_v(m_x^+, n_x | x', y') \sum_{p,q} [\rho(p_x^+, q_x) \Pi(p_x^+, q_x | x', y')] \right\} dx' dy' \\
 &= -\frac{1}{j\omega} \sum_{p,q} \left\{ \rho(p_x^+, q_x) \int_{-\frac{b}{2}-\frac{a}{2}}^{+\frac{b}{2}+\frac{a}{2}} \int G_v(m_x^+ - p_x^+, n_x - q_x - | x', y') dx' dy' \right\}
 \end{aligned}
 \tag{2.2.6}$$

and similarly for $V(m_x^-, n_x)$.

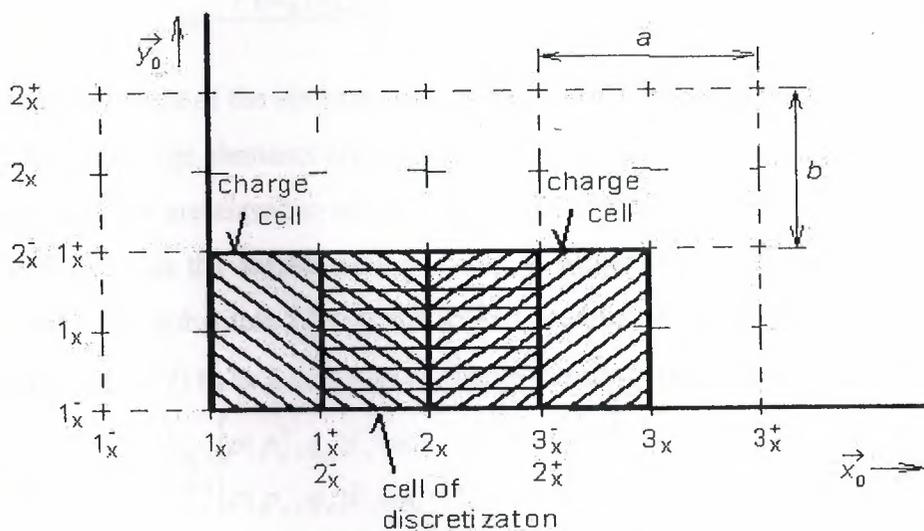


Figure 2.12 Validity region of charge density value (ρ), scalar potential (v) to the magnitude of electric intensity of wave, radiated by antenna

Before continuing, let us discuss the relation (2.2.6). Rearranging it, we swapped integration and summation, and the integral of the product of the unitary pulse Π and the scalar Green function G_v over the whole dipole was substituted by the integral of the single scalar Green function over the surface of this charge element, where Π is non-zero.

Dealing with indexes, (m, n) determines position of the *observation* element, over which the value of scalar potential is computed, and indexes (p, q) specifies position of the source element, whose charges contribute to the scalar potential of the element (m, n) .

Green function G_v is the only continuous function in (2.2.6), and therefore, the integral has to be evaluated for this function only. Evaluating this integral for various distances between the source element and the observation one, position of the observation element is changed only, and the source element stays in the origin of the coordinate system. Therefore, integration limits stay the same in all the cases (from $-a/2$ to $+a/2$ for the coordinate x' and from $-B/2$ to $+B/2$ for the coordinate y').

Therefore, constant values of scalar potential are known over all charge elements. Therefore, we can compute contribution of charges, which are represented by scalar potential, to the values of x -component of electric field intensity

$$E_x(m_x, n_x) = -\frac{\partial V}{\partial x} \cong -\frac{V(m_x^+, n_x) - V(m_x^-, n_x)}{a} \quad (2.2.7)$$

where, a denotes height of the element, B is the width of the element, and values of scalar potential V over charge elements are given by the relation (2.2.6).

As already said, we are aimed to express an approximation of electric field intensity of the radiated wave on the surface of the dipole as a function of current density on this dipole. Hence, we substitute (2.2.6) for scalar potential V on the border of elements, which enforces (2.2.7) to be a function of charge density ρ on edges of elements

$$E_x^s(m_x, n_x) = \frac{1}{j\omega} \left\{ \sum_{p,q} \left[\rho(p_x^+, q_x) \Gamma_v(m_x^+ - p_x^+, n_x - q_x) \right] - \left[\rho(p_x^-, q_x) \Gamma_v(m_x^- - p_x^-, n_x - q_x) \right] \right\} \quad (2.2.8)$$

From the above-given relation, a denotes height of element and B is its width. Angular frequency ω corresponds to the frequency, on which antenna is analyzed. The symbol E_x^s denotes a contribution of scalar potential V to the magnitude of x -component of the vector of electric intensity of the radiated wave. Finally, Γ_v represents integral of scalar Green function G_v over the surface of the element

$$\Gamma_v(m-p, n-q) = \int_{-B/2}^{+B/2} \int_{-a/2}^{+a/2} G_v(m-p, n-q | x', y') dx' dy' \quad (2.2.9)$$

As the next step in expressing electric field intensity as a function of current density, values of charge density from (2.2.5) are substituted to (2.2.8)

$$\begin{aligned}
{}^v E_x^S(m_x, n_x) = & + \frac{1}{\varpi^2 a^2} \sum_{p,q} \{ [J_x(p_x + 1, q_x) - J_x(p_x, q_x)] \Gamma_v(m_x^+ - p_x^+, n_x - q_x) \} - \\
& \frac{1}{\varpi^2 a^2} \sum_{p,q} \{ [J_x(p_x, q_x) - J_x(p_x - 1, q_x)] \Gamma_v(m_x^- - p_x^-, n_x - q_x) \}
\end{aligned} \quad (2.2.10)$$

The relation describes the contribution to the component of the vector of electric field intensity using unknown values of the component of current distribution vector J_x and known coefficients Γ_v , given by (2.2.9). From the point of view of scalar potential, the aim was reached, and therefore, the attention is turned to the vector potential.

In order to evaluate the contribution of the current to the electric field intensity, we have to compute vector potential substituting piecewise constant approximation of current density to (2.2.2.a)

$$\begin{aligned}
A_x(m_x, n_x) = & \iint_S \left\{ G_A^{xx}(m_x, n_x | x', y') \sum_{p,q} [J_x(p_x, q_x) \Pi(p_x, q_x | x', y')] \right\} dx' dy' \\
= & \sum_{p,q} \left\{ J_x(p_x, q_x) \int_{-B/2 - a/2}^{+B/2 + a/2} \int [G_A^{xx}(m_x - p_x, n_x - q_x | x', y')] dx' dy' \right\} \\
= & \sum_{p,q} \{ J_x(p_x, q_x) \Gamma_A^{xx}(m_x - p_x, n_x - q_x) \}
\end{aligned} \quad (2.2.11)$$

In the above-given relation,

$$\Gamma_A^{xx}(m - p, n - q) = \int_{-B/2 - a/2}^{+B/2 + a/2} \int G_A^{xx}(m - p, n - q | x', y') dx' dy' \quad (2.2.12)$$

Next, a is height of the discretization element and B is its width. G_A^{xx} denotes x diagonal component of dyadic Green function. Function $\Pi(p_x, q_x | x', y')$ is unitary over an element with the center in the point (p_x, q_x) and is zero elsewhere. Values J_x represent piecewise constant current density in the element mesh (p_x, q_x) .

During derivation, integration and summation were swapped, and integral over the whole dipole surface S was replaced by the integral over a single element (due to multiplying by Π the integrand is non-zero over a single discretization element only).

Finally, substituting vector potential to (2.2.3) and replacing derivatives of scalar potential by contributions (2.2.10), we obtain the final equation

$$\begin{aligned}
E_x^S(x_m, y_n) = & -j\omega \sum_{p,q} \{J_x(p_x, q_x) \Gamma_A^{xx}(m_x - p_x, n_x - q_x)\} + \\
& + \frac{1}{\omega^2 a^2} \sum_{p,q} \{J_x(p_x + 1, q_x) - J_x(p_x, q_x)\} \Gamma_v(m_x^+ - p_x^+, n_x - q_x) \} = \\
& - \frac{1}{\omega^2 a^2} \sum_{p,q} \{J_x(p_x, q_x) - J_x(p_x - 1, q_x)\} \Gamma_v(m_x^- - p_x^-, n_x - q_x) \}
\end{aligned} \tag{2.2.13}$$

Since equation (2.2.13) is rather complicated, we rewrite it into a matrix form

$$U_x = Z_{xx} I_x \tag{2.2.14}$$

Here, U_x is column vector of voltages in the direction x over elements. Voltages are computed by multiplying x -component of electric field intensity by x -dimension of the discretization element

$$U_x(m, n) = E_x(m, n)a \tag{2.2.15}$$

The symbol a denotes height of the discretization element (i.e. the element dimension in the direction x).

Since the microstrip dipole is assumed to be fabricated from perfect electric conductor, the vector of voltages consists of zeros only (except of excitation elements).

Next, I_x is column vector of unknown current in the direction x . Elements of I_x are related to current density J_x by the equation

$$I_x(m, n) = J_x(m, n)B \tag{2.2.16}$$

(B is width of the dipole, and therefore, even the width of the discretization element).

Impedance matrix Z_{xx} describes contributions of currents I_{xx} and contributions of charge densities ρ to voltages U_x over elements. Single elements of the impedance matrix Z_{xx} are obtained comparing (2.2.13) to (2.2.16)

$$Z_{xx}(m, n) = \frac{j\omega a}{B} \Gamma_A^{xx}(m, n) + \frac{1}{j\omega a B} \begin{bmatrix} \Gamma_v(m^+, n^+) - \Gamma_v(m^-, n^+) \\ -\Gamma_v(m^+, n^-) + \Gamma_v(m^-, n^-) \end{bmatrix} \tag{2.2.17}$$

In order to evaluate impedance matrix Z_{xx} , we have to know values of integrals of Green functions over the surface of the discretization element for different distances between source elements and observation ones.

2.3 Microstrip Antenna

We become familiar with the patch antenna in the paragraph describing the analysis of the wire dipole. Nevertheless, we left this antenna and concentrated on the one-dimensional moment method for the analysis of a dipole.

In the introductory part, architecture and properties of patch antennas are described. Then, two ways of the antenna analysis are explained. Finally, accuracy and computational demands of those two methods are compared.

2.3.1 Introduction

First notes on patch antennas appeared in the literature at the beginning of fifties of 20th century. This time, planar antennas, which can meet the surface of airplanes and military equipment, were demanded. Later, patch antennas were intensively exploited even in the civic are (in communication especially).

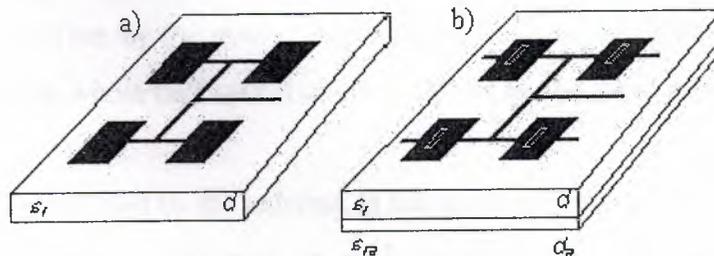


Figure 2.13 Classical patch antenna (a) aperture-fed patch antenna, (b) feeding microstrip at the bottom side of the substrate is drawn in gray, apertures in the ground plane in the black slots.

