FUNDAMENTAL LIMITS ON ANTENNA SIZE FOR FREQUENCY AND TIME DOMAIN APPLICATIONS

TAE-YOUNG YANG

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

William A. Davis, Co-Chair
Warren L. Stutzman, Co-Chair
Steven W. Ellingson
Jeffrey H. Reed
Christopher A. Beattie

SEPTEMBER 3, 2012
BLACKSBURG, VIRGINIA

KEYWORDS: Antenna Radiation Physics, Antenna Transfer Function, Fundamental-Limit Theory on Antenna, Near-Field Interaction, Ultra-Wideband Antenna

FUNDAMENTAL LIMITS ON ANTENNA SIZE FOR FREQUENCY AND TIME DOMAIN APPLICATIONS

TAE-YOUNG YANG

Abstract

As ubiquitous wireless communication becomes part of life, the demand on antenna miniaturization and interference reduction becomes more extreme. However, antenna size and performance are limited by radiation physics, not technology.

In order to understand antenna radiation and energy storage mechanisms, classical and alternative viewpoints of radiation are discussed. Unlike the common sense of classical antenna radiation, it is shown that the entire antenna fields contribute to both radiation and energy storage with varying total energy velocity during the radiation process. These observations were obtained through investigating impedance, power, the Poynting vector, and energy velocity of a radiating antenna.

Antenna transfer functions were investigated to understand the real-world challenges in antenna design and overall performance. An extended model, using both the singularity expansion method and spherical mode decomposition, is introduced to analyze the characteristics of various antenna types including resonant, frequency-independent, and ultra-wideband antennas. It is shown that the extended model is useful to understand real-world antennas.

Observations from antenna radiation physics and transfer function modeling lead to both corrections and extension of the classical fundamental-limit theory on antenna size. Both field and circuit viewpoints of the corrected limit theory are presented. The corrected theory is extended for multi-mode excitation cases and also for ultra-wideband and frequency-independent antennas.

Further investigation on the fundamental-limit theory provides new innovations, including a low-Q antenna design approach that reduces antenna interference issues and a generalized approach for designing an antenna close to the theoretical-size limit. Design examples applying these new approaches with simulations and measurements are presented.

The extended limit theory and developed antenna design approaches will find many applications to optimize compact antenna solutions with reduced near-field interactions.

Acknowledgments

I would first like to thank Dr. William A. Davis and Dr. Warren L. Stutzman for serving as my co-advisors during both my M.S. and Ph.D. degrees. Their generous advice and encouragement through constructive discussions have been invaluable. In fact, the presented work in this dissertation was achieved through joint team effort of Dr. William A. Davis, Dr. Warren L. Stutzman, and me. In particular, it is Dr. Davis' original idea correcting the classical fundamental-limit theory with the energy velocity approach.

I would also like to thank Dr. Steven W. Ellingson, Dr. Jeffrey H. Reed, and Dr. Christopher A. Beattie for serving on my committee. Thanks are extended to Dr. Ahmad Safaai-Jazi, Dr. Ioannis M. Besieris, Dr. Gary S. Brown, Dr. Dong S. Ha, Dr. Carl B. Dietrich, and Dr. Majid Manteghi for their persistent help with my research and professional development.

I would like to express my gratitude to Sony Ericsson USA for funding the hearing-aid compatibility (HAC) research. Especially, Dr. Minh-Chau Huynh at Sony Ericsson USA provided great help with performing the near-field scan and specific-absorption-rate (SAR) measurements for mock cell phones.

Special recognition is due all the previous and current members of Virginia Tech Antenna Group (VTAG) that I have worked with. The opportunity to interact with such a group of talented and diverse individuals has been enlightening. In addition, I am indebted to Randall Nealy for the great advice and assistance in antenna construction and measurement.

Finally, and most importantly, I recognized my parents, younger brother and his wife, parents-in-law, brother-in-law and wife for their never ending patience, love and support during my doctoral journey.