These days, a plenty of various types of patch antennas are at our disposal. These antennas differ in the shape of antenna elements (rectangular, circular), in polarization they operate with, in feeding. In our description, we concentrate on rectangular patch antennas fed by a microstrip transmission line. We subdivide path antennas into two groups depending if the feeding microstrip is placed on the same side of the substrate as



the antenna element (classical patch antenna, (Fig. 2.13) or not. The second (newer) case is going to be investigated more intensively.

Within last 15 years, indirect feeding of patch antennas is frequently used. For this purpose, small coupling slots in the ground plane are utilized. These slots are usually positioned in the center of antenna elements. Below the slot, where the feeding microstrip is placed, a certain (inductive) impedance can be observed. This impedance physically represents an antenna element, which exhausts energy from the microstrip transmission line. The microstrip crosses the slot and continues further (Fig. 2.13) in order to compensate the reactive part of the above-mentioned impedance by the shunt. Antennas, which are fed by the described way, are for simplicity called aperture-fed patch antennas.

Both groups of antennas exhibit some similarities and some differences (Table 2.1).

Small impedance bandwidth is a common property of patch antennas. For classical patch antennas, the bandwidth is about 2 to 3%, for aperture-fed antennas about 4 to 6%. High quality factor of patches is the reason. Patches behave as a resonator of a quality factor of several tens.

Antenna efficiency is another important parameter. Radiation efficiency (computed for loss less antenna) is given by the ratio of radiated power (integrating Pointing vector in the far region over the whole half-space) and real power on the input port of the antenna (element).

Energy, which is not radiated by the antenna, is taken away in the form of surface waves along the infinite dielectric substrate (in real situation, the substrate is limited, of course).

From the practical point of view, the total efficiency is important. Except of losses in dielectrics, surface-wave losses contribute to its value. Surface-wave losses are extremely undesired. Loss dielectrics guiding surface waves can be imagined as a loss transmission line. Exploiting a limited substrate, standing waves are formed in the substrate. Higher losses are related to higher standing wave ratio. Increase of losses due to the increase of standing wave ratio is called surface-wave losses.

The total efficiency of patch antennas is for 1 to 2 dB worse than the efficiency of reflector antennas.

Patch antennas exhibit very good cross-polarization. This parameter is not important for antennas operating linearly polarized waves. Designing a dual antenna exhibiting high polarization isolation, this parameter is of high importance.

	Classical antenna	Aperture-fed antenna	Note
Impedance bandwidth	2 to 3 %	4 to 6 %	In % of resonant frequency
Radiation efficiency	80 %	80 %	Surface-wave losses only
Total efficiency	40 to 60 %		Considering losses in dielectrics
Cross- polarization	35 to 40 dB		
Patch coupling	Indispensable	Small	
Feeder radiation	Middle (*)	Small	(*) Reducible by feeding topology

Table 2.1 Overview of parameters of patch antennas

2.3.2 Analysis of Stand-Alone Antenna Element

There are many methods for the analysis of a stand-alone antenna element. We are going to describe three methods, which are of the most frequent use. Two of the methods are going to be explained in detail.

The first method of the analysis understands the antenna element as a wide microstrip transmission line. As the radiation source, fringe electric fields are considered (Fig 3.2.2.a).

The second method of analysis consists in replacing the space below the antenna element by a loss resonator. Then, input impedance can be computed using the theory of cubical cavity resonators. In that case, zero normal component of electric field intensity is assumed on sidewalls of the resonator (Fig. 2.14). Hence, sidewalls can be understood as perfect magnetic conductors.

Numerical analysis of the antenna element is the third method. The numerical analysis is based on the moment method. Even here, we concentrate on the moment method in the spatial domain.

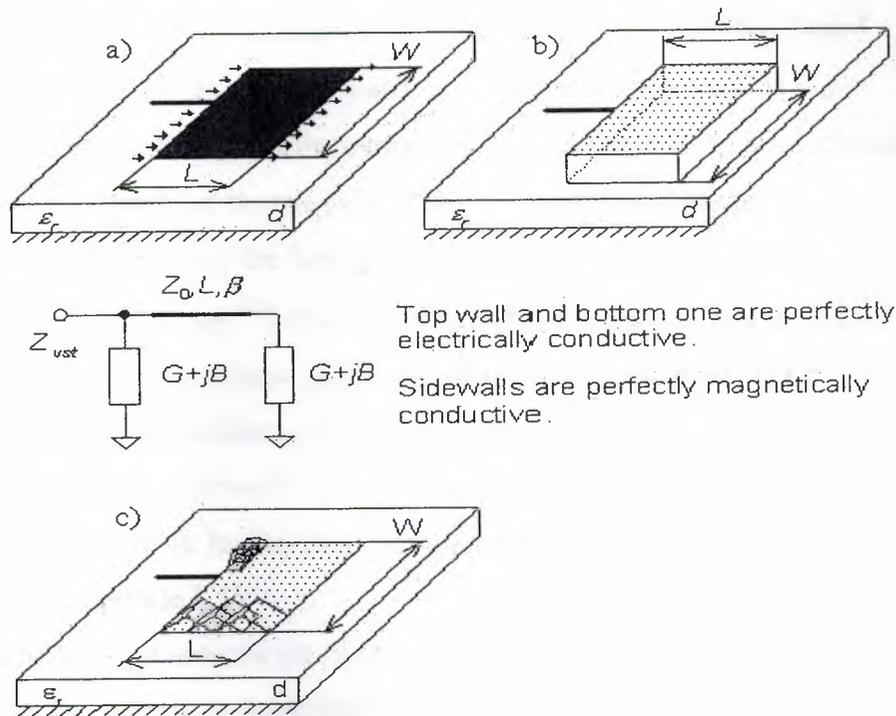


Figure 2.14 Schematics of models for the analysis of a stand-alone element (a) transmission line method, (b) cavity resonator, (c) moment method (small basis function)

2.3.2.1 Transmission Line Method

The method excels in simplicity on one hand, and enables to reach a good accuracy on the other hand. Accuracy of computing input impedance of the antenna element depends of the choice of antenna model. In this paragraph, the simplest model considering radiating slots between two neighboring antenna elements only is assumed (tangential fringe components of electric field intensity are considered as a single source of radiation). The other effects (mutual coupling among radiation sources, radiation of marginal edges) are neglected.

Designing the antenna, the admittance of the radiating slot Y_S (arrows in Fig. 2.14) has to be known

$$Y_S = \frac{\pi W}{\lambda_0 Z_0} \{1 + j[1 - \ln(2) \ln(k_0 w)]\} \quad (2.3.1)$$

Here, W denotes length of the slot [m], w width of the slot [m] (it equals to the height of the substrate d approximately), λ_0 denotes free-space wavelength [m], k_0 is free-space wave number ($k_0 = 2\pi / \lambda_0$) [m^{-1}] and Z_0 is free-space impedance [Ω].

In the relation, influence of the permittivity of the substrate is not included because the constant magnitude of electric field intensity is considered along the whole slot.

Designing the antenna, the length of the antenna element is chosen $(0.48 \div 0.49) \lambda / 2$, where $\lambda = c / [f (\epsilon_{\text{eff}})^{1/2}]$ and where the effective permittivity of the substrate ϵ_{eff} is computed using quasi-static relations or numerically (computing effective permittivity or characteristic impedance of microstrip transmission line, a proper computer program has to be chosen (several programs do not properly model the dispersion of the transmission line on higher frequencies; hence we recommend TX line from AWR, or TRL85 in Serenade 8.5SV).

The length of the antenna element does not equal exactly to the half of the wavelength because the capacitive impedance of the right-hand slot has to be transformed to the inductive impedance so that the capacitive impedance of the first slot on the frequency f can mutually compensate the inductive one.

In the second step, such width W of the antenna element is chosen so that the desired resonant resistance of the antenna element is obtained. This resistance can be expressed as $R_{\text{rez}} = (2 G_S)^{-1}$, where G_S is the real part of the slot admittance

$$G_S = \text{Re}\{Y_S\} = \frac{\pi W}{\lambda_0 Z_0}$$

Substituting free-space wavelength λ_0 and free-space characteristic impedance Z_0 , we get

$$R_{\text{res}} = \frac{120 \lambda_0}{(2W)} \quad (2.3.2)$$

Choosing $W = \lambda_0 / 2$, the resonant resistance equals to 120 Ω .

The described model of the antenna element is valid for electrically thin substrates only (we assume $d / \lambda \ll 0.01$, where d is the height of the substrate).

If the resonant resistance is required to be higher than 120Ω , the length of the antenna element W has to be smaller than one half of the wavelength. Resonant resistance cannot be significantly increased because its maximum value cannot be (respecting feasibility of the feeding microstrip) higher than 150Ω . If impedance transformer is exploited, then the antenna element of the impedance about 240Ω can be used.

Optimizing the width of the antenna element by extending its length is not recommended due to the spatial reasons and functional ones (parasitic resonance and parasitic lobes can appear). Satisfactorily small radiation resistance can be achieved by feeding the antenna element near its center instead of the edge feeding.

2.3.2.2 Moment Method

Applicability to the limited number of shapes of the antenna element and to the electrically thin substrates is drawback of the transmission-line method. In order to analyze elements of a general shape and elements on electrically thick substrates, the moment method has to be used.

The planar antenna element is assumed to be fed by a microstrip transmission line. Computing the input impedance of the antenna in the position of the microstrip input is the aim of the analysis.

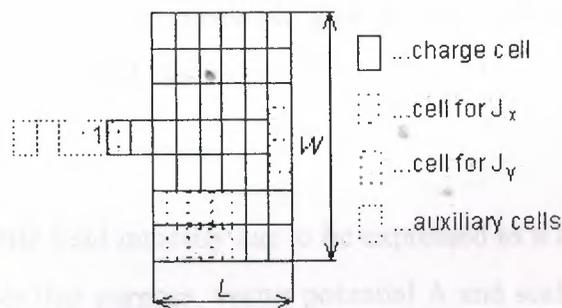


Figure 2.15 The mesh of current cells and charge ones

On the antenna element, there is a given current and charge distribution (in reality, a current is distributed over the antenna element). In order to compute the input impedance, currents and charges have to be computed first. Then, another quantity (electric field intensity, magnetic field intensity, input impedance, etc.) can be computed.

Since the moment analysis was described, a brief overview is presented here only:

1. The analyzed structure is divided to charge cells and current cells. The shape of cells is chosen as simple as possible (rectangular one). Next, both current and charge are assumed being constant over the cell (their value can differ from a cell to a cell). We concentrate on current cells because the charge can be computed using continuity equation. The mesh of current cells is shifted for one half of a cell with respect to charge cells.

2. We formulate a relation among electric field intensity in the center of i -th current cell and currents over all elements (including i -th one). Tangential components of electric field intensity in the center of elements are required to be zero except for source elements (the structure is assumed to be fed by flowing ideal voltage sources).

3. On the basis of computed current distribution over the antenna element, input impedance can be computed. Input impedance cannot be computed directly because the structure is fed by sources in the horizontal direction, which contacts the feeding microstrip by the first pole, and the auxiliary shunt by the second pole. The described voltage source is placed to the slot, which is denoted by one in (Fig. 2.15). The shunt is depicted by a dotted line. The ratio of a voltage and a current in the slot does not equal to the correct value of input impedance, and several additional steps have to be performed (we call them de-embedding).

2.4 Patch Antenna

As already said, electric field intensity has to be expressed as a function of currents and charges over cells. For that purpose, vector potential A and scalar potential φ are used. Moreover, the positive convention $\exp(+jkr)$ is considered

$$E = j\omega\mu A - \nabla\varphi \quad (2.4.1)$$

Vector potential A and scalar potential ϕ of elementary planar currents and charges physically represent contributions of sources to the electric field in a given point.

From the mathematical point of view, potentials are the solution of Sommerfeld integrals. For simple situations, potentials can be composed on the basis of the physical point of view. This approach is used even here.

Assume a two-dimensional antenna element over the ground plane (Fig. 2.16). Relative permittivity of the dielectrics is considered (for this moment) to be one, and contribution of an elementary current facet and charge one in the position r' to the intensity in r is computed. Using Coulomb law and respecting influence of ground plane (mirror principle), we get

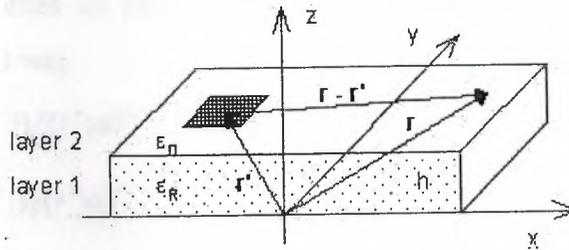


Figure 2.16 Coordinate system for computing contributions of elementary current elements

$$A(r) = \frac{\mu_0}{4\pi} \left[\frac{\exp(-jkr_0)}{r_0} - \frac{\exp(-jkr_1)}{r_1} \right] dr'$$

$$\phi(r) = \frac{1}{4\pi\epsilon_0} \left[\frac{\exp(-jkr_0)}{r_0} - \frac{\exp(-jkr_1)}{r_1} \right] dr'$$

where the distances r_0 and r_1 are

$$r_0 = \sqrt{(x-x')^2 + (y-y')^2}$$

$$r_1 = \sqrt{(x-x')^2 + (y-y')^2 + (2h)^2}$$

Here

$r = (x, y)$, $r' = (x', y')$ and h denotes the height of the substrate. Symbols $J[A/m]$ and $\sigma [C/m^2]$ denote current density and charge density.

Including the influence of the dielectrics is more complicated. The physical notion tells us that wave can propagate to the observation point r by the infinite number of paths:

$$\varphi(r) = \frac{1-\eta}{4\pi\epsilon} \left[\frac{\exp(-jk_0 r_0)}{r_0} - (1+\eta) \sum_{i=1}^{\infty} \frac{(-\eta)^{i-1} \exp(-jk_0 r_i)}{r_i} \right]$$

where

$$\eta = \frac{\epsilon_r - 1}{\epsilon_r + 1}$$

$$r_i = \sqrt{x^2 + y^2 + (2ih)^2}$$

and ϵ_r is relative permittivity of the substrate.

Vector potential A stays unchanged.

In the case of sources of finite dimension (rectangular cells), the contribution is expressed in integral way

$$A(r) = \iint G_A(|r-r'|) J(r') dS'' \quad (2.4.2)$$

$$\varphi(r) = \iint G_q(|r-r'|) q(r') dS' \quad (2.4.3)$$

where G_A and G_q are Green functions for magnetic potential and electric one, σ is charge density.

Now, equation (2.4.2) is discretized. In order to meet this task, a contribution of an arbitrary current (charge) cell to the potential in the center of another cell has to be expressed. Therefore, we denote

$$m_{ij}^{Ax} = \iint G_{xx}^A(|r_j-r'|) J_x(r') dS'' \quad (2.4.4.a)$$

$$m_{ij}^{Ay} = \iint G_{yy}^A(|r_j-r'|) J_y(r') dS' \quad (2.4.4.b)$$

$$m_{ij}^p = \iint G^q(|r_j-r'|) \sigma(r') dS'' \quad (2.4.5)$$

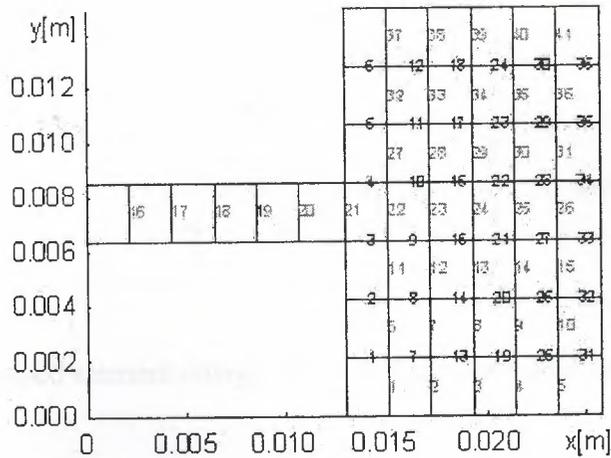


Figure 2.16 Example of charge-cell mesh

Expression m_{ij}^{Ax} tells how j -th current cell for x -component contributes to vector potential in the center of i -th cell. Similar meaning is observed at m_{ij}^{Ay} for J_y and at $m_{ij\phi}$ for scalar potential. Meshes of current cells and charge ones are mutually shifted for one half of a cell, and hence, a different numbering is used. Moreover, we independently number current cells for x -component and y -component. This fact is illustrated by numbering in (Fig. 2.4.2) which corresponds to the structure in (Fig. 2.16).

Composing equations for unknown current distribution, we have to know a function describing contributions between two arbitrary cells (of the same type). Due to the limited space, we do not describe this function here. We state only that all the contributions are evaluated at the beginning of the program. Moreover, several contributions repeat (e.g., cells 1 and 2 for the component J_x are of the same contribution as cells 3 and 4). The mutual contributions are stored in a moment table (matrix). Calling the function, a contribution on a respective position of the table is read.

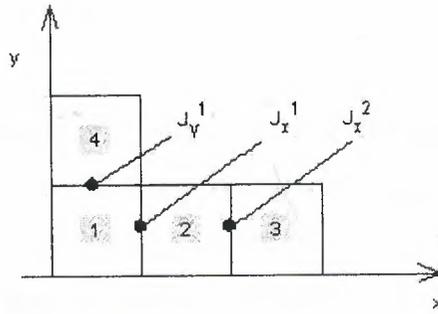


Figure 2.17 L-shaped element charge cell 1 to 4 current nodal values j_x^1, j_x^2, j_y^1 .

Now, the set of equations can be composed. We do not introduce a general relation, but a given simple situation is described. (Fig. 2.17) contains a L-shaped antenna element, consisting of four charge cells and three current cells (two of them are related to J_x , one of them to J_y).

Next, equations for E_1^x are built:

$$E_1^x = j\omega \sum_{j=1}^3 m_{1j}^{Ax} - \frac{1}{\Delta x} \left[\begin{aligned} &-\frac{j}{\omega} \frac{J_1^x - 0}{\Delta x} + \frac{J_1^y - 0}{\Delta y} (m_{21}^p - m_{11}^p) - \frac{j}{\omega} \frac{J_2^x - J_1^x}{\Delta x} + \frac{0 - 0}{\Delta y} (m_{22}^p - m_{12}^p) - \\ &-\frac{j}{\omega} \frac{0 - J_2^x}{\Delta x} + \frac{0 - 0}{\Delta y} (m_{23}^p - m_{13}^p) - \frac{j}{\omega} \frac{0 - 0}{\Delta y} + \frac{0 - J_2^y}{\Delta y} (m_{24}^p - m_{14}^p) \end{aligned} \right]$$

The first term represents the contribution of all current cells to the current cell J_x^1 . The second term represents the contribution of charge cells; expression $-1/\Delta x$ denotes replacing the derivative $\partial \phi / \partial x$ by the central difference. Brackets contain three terms since there are four charge cells. E.g., the first term represents the contribution of the first charge cell to the center of the second cell minus the contribution of the first charge cell to the center of the first cell. i.e., the first index at m denotes the center of the observation charge cell, where the contribution is computed, and the second index denotes the source cell. Charge in each cell is evaluated using nodal current densities and exploiting continuity equation

$$\nabla \cdot J = -j\omega\sigma \tag{2.4.6}$$

where $J[A/m]$ and $\sigma [C/m^2]$ are current density and charge density.

Similarly, equations for E_2^x and E_1^y can be composed. Next, we turn our attention to feeding the structure.

Plane wave feeding is the simplest way of feeding. For each current cell, we substitute $E_x = -E_x^I$ and $E_y = -E_y^I$, where E_x^I and E_y^I represent components of incident wave. In our situation, the antenna is assumed to be fed by a voltage source, which is connected between the analyzed structure and an auxiliary microstrip (the incident wave is not considered). In detail, the situation from (Fig. 2.16) is depicted in (Fig. 2.18).

In the position, where 1 denoted the input port in (Fig. 2.16), a voltage source of a voltage V_S is considered. The source supplies the structure by the current I_S . One pole of the source is connected to the analyzed structure; the second pole is connected to the auxiliary microstrip (*stub*), which represents certain impedance Z_{stub} related to ground. Input impedance Z_1 is then given by the relation

$$Z_1 = Z_S - Z_{stub} \quad (2.4.7)$$

where the impedance Z_S is determined as the ratio of the voltage V_S (its value is eligible) and the current I_S . The impedance Z_{stub} is computed analytically as $-Z_0 \cotg(\beta l_{stub})$.

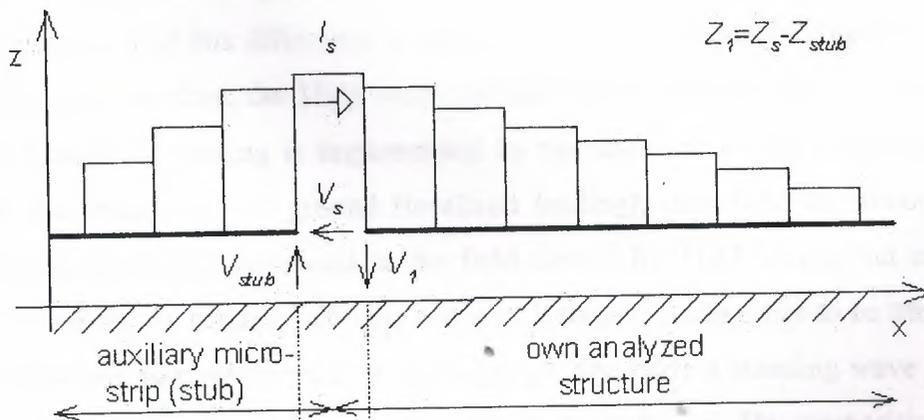


Figure 2.18 Approximation of current distribution along structure from figure 2.16 horizontal de-embedding.