FUNDAMENTAL LIMITS ON ANTENNA SIZE FOR FREQUENCY AND TIME DOMAIN APPLICATIONS

TAE-YOUNG YANG

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The demand on antenna miniaturization and interference reduction has substantially increased in recent years. However, antenna size and performance are limited by radiation physics, not technology. Not all of the energy around an antenna contributes to far-field radiation [*Chu*, 1948; *R Collin and Rothschild*, 1964; *Davis et al.*, 2011; *Grimes and Grimes*, 1999; *McLean*, 1996]. Physics requires a minimum amount of non-radiating energy in the antenna radiation process, affecting antenna operational bandwidth and radiation efficiency. The amount of non-radiating energy is also responsible for the interference issues. On the other hand, some wireless applications can have an extended near-field communication range if the non-radiating energy is maximized.

The fundamental-limit theory for antennas shows that size, efficiency, and bandwidth are trade-offs in the design process. The theoretical minimum radiation Q curve versus ka enables the designer to evaluate the feasibility of realizing specifications and determine how close the designed antenna is to the limit. Unfortunately, over the last half century, previous research activities on the fundamental limit theory and related research have focused on the accuracy of the fundamental-limit theory itself rather than discussing more why conventional antennas are not close to the theoretical limit and how we can make the performance of the antenna close to the limit. In addition, the trade-off relationship of the limit theory has often been focused on electrically small antennas. However, the basic radiation-Q formula is directly related to the non-

propagating near-field energy and it provides a clue to solve some near-field problems. This has been overlooked for a long time.

This dissertation provides integrated knowledge and understandings on antenna radiation physics and fundamental limits on antenna size and performance. A couple of antenna design examples are provided to demonstrate the advantage of the integrated knowledge and understandings on fundamental antenna matters.

Chapter 2 provides fundamental physical insights about how energies get decoupled from an antenna and contribute to far-field radiation through multiple virtual boundaries, including the radian sphere and causal surfaces. The fundamental physical insights are useful to understand the fundamental limits on antenna size and near-field interactions issues. We review the classical antenna radiation models, discuss what has been assumed and missed in the models, and provide alternative viewpoints in terms of wave impedance, power, Poynting vector, and energy velocity of an ideal dipole antenna. The radiation and energy storage mechanism in the near-field region is important for understanding the fundamental limits on antenna size and the near-field interactions issues.

Chapter 3 introduces an antenna transfer function model as a tool to understand correlation between antenna physical parameters and performance. The singularity-expansion method (SEM) and spherical-wave expansions (SWE) are used to model the antenna transfer function. A previous SEM-based model using antenna effective height is reviewed. An alternative, decomposed SEM-based model using a network description of excited spherical modes is introduced. The transfer functions of resonant, frequency-independent, and ultra-wideband antennas are investigated to understand the correlation between antenna physical structure and performance parameters.

Chapter 4 investigates the classic fundamental-limit theory on antenna size and performance. The fundamental-limit theory on antenna size and performance not only provides a figure of merit to quantify the amount of non-radiating energy created from an antenna, but also suggests some clues to address the near-field interference and communication range issues. Assumptions and errors in the classical fundamental-limit are explained, based on the observation of varying energy velocity in Chapter 2. The classic limit theory is corrected and extended for antennas with

an arbitrary polarization. The antenna model developed in Chapter 3 is used to consider the theoretical size limit of frequency-independent and ultra-wideband antennas.

Chapter 5 demonstrates the low-Q antenna solution approach to mitigate near-field interaction problem between a cellular phone and a hearing aid without sacrificing antenna far-field radiation performance. Interaction issues between the cellular phone and the human head are also addressed with the same low-Q antenna solution. Analytic analysis, simulation, and measurement results are provided to show the advantage of the low-Q antenna approach, compared to high-Q antennas.

Chapter 6 addresses challenges and trade-offs in a compact ultra-wideband antenna design. A general design strategy is developed, based on observations from the fundamental-limit theory. An example antenna design for ultra-wideband and software-defined radio applications is introduced. Performance of the designed antennas is characterized through simulations and measurements.

Finally, Chapter 7 presents an overall summary and conclusions. The original contributions of this dissertation are summarized, followed by a short discussion on the possibilities for future work.