If impedance is computed by the above-described way, the result does not correspond to the reality because influence of edge capacitances is not considered (Fig. 2.19).

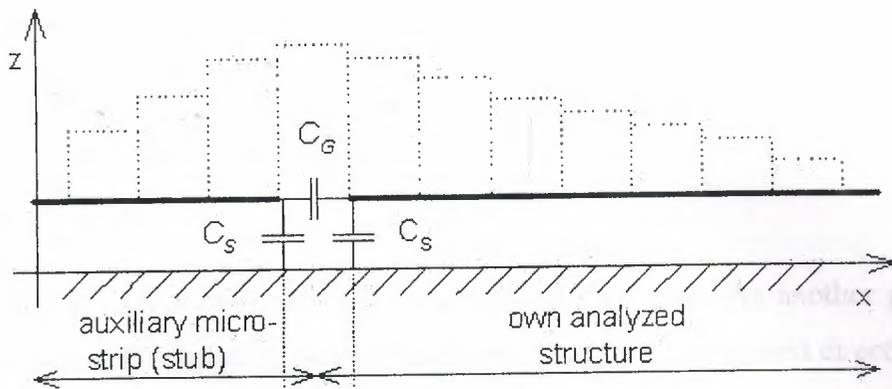


Figure 2.19 Edge capacitances.

Respecting capacitances, the impedance Z_{stub} has to be added in parallel to $Z_S = 1/j\omega C_S$ (C_S is not open-end capacitance, but it is smaller). That way, we get $Z_{stub}' = Z_{stub} \parallel Z_S$. Capacitance C_G is already included in the impedance Z_S and computing impedance $Z_1' = Z_S - Z_{stub}'$, we have to subtract Z_S . The final impedance is $Z_1'' = Z_1' Z_S / (Z_S - Z_1')$. Finally, we give a note about feeding the structure.

We can assume the structure to be fed by a voltage source connected between the microstrip and ground. This approach causes the direct computation of input impedance from the voltage and the current to provide results, which differ from the measured value. The reason of this difference is hidden in the fact that long cables are used for measuring, and therefore, the TEM wave (forward one, backward one) is present in the point of interest. If feeding is implemented by the above-described source connected between the microstrip and ground (localized feeding), then field distribution at the positive pole does not correspond to the field caused by TEM waves, but end effect influences results. In practical life, the result of localized feeding has to be transformed to the long-cable feeding (vertical de-embedding). The current standing wave has to be de-composed to the forward wave and the backward one. De-composition is not performed at the source port, but is shifted to the right. If both the waves are known, then reflection coefficient can be evaluated at the position of measurement. Then, the phase of reflection coefficient is shifted to the computation position.

The set of equations for unknown nodal values was built respecting physical principles. Now, the mathematical approach, which corresponds to the above-described notions, is given:

1. Current density is approximated in terms of basic functions

$$J_x(x, y) = \sum_{n=1}^{N_x} J_{x_n} B_{x_n}$$

$$J_y(x, y) = \sum_{n=1}^{N_y} J_{y_n} B_{y_n}$$

In our approach piecewise constant basis functions are used. As another possibility, rooftops can be used (linear in the first direction, constant in the second direction).

2. Integral expression of potentials based on Green functions is substituted to (2.4.1)

$$\left\{ -j\omega \iint_{\Omega} G_A(x-x', y-y') J(x', y') dx' dy' \right\} + \left\{ -\nabla \left[\iint_{\Omega} G_q(x-x', y-y') \rho(x', y') dx' dy' \right] \right\} + E^I = 0$$

Here, Ω denotes the metallic area and E^I is intensity of incident wave. In most cases, the structure is fed via microwave ports and E^I is zero. Analyzing selective surfaces e.g., plane-wave feeding is assumed.

3. All the integrals are efficiently evaluated and set of equations is composed. Solving it, unknown nodal values of current density are obtained.

On the basis of known current distribution, other antenna parameters can be computed (reflection coefficient at a given port, input impedance, etc.).

CHAPTER 3

ANALYSIS OF FREQUENCY SELECTIVE SURFACES

In this chapter, reflection coefficient and transmission coefficient are computed for an infinite periodic frequency selective surface. The attention is turned to the surface consisting of rectangular elements (metallic, slot). First, the surface is assumed to be fabricated from such dielectrics, which properties are identical with surrounding. At the end, a brief note on the analysis of real substrates is provided.

3.1 Numeric Analysis of Frequency Selective Surfaces

There are two approaches to the analysis of frequency selective surfaces. The first one is based on the method of induced electromotoric forces and enables to analyze surfaces of both the finite extent and the infinite one (in most cases, the method is applied to infinite structures).

Spectral domain moment method is the second way of analyzing frequency selective surfaces. We concentrate on this method only.

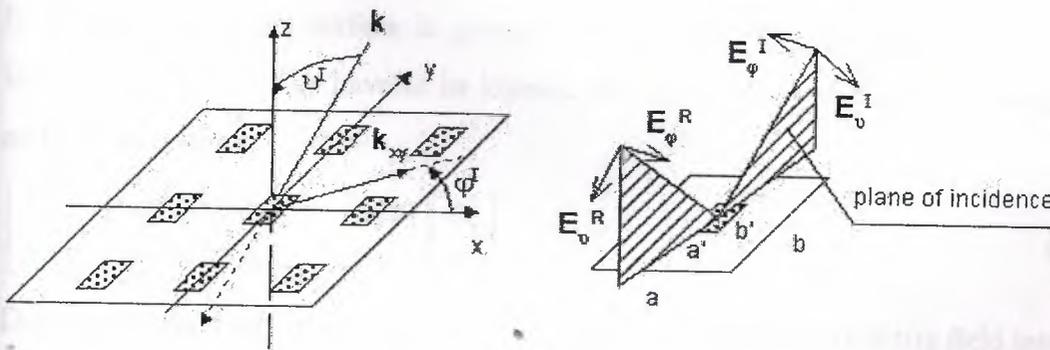


Figure 3.1 (a) Coordinate system for the analysis of a frequency selective surface, (b) an elementary cell of the dimensions a , b .

3.2 Metallic Elements

We assume a periodic surface as depicted in (Fig. 3.1.1) The surface is illuminated by a plane wave. Projection of wave vector k to the plane of the surface is denoted by k_{xy} . Electric field intensity of incident wave can be expressed then as follows:

$$E^I = E_0 \exp[+j(\alpha_0 x + \beta_0 y)]. \quad (3.2.1)$$

Here, α_0, β_0 are negative projections of wave vector of incident wave to the directions of coordinate axes x and y (the exponential term $\exp(+jkr)$ is assumed for increasing phase in the direction of propagation). Projections of wave vector in the described coordinate system are given as follows:

$$\alpha_0 = k \sin(\vartheta) \cos(\varphi)$$

$$\beta_0 = k \sin(\vartheta) \sin(\varphi)$$

The symbol E_0 denotes electric-field intensity vector in the origin of the coordinate system and k is free-space wave number.

In general, the incident wave is of both the parallel polarization and the perpendicular one. If field intensity of the incident wave is expressed in spherical coordinate system (Fig. 3.1), then components E_{ϑ}^I and E_{φ}^I represent directly the parallel component and the perpendicular one.

If frequency selective surface is going to be analyzed, components of electric field intensity in the plane xy have to be known. We can evaluate them re-computing E_{ϑ}^I and E_{φ}^I as follows:

$$\begin{bmatrix} E_x^I \\ E_y^I \end{bmatrix} = \begin{bmatrix} \sin(\vartheta) \cos(\varphi) & -\sin(\varphi) \\ \sin(\vartheta) \sin(\varphi) & \cos(\varphi) \end{bmatrix} \begin{bmatrix} E_{\vartheta}^I \\ E_{\varphi}^I \end{bmatrix} \quad (3.2.2)$$

Due to the periodicity of the surface, all the necessary quantities (electric field intensity, current density) in the plane xy can be expressed using Fourier series. Using Fourier expansion, density of electric current \mathbf{J} [A/m] on metallic element is expressed:

$$J = \sum_{m=-hd}^{hd} \sum_{n=-hd}^N J(\alpha_m, \beta_n) \exp[+j(\alpha_m x + \beta_n y)] \quad (3.2.3)$$

where

$$\alpha_m = \alpha_0 + \frac{2\pi}{a} m \quad \beta_n = \beta_0 + \frac{2\pi}{b} n \quad (3.2.4)$$

are spatial frequencies. Those spatial frequencies are a spatial analogy to the temporal angular frequency Ω , which are used for temporal analysis of signals. The only difference is hidden in the fact that time period T corresponds to spatial periods a and b . Coefficients of Fourier series can be obtained by integrating current densities over the whole surface of a cell (inverse Fourier transform)

$$J(\alpha_m, \beta_n) = \frac{1}{ab} \iint J(x, y) \exp[-j(\alpha_m x, \beta_n y)] \quad (3.2.5)$$

In analogy to current density, Fourier series can express arbitrary periodic quantity. We do that in the next part, where a plane wave excitation of a selective surface is expressed in the spectral domain (i.e., in the domain of spatial frequencies). Determination of reflective properties of the selective surface is the aim of the computation.

3.3 Problem Formulation

Assume a plane wave falling to a frequency selective surface under angles (ϑ, φ) . This wave excites currents in metallic elements, which form a scattered field E^S .

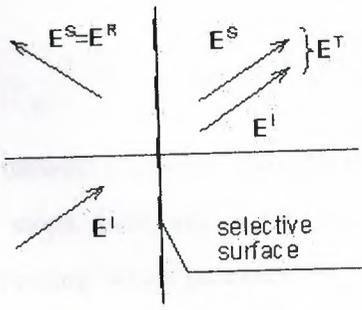


Figure 3.3.1 Frequency selective surface (reflected and transmitted wave)

Whereas the magnitude (not the phase) of the intensity of incident wave E^I is constant over the whole elementary cell, the magnitude of the intensity of the scattered wave E^S is different in the different points of the cell. Relation between E^S and E^I change according to the surface impedance in a given point. For the surface of conductive elements, we get

$$E^I + E^S = J/\gamma \quad (3.3.1)$$

Here, γ [$\text{S}\cdot\text{m}^{-1}$] is electric conductivity of elements. Assuming the electric conductivity of metallic elements to be perfect, the right-hand side of (3.3.1) approaches zero, and we get

$$E^I + E^S = 0 \quad (3.3.2)$$

Intensity of scattered electric field E^S in a point determined by the position vector r can be evaluated from vector potential A

$$E^S(r) = -j\omega\mu \left\{ A + \frac{1}{k^2} \nabla[\nabla \cdot A(r)] \right\} \quad (3.3.3)$$

As a source of vector potential A , currents J induced by incident wave to metallic elements of selective surface can be considered. Contribution of current in r' to the potential in r is described by Green function G . For vector potential, we get

$$A(r) = \iint G(r | r') J(r') dS' \quad (3.3.4)$$

In free space, Green function G is of the form

$$G = \frac{\exp(-jk |r - r'|)}{4\pi |r - r'|} \quad (3.3.5)$$

Combining relations (3.3.4), (3.3.3) and (3.3.2), integral equation for space-domain electric field is obtained:

$$-j\omega\mu \left\{ \iint G(r | r') \cdot J(r') dS' + \frac{1}{k^2} \nabla[\nabla \cdot \iint G(r | r') \cdot J(r') dS'] \right\} = -E^I \quad (3.3.6)$$

Eqn. (3.3.6) is solved in the domain of spatial frequencies. Eqn. (3.3.6) is mapped to the spectral domain within two steps. First, eqn. (3.3.3) is mapped by its rewriting to the component form and by expressing vector potential by Fourier series. That way, we get

$$\begin{bmatrix} E_x^S(\alpha_m, \beta_n) \\ E_z^S(\alpha_m, \beta_n) \end{bmatrix} = -j\omega\mu \begin{bmatrix} 1 - \frac{\alpha_m^2}{k^2} & -\frac{\alpha_m \beta_n}{k^2} \\ -\frac{\alpha_m \beta_n}{k^2} & 1 - \frac{\beta_n^2}{k^2} \end{bmatrix} \begin{bmatrix} A_x(\alpha_m, \beta_n) \\ A_y(\alpha_m, \beta_n) \end{bmatrix} \quad (3.3.7)$$

In the second step, the relation is mapped to the domain of spatial frequencies (3.3.4)

$$A(\alpha_m, \beta_n) = G(\alpha_m, \beta_n) \cdot J(\alpha_m, \beta_n) \quad (3.3.8)$$

Evaluating Fourier transform of the function (3.3.5), we get Green function in spectral domain

$$G(\alpha_m, \beta_n) = \frac{-j}{2\sqrt{k^2 - \alpha_m^2 - \beta_n^2}} \quad (3.3.9)$$

where the square-root of negative imaginary part is assumed only.

Combining (3.2.8), (3.2.9) and (3.2.7), the final relation for electric field in spectral domain is obtained (3.2.10)

$$-\frac{1}{2\omega\epsilon} \sum_{m,n} \begin{bmatrix} \frac{k^2 - \alpha_m^2}{\sqrt{k^2 - \alpha_m^2 - \beta_n^2}} & \frac{-\alpha_m \beta_n}{\sqrt{k^2 - \alpha_m^2 - \beta_n^2}} \\ -\alpha_m \beta_n & \frac{k^2 - \alpha_m^2}{\sqrt{k^2 - \alpha_m^2 - \beta_n^2}} \end{bmatrix} \begin{bmatrix} J_x(\alpha_m, \beta_n) \\ J_y(\alpha_m, \beta_n) \end{bmatrix} \quad (3.2.10)$$

$$\exp[j(\alpha_m x + \beta_n y)] = - \begin{bmatrix} E_z^I(x, y) \\ E_y^I(x, y) \end{bmatrix}$$

3.4 Problem Solution

Solving the problem, an unknown current distribution in (3.2.10) is approximated by the linear combination of properly chosen basis functions and unknown approximation coefficients. Such formal approximation is substituted into the solved eqn. (3.2.10). Since the approximation of the solution does not meet the initial relation exactly, we respect this fact by introducing a residual function (residuum). Lower the residuum is, more accurate the approximation is. The residuum is minimized by Galerkin method (residuum is sequentially multiplied by as many basis functions as many unknown approximation coefficients is computed; that way, N linear algebraic equations for N unknown approximation coefficients is obtained).

There are two approaches for choosing basis functions. The first one exploits basis functions, which are non-zero over the whole analyzed area. Functions are elected to physically represent standing waves of a current on an element.

The second approach divides the analyzed region to sub-regions, where current is approximated in terms of basis functions, which functional value is non-zero over a given sub-region only. This approach is advantageous in simple analysis of selective surfaces consisting of arbitrarily shaped elements.

3.5 Harmonic Basis Functions

Assume a frequency selective surface consisting of rectangular perfectly conductive elements of the dimensions a' and b' . If current distribution over an element is approximated in terms of harmonic basis functions, we obtain for the mode (1, 1) for the parallel polarization such current distribution, which is depicted in (Fig. 3.5.1.a), and for the perpendicular polarization, the distribution from (Fig. 3.5.1.b) is obtained

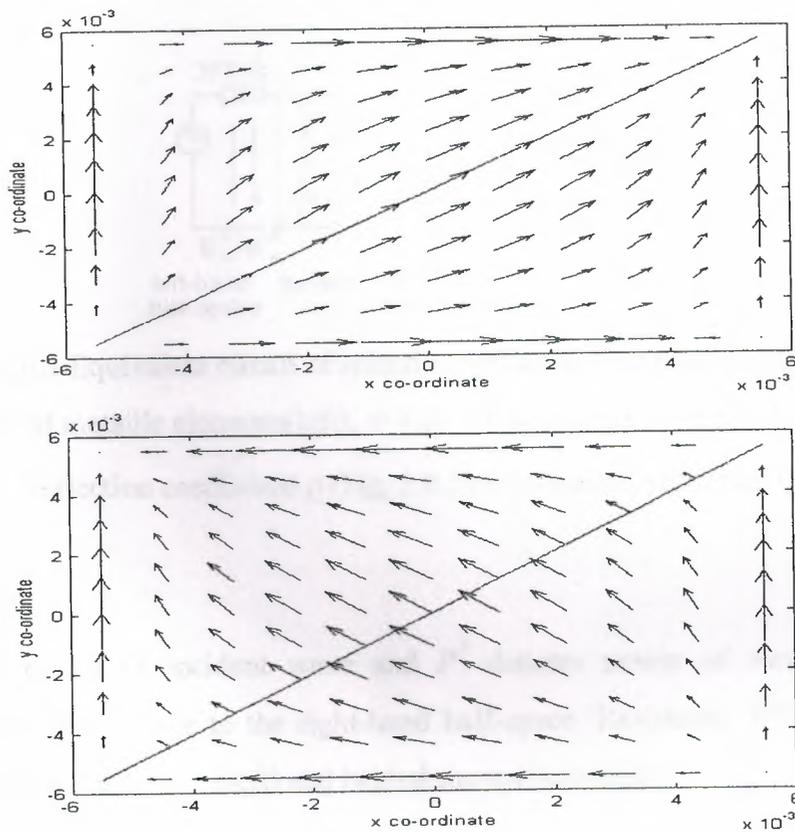


Figure 3.5.1 Directional field of current density of the mode (1,1) for parallel polarization (left) and for perpendicular one (right). The cell is of dimensions $a=b=12$ mm, the metallic element is of dimensions $a'=b'=12$ mm; $f=10$ GHz, $\theta=1^\circ$, $\varphi=45^\circ$, $d=1.5$ mm, $\epsilon_r=3.7$. Magnitude of marginal currents was increased with respect to reality. Plane of incidence is depicted in red.

3.6 Combination of Harmonic Functions and Chebyshev Polynomials

In contrast to purely harmonic basis functions, the function cosine is replaced by Chebyshev polynomial, which is of infinite magnitude at the edges of the element. That way, strong currents flowing along edges are correctly modeled. An equivalent circuit can describe behavior of frequency selective surfaces, which consist of metallic elements, where inductances of elements together with capacitances among ends of elements create a serial resonant circuit (Fig. 3.6.1).

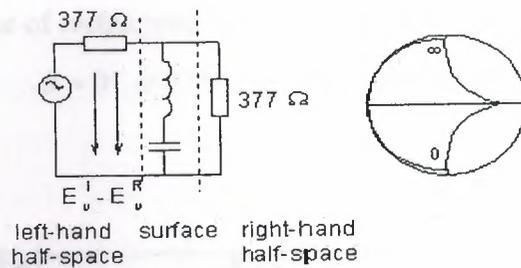


Figure 3.6.1 Equivalent circuit of selective surface in frequency domain: surface consists of metallic elements(left); $\nu = \varphi = 0$ directivity in smith chart (right)

Reflection coefficient ρ (Fig. 3.6.1) is evaluated according to.

$$P^T = P^I [1 - |\rho|^2]$$

Here, P^I is power of incident wave and P^T denotes power of wave, which was transmitted by the surface to the right-hand half-space. Resistance 377Ω models free space in front of the surface (left) and behind the surface (right).

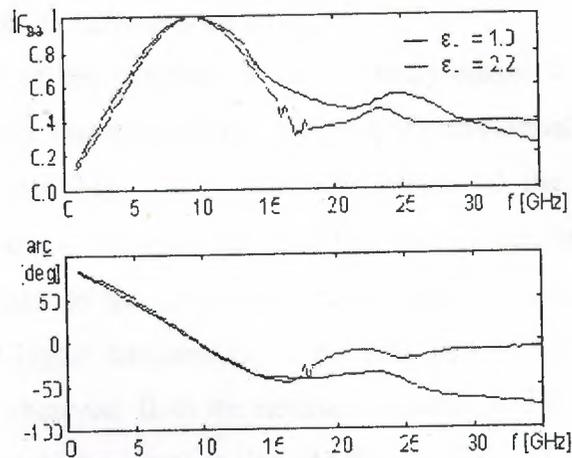


Figure 3.6.2 Surface of rectangular elements, tuned for 10 GHz, normal incidence
 $\vartheta = 0^\circ, \varphi = 90^\circ$, parallel polarization.

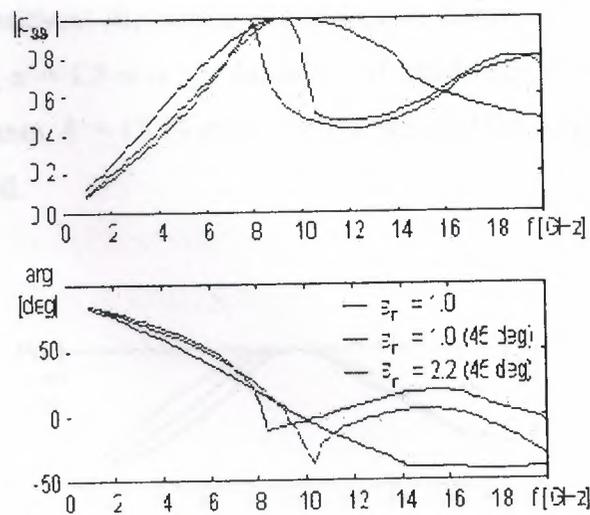


Figure 3.6.3 Surface of rectangular elements, tuned for 10 GHz, unnormal incidence
 $\vartheta = 45^\circ, \varphi = 90^\circ$, parallel polarization.