Chapter 2

Antenna Radiation and Energy Storage Mechanism

In wireless communication systems, the antenna is often treated as a black box located at the end of a radio-frequency (RF) front-end module. However, to meet strict volume constraints, antennas must be treated as sophisticated devices using innovative design approaches as introduced in this chapter. Realized gain, gain and polarization bandwidth, and radiation directivity of antennas [*Balanis*, 2005; *R E Collin and Zucker*, 1969; *Elliott*, 2003; *Schelkunoff*, 1952; *Stutzman and Thiele*, 2012] are typical performance antenna characterization parameters. In this sense, how much transmitted power is delivered to a receiving system would be the primary interest in a system-level viewpoint. Thus, far-field aspects of antenna radiation are emphasized in many antenna textbooks and are appropriate for most practical uses of antennas. But, antenna radiation phenomenon and energy storage mechanism in the near-field region have received little attention. The radiation and energy storage mechanism in the near-field region is important for understanding the fundamental limits on antenna size and near-field interactions issues.

In this chapter, we review classical antenna radiation models, discuss what has been assumed and missed in the models, and provide alternative viewpoints.

2.1 Classical Viewpoints

The definition of *electromagnetic radiation*, in the institute of electrical and electronics engineers (IEEE) standard definitions of terms for antennas [*IEEE*, 1993] is "The emission of electromagnetic energy from a finite region in the form of unguided waves." For an antenna, the finite region is the effective antenna aperture, which is an interface between bounded and unbounded regions. Antennas transform input signals inside the bounded region to unbounded electromagnetic energy.

Often, the unbounded region is free space having the impedance of 376.7 Ω (η_0) and is defined as the ratio between electric and magnetic-field amplitudes [*Stutzman and Thiele*, 2012] of a transverse-electromagnetic-mode (TEM) wave. Considering the typical 50 Ω reference impedance (Z_0) in wireless communication systems, antennas need to provide an impedance transformation from the reference impedance to the free-space impedance in order to minimize the return loss due to the impedance mismatch. In this sense, the antenna functions as an impedance transformer. The radiation of the unbounded energy can be limited to a confined angular range, i.e. be directive. So, power density in the limited angular range becomes larger than that of the isotropic radiator with the same input power. In this sense, an antenna can be considered as a spatial amplifier. The figure of merit for directive radiation is antenna directivity (D). Ohmic loss in the antenna structure reduces the total radiated power, which is quantified with the antenna radiation efficiency (e_r). Antenna gain is the product of the directivity and the radiation efficiency. If return loss is included in the gain, it is referred as *realized gain*. In a wireless communication link, polarization mismatch between transmitting and receiving antennas is included using additional efficiencies. It is common in wireless communication to quantify operational bandwidth using 3-dB gain bandwidth. These antenna parameters characterize antenna performance for far-field wireless communications. In this sense, the antenna is seen as a black box that interfaces the wireless communication system to air. Far-field performance of the antenna has been emphasized in practice. As a result, near-field radiation phenomena, including energy storage mechanisms, have often been overlooked in both academia and industry. However, as we will see later in Chapter 4, the fundamental-limit theory on antenna size and performance is closely related to the antenna radiation process and energy storage mechanisms. Therefore, knowledge of the antenna radiation process and energy-storage

mechanisms in the near field can be used in antenna design for satisfying far-field performance specifications as well as understanding the size limit theory. In addition, such knowledge also explains the actual source of antenna radiation and what is limiting antenna performance.

In order to have a broad picture of an antenna radiation, consider electric fields from an electric dipole moment ($\vec{p} = q\vec{dl}$, where q is electric charge and $\vec{dl}$ is a time-dependant displacement vector between the charges) in a retarded time form [*Jefimenko*, 1966; *Panofsky and Phillips*, 2005; *H G Schantz*, 2001; *Glenn S. Smith*, 1997] in spherical coordinates is given by

$$\vec{E}(r,t) = \frac{1}{4\pi\varepsilon_0}\left[\left(\frac{p(r,t-\frac{r}{c})}{r^3} + \