Frequency course of reflection coefficient of the surface for parallel polarization is depicted in Fig. 3.6.2 and 3.6.3. Considering the surface without dielectrics, ($b = 21$ mm, $a = 7.5$ mm, $b' = 19.75$ mm, $a' = 1.5$ mm), there is a break of the steepness of the reflection coefficient course around 14.3 GHz; this phenomenon is caused by the excitation of the first parasitic mode (grating lobe).

Adding dielectrics and modifying dimensions ($b = 17$ mm, $a = 7.5$ mm, $b' = 15.75$ mm, $a' = 1.5$ mm, height of the substrate $d = 1.57$ mm) causes a small decrease of the selectivity on one hand, but the stability of tuning for un-normal incidence is increased on the other hand (see Fig. 3.6.3). Further increase of the stability of reflection coefficient with respect to the angle of incidence can be reached by adding an upper dielectric layer. Thanks to the dielectrics, cell dimensions are smaller and parasitic modes are excited at higher frequencies (17.6 GHz). In Fig. 3.6.3, no significant effect of this mode can be observed. Both the normal incidence and the un-normal one excite parasitic resonance at higher frequencies. At frequencies, which are higher than the frequency of the first parasitic mode, the course of reflection coefficient does not express the total intensity of the reflected wave, but only the intensity of the basic mode (0,0). Considering perpendicular polarization of incident wave (in contrast to parallel one), reflection coefficient is nearly independent on the angle of incidence (see Fig. 3.6.4). If a surface without dielectrics is considered, dimensions $b = 21$ mm, $a = 7.5$ mm, $b' = 19.75$ mm, $a' = 1.5$ mm, are assumed., If dielectrics is considered, dimensions $b = 17$ mm, $a = 7.5$ mm, $b' = 15.75$ mm, $a' = 1.5$ mm and the height of the substrate $d = 1.57$ mm are assumed.

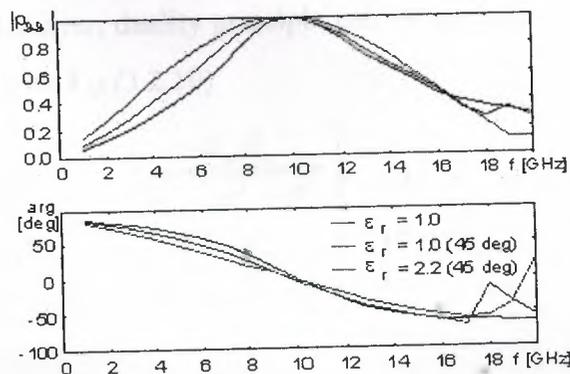


Figure 3.6.4 Surface of rectangular elements, tuned for 10 GHz, unnormal incidence $\vartheta = 45^\circ, \varphi = 0^\circ$, perpendicular polarization.

3.7 Slot Elements

In contrast to the surfaces of metallic elements, frequency selective surfaces of slots exhibit an *opposite* dependency of reflection coefficient, i.e. they behave as a band-stop filter. If no dielectrics and a single metallic layer are assumed, then slot surfaces are true complements of metallic-element surfaces.

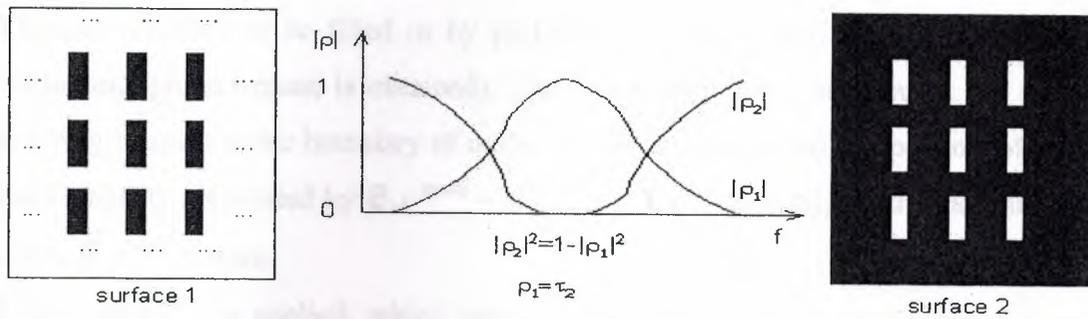


Figure 3.7.1 Complementary of slot surfaces and metallic-element ones. Reflection coefficient of metallic-element surface (1) equals to transmission coefficient of slot surface (2)

Now, we turn our attention to the derivation of equations for magnetic intensity over the aperture of slot surface in spectral domain. Considering the case without dielectrics and with a single metallic layer, duality principle can be applied: in (3.2.9), we exchange \mathbf{E} with \mathbf{H} , ϵ with μ , \mathbf{J} with \mathbf{J}_M (3.2.10)

$$-\frac{1}{2\omega\mu} \sum_{m,n} \begin{bmatrix} \frac{k^2 - \alpha_m^2}{\sqrt{k^2 - \alpha_m^2 - \beta_n^2}} & \frac{-\alpha_m \beta_n}{\sqrt{k^2 - \alpha_m^2 - \beta_n^2}} \\ -\alpha_m \beta_n & \frac{k^2 - \alpha_m^2}{\sqrt{k^2 - \alpha_m^2 - \beta_n^2}} \end{bmatrix} \begin{bmatrix} 2J_x^{h1+}(\alpha_m, \beta_n) \\ 2J_y^{h1+}(\alpha_m, \beta_n) \end{bmatrix} \exp[j(\alpha_m x + \beta_n y)]$$

$$= - \begin{bmatrix} 2H_x^i(x, y) \\ 2H_y^i(x, y) \end{bmatrix}$$

Comparing (3.2.10) and (3.2.9), a certain inconsistency can be revealed: coefficient 2 appeared at magnetic current density and exciting magnetic intensity. Next to the duality principle, the following fact had to be considered:

Field radiated by electric currents of the metallic-element surface is not scattered by those elements, whereas magnetic currents in slots produce a field, which is scattered by the conductive plane. In order to include the influence of the surrounding of the conductive plane, following steps have to be done:

$\mathbf{J}_M^+ = \mathbf{z} \times \mathbf{E}_{ap}$, where \mathbf{E}_{ap} is electric field intensity in the aperture (similarly to electric current, even the magnetic conductor of a real thickness is flown by the current from one side only - from left one).

The slot assumed to be filled in by perfect electric conductor (that way, an infinite continuous ground plane is obtained). This assumption does not change the situation any way because at the boundary of metal and air, the tangential components of electric field intensity are related by $\mathbf{E}_t^{metal} - \mathbf{E}_t^{air} = \mathbf{J}_M^+ \times (-\mathbf{z})$, and field intensity in metal \mathbf{E}_t^{metal} is zero.

Mirror principle is applied, which converts the problem to free-space one (Fig. 3.7.2 right).

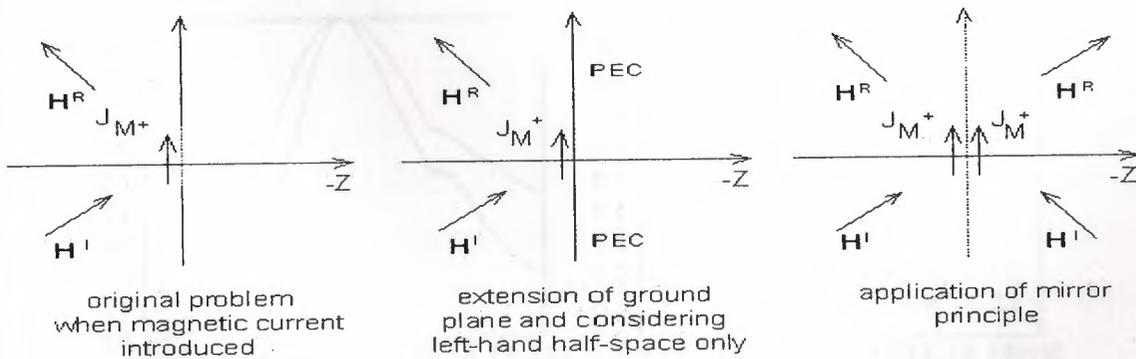


Figure 3.7.2 Explanation of equivalence principle.

In analogy to metallic elements, dual formulation of transmission coefficient is possible τ . In (Fig. 3.8.1), an example for normal incidence of a wave to a frequency selective surface with slots in vacuum ($b = 21$ mm, $a = 7.5$ mm, $b' = 19.75$ mm, $a' = 1.5$ mm) and on dielectric substrate ($b = 14$ mm, $a = 7.5$ mm, $b' = 13.25$ mm, $a' = 1.5$ mm and the height of the substrate $d = 1.57$ mm). are given. Obviously, selectivity of transmission coefficient is better for dielectrics (in case of metallic-element surface, the situation was

opposite). Next, influence of dielectrics is different compared to metallic-element surface, and therefore, slots are of different dimensions compared to metallic-element surface with dielectrics. Dielectrics even cause non-zero value of reflection coefficient on resonant frequency. Even here, parasitic modes appear.

3.8 Conclusion

The method of the analysis of frequency selective surfaces, given in this chapter, does not include the influence of dielectrics, but given examples do that. Real applications use several sequentially lined selective surfaces. That way, selectivity can be significantly improved. Design of such surfaces is rather complicated. Frequency selective surfaces can be modeled in commercial programs based on finite elements (HFSS) or on moment method (IE3D). In both cases, only a single element is analyzed thanks to the periodicity. Designing frequency selective surfaces, physical basis of their behavior and limitations have to be known.

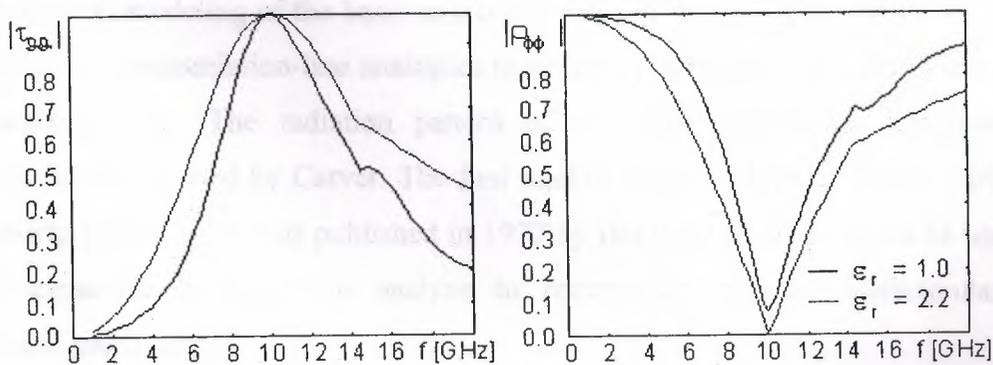


Figure 3.8.1 Surface of rectangular slots, tuned for 10 GHz, normal incidence $\theta = 0^\circ, \varphi = 90^\circ$, perpendicular polarization, transmission (left), reflection (right).

CHAPTER 4

DESIGN OF MICROSTRIP ANTENNA FOR (WLAN)

4.1 Abstract

This study presents theoretical investigation and simulation of Input Impedance behavior of nearly square single fed aperture coupled microstrip patch antenna that can satisfy the narrowband WLAN applications with 2.4 GHz band with 80 MHz bandwidth. For maximum coupling the patch should be centered over the slot, moving the patch relative to the slot in the H-plane direction has little effect, while moving the patch relative to the slot in the E-plane (resonant) direction will decrease the coupling level, also for maximum coupling, the feed line should be positioned at right angles to the center of the slot. Skewing the feed line from the slot will reduce the coupling, as will positioning the feed line towards the edge of the slot. Position of the feed line in different places relative to patch also was presented.

4.2 Introduction

Mathematical modeling of the basic microstrip radiator was initially carried out by the application of transmission-line analogies to simple rectangular patch fed at the center of radiating wall. The radiation pattern of a circular patch was analyzed and measurements reported by Carver. The first mathematical analyze of a wide variety of microstrip patch shapes was published in 1977 by Bahl and Bhartia, which he used the modal-expansion technique to analyze to rectangular, circular, semicircular and triangular patch shapes.

High speed, broadband and high capacity in or outdoor wireless local area networks (WLAN) are becoming more and more predominant today, its interesting to become familiar with some of the aspects of wireless design that must be faced and overcome. The advantages of microstrip antennas have made them a perfect candidate for use in the wireless local area network (WLAN) applications. Though bound by certain disadvantages, microstrip patch antennas can be tailored so they can be used in the new high-speed broadband WLAN systems and other applications, e.g. PCS, Blue tooth,

RFID, etc. Foundations for microstrip design: A microstrip patch antenna is a radiating patch on one side of a dielectric substrate, which has a ground plane on the underside. The EM waves fringe off the top patch into the substrate, reflecting off the ground plane and radiates out into the air. Radiation occurs mostly due to the fringing field between the patch and ground (Fig. 4.1). The radiation efficiency of the patch antenna depends largely on the permittivity of the dielectric. Ideally, a thick dielectric, low, and low insertion loss is preferred for broadband purposes and increased efficiency. The advantages of microstrip antenna that they are low-cost, conformable, lightweight and low profile, while both linear and circular polarization easily achieved. These attributes are desirable when considering antennas for WLAN systems. Disadvantages of microstrip antenna include such as a narrow bandwidth, a low gain (~6 dB) and polarization purity is hard to achieve.

4.3 Polarization Types

Usually the polarization of the wave radiated by the antenna in a particular direction. This is usually dependant on the feeding technique. When the direction is not specified, it is in the direction of maximum radiation. Two most widely polarizations (linear and circular) are shown in Fig. 4.2 and 4.3.

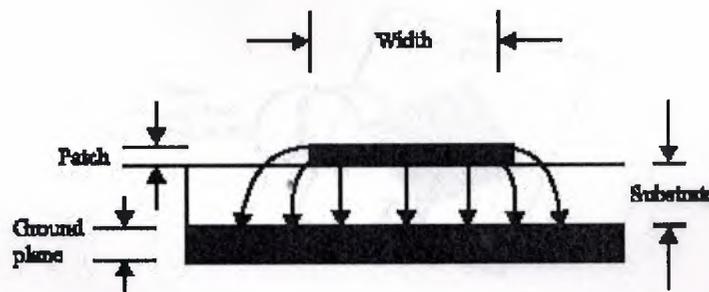


Figure 4.1 Operation of microstrip patch.

4.4 Linear Polarization

A slot antenna is the counter part and the simplest form of a linearly polarized antenna. On a slot antenna the E field is orientated perpendicular to its length dimension. The

usual microstrip patches are just different variations of the slot antenna and all radiate due to linear polarization. Figure 4.2 illustrates the operations of a linearly polarized wave radiating perpendicular to the patch plane.

Circular polarization: Circular polarization (CP) is usually a result of orthogonal fed signal input. When two signals of equal amplitude have 90° phases, the resulting wave is circularly polarized. Circular polarization can result; left hand circularly polarized (LHCP) where the wave with anticlockwise, or right hand circularly polarized (RHCP) wave with clockwise rotation. The main advantage of CP is that regardless of receiver orientation, it will be always able receiving a component of the signal. This is due to the resulting wave having an angular variation.

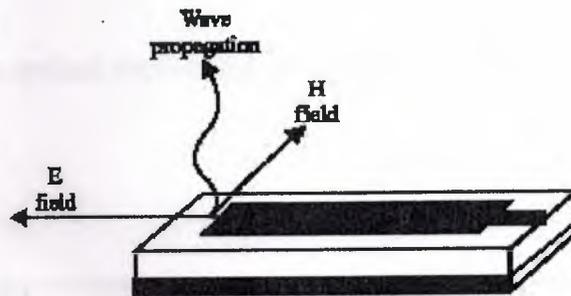


Figure 4.2 Microstrip antenna patch with linear polarization wave.

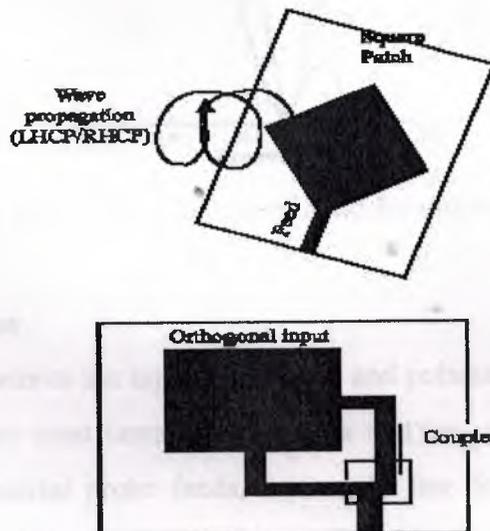


Figure 4.3 Two type of microstrip patch with circular polarization wave.

4.5 Bandwidth

The bandwidth of the patch is defined as the frequency range over which it is matched with that feed line within specified limits. In other words, the frequency range over which the antenna will perform satisfactorily. This means the channels have larger usable frequency range and thus results in increased transmission. The bandwidth of an antenna is usually defined by the acceptable standing wave ratio (SWR) value over the concerned frequency range.

$$BW = \frac{SWR - 1}{Q\sqrt{SWR}} \quad (4.5.1)$$

Where, Q is a quality factor.

The Fig. 4.4 shows a typical narrow and broadband phenomenon in terms of frequency band usage.

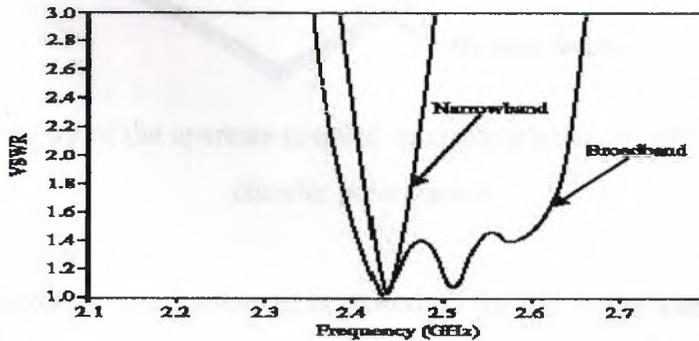


Figure 4.4 Narrowband and broadband for microstrip antenna.

4.6 Feeding Techniques

Feeding technique influences the input impedance and polarization characteristics of the antenna. There are three most common structures that are used to feed planar printed antennas. These are coaxial probe feeds, microstrip line feeds and aperture coupled feeds. For aperture-coupled a feed, which is used in this study the microstrip feed line, is etched on the bottom of the feed substrate.

4.7 Aperture Coupled Microstrip Antennas

In an aperture-coupled feed, which is another type of electro magnetically coupled (EMC) feed; the RF energy from the feed line is coupled to the radiating element through a common aperture in the form of a rectangular slot. This type of feed was first proposed by Pozar. The aperture coupled feeding mechanism is shown in Fig. 4.5. Figure 4.5 shows the geometry of the basic aperture coupled patch antenna that is designed in this work and considered in simulation parameters.

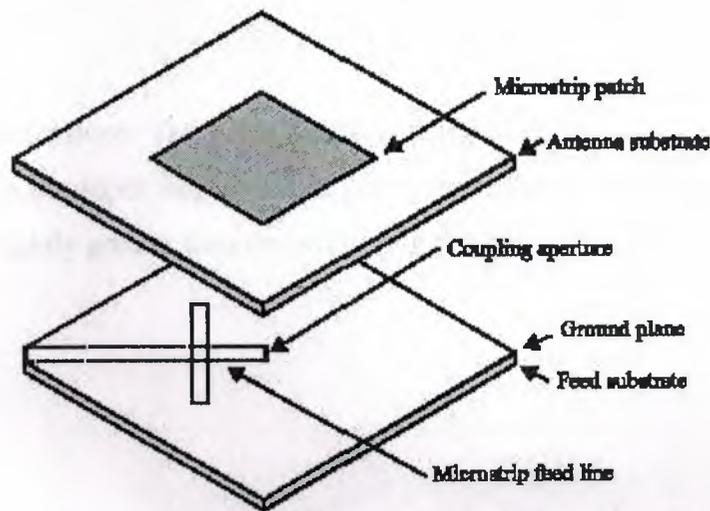


Figure 4.5 Geometry of the aperture coupled microstrip antenna with corner feeding of circular polarization.

The radiating microstrip patch element is etched on the top of the antenna substrate and the microstrip feed line here is etched of the corner of bottom of the feed substrate.

Rectangular slot 45° is also used the nearly square ground plane.

Designs and simulation procedure: The aim of this design is to provide a good performance in input impedance resonant frequency matching and perfect circular polarization for WLAN requirements.

4.8 Substrate Thickness

Substrate thickness should be chosen as large as possible to maximize bandwidth and efficiency, but not so large as to risk surface-wave excitation. For maximum operating frequency f the substrate thickness (h) should satisfy:

$$h \leq \frac{0.3c}{2 \cdot f_u \sqrt{\epsilon_r}} \quad (4.6.1)$$

Where, ϵ_r is the relative dielectric constant obtained using (r) resonant line method and (c) is the light speed.

Antenna patch dimensions: The patch length is determined by condition for resonance. This occurs when the input impedance is purely real. For nearly square geometry the length (l) must slightly greater than the width (w), to obtain an initial value of 1

$$L = \frac{c}{2f_r \sqrt{\epsilon_r}} \quad (4.7.1)$$

$f = 2.45$ GHz. Using equation (4.7.1) we can calculate the effective relative permittivity (ϵ_{eff}) as follows:

$$\epsilon_{eff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left(1 + 12 \frac{h}{w} \right)^{-0.5} \quad (4.7.2)$$

To calculate patch width the following equation used.

$$w = \frac{c}{2f_r} \left(\frac{\epsilon_r + 1}{2} \right)^{0.5} \quad (4.7.3)$$

For improved value of l we can find the fringe factor dL .

$$dL = 0.412h \frac{(\epsilon_{eff} + 0.3)(w/h + 0.264)}{(\epsilon_{eff} + 0.258)(w/h + 0.8)} \quad (4.7.4)$$

The foregoing procedure may be repeated with a result that improves l .

4.9 Excitation

Figure 4.5 shows that it is possible to excite the two modes (TM_{10}, TM_{01}) using one feed by introducing a small perturbation in the patch and it radiates Right-Hand Circular Polarization (RHCP). For (LHCP) feed is placed on other diagonal. In this simulation different location for the feeder were used.

4.10 Results and Discussion

Analytical results are obtained using equations with the help of FORTRAN simulation software and compared with for verification.

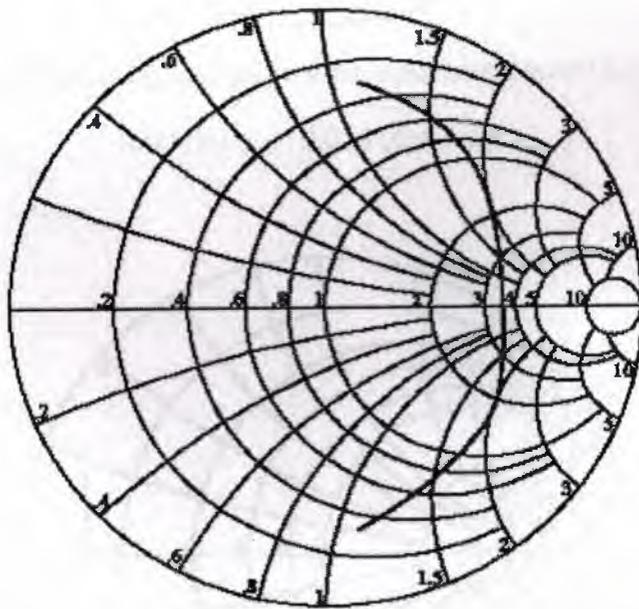


Figure 4.6 Input impedance for microstrip rectangular patch.

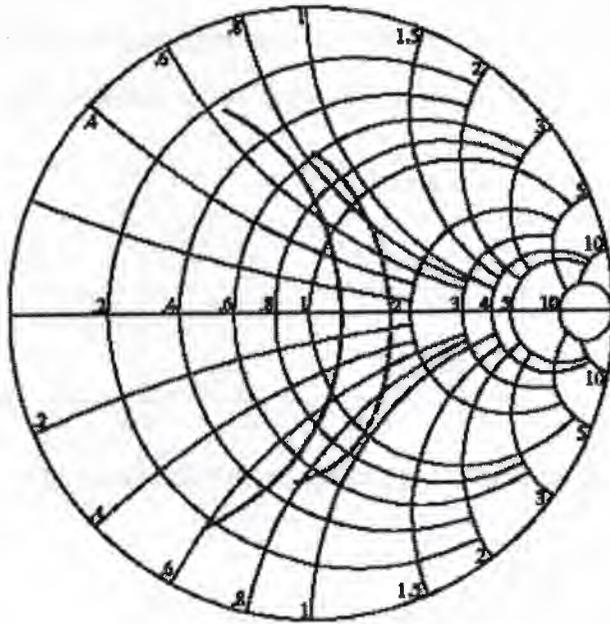


Figure 4.7 Trade-off inseting feed point with input impedance (A) $y_0 = 0.079$ (solid line) (B) $y_0 = 0.026$ (dot line).

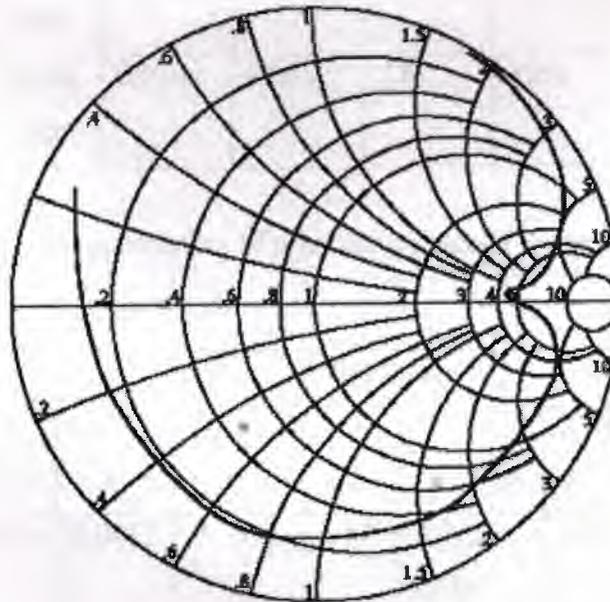


Figure 4.8 The input impedance for corner fed microstrip patch.

Smith chart shown in Fig. 4.6, show the input impedance for microstrip rectangular patch with these following parameters and BW which it calculates from equation (4.5.1) as a relative bandwidth as shown in the Table 4.1, which it same as appeared in Brown.

Dimension					Excitation			
w	l	h	y_0	x_0	\bullet_r	f_t	SWR	BW
0.495	0.330	0.0064	0	0.013	2.2	1.57 GHz	3.6	16%

Table 4.1 The parameters of microstrip antenna rectangular patch.

	Excitation			
	y_0	x_0	SWR	BW
Solid line	0.079	0.013	2.0	8%
Dots line	0.026	0.013	2.6	12%

Table 4.2 The parameters of microstrip antenna rectangular type.

Dimension					Excitation			
w	l	h	y_0	x_0	\bullet_r	f_t	SWR	BW
0.224	0.217	0.0083	0	0	2.62	2.42 GHz	4.5	18.7%

Table 4.3 The parameters of center feed microstrip antenna rectangular patch.

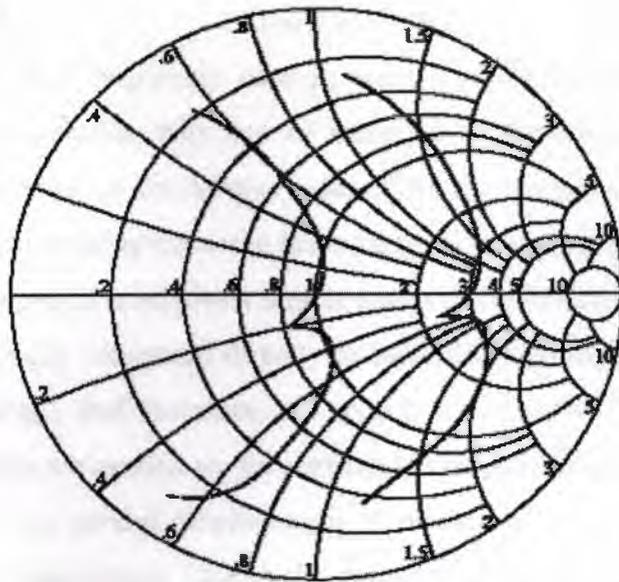


Figure 4.9 The input impedance for LHCP Microstrip antenna nearly square patch
 (a) $x_0=0.11$, $y_0=0.11$ (dot line), (b) $x_0=0.068$, $y_0=0.068$ (solid line).

Also Fig. 4.7 shows the input impedance with increased feed points at $y = 0.0798$ and $y = 0.0268$ has been studied. All other parameters same as in Fig. 4.6 and Table 4.2 describes these. For a simple circular polarization for single feed microstrip antenna with an aperture coupled has been demonstrated. Properly adjusting the length of the patch and using inset feed line can easily obtain the right-hand or left-hand. Smith chart shown in Fig. 4.8 the input impedance for microstrip rectangular patch corner feed with these following parameters and BW, as shown in the Table 4.3.

As shown in Table 4.3 the microstrip antenna corner feed give center frequency 2.425 GHz, which it is approximately equal to unlicensed ISM WLAN IEEE802.11b or Wi-Fi band. But with narrower bandwidth than it, with carefully select dielectric constant better bandwidth can be obtained. Figure 4.9 shows the variation of the feed line positions to give optimum state circular polarization for two modes (TM_{10}, TM_{01}) for equal amplitude and 90° phase differences, but with cost pay for the antenna bandwidth.

CONCLUSION

Microstrip antennas are frequently used in today's wireless communication systems. Thanks to their low profile, they can be mounted to the walls of buildings, to the fuselages of airplanes or to the reverse sides of mobile phones. Moreover, microstrip antennas are fabricated using the same technology as producing printed circuit boards. Therefore, the fabrication is relatively simple and well reproducible. Finally, microstrip antennas can be simply integrated directly to microwave circuits, which are based on microstrip technology, and therefore, no special transmission lines, symmetrization circuits or connectors are needed on the contrary to classical antennas.

Nevertheless, there are several disadvantages of microstrip antennas. Narrow operation band is the main disadvantage. Due to this property, the design of microstrip antenna arrays exhibiting sufficiently low level of side-lobes is a really hard nut to be broken. Even the parasitic radiation of the feeding microstrip network, which can deform the directivity pattern. One of the most frequently used types of microstrip antennas, the patch antenna. The antenna consists of a conductive rectangle of the dimensions $A \times B$, which is etched on a dielectric substrate. The antenna is fed by the microstrip transmission, the microstrip goes from the front edge of the substrate crossways from the left. The second side of the substrate is continuously electroplated. The electroplated side plays the role of a reflector in the sense of zero potential (from the point of view of feeding) and in the sense of limiting radiation in the direction behind the reflector. Further, we call the electroplated side the ground plane. The microstrip antenna that is fed by the microstrip transmission line can be considered as an open (non-shielded) open-ended transmission line, which is significantly widened at its end. If electromagnetic wave propagates along such transmission line, electromagnetic energy is primarily radiated into surrounding at the non-homogeneities (spontaneous widening of the microstrip at the border of the feeding line and the antenna element and the open end of this element) of the transmission line. The structure therefore behaves as a transmitting antenna. Moreover, if the length of the microstrip antenna element equals to the half of the wavelength on this widened transmission line, then input impedance of such an antenna is purely real. Then, the antenna is said to be in resonance.

Radiation of a microstrip antenna can be explained in different ways. We can come out of the current distribution on the antenna element, which can be consequently understood as a wire antenna consisting of a very wide and a very thin antenna conductor. Next, we can come out from the line of electric intensity at the front side and at the back side of the antenna element (from the point of view of the feeding transmission line) and we can explain the radiation as an effect of a strong horizontal component (i.e. of the component, which is oriented in parallel with the ground plane) of electric field intensity vector at those edges.

REFERENCES

- [1] Assoc. Prof. Dr. SAMERR IKHADIR, Lecture notes
- [2] ZBYNK RAIDA, "Electromagnetic and Microwaves Techniques"
<http://www.urel.feec.vutbr.cz>
- [3] Carver, K.R. and J.W. Wink, 1981. Microstrip Antenna Technology, IEEE Trans. Antennas and Propagation.
- [4] Kumar, G. and K.P. Ray, 2003. Broadband Microstrip Antenna: A Handbook. Artech House.
- [5] Wikipedia, free encyclopedia"<http://en.wikipedia.org>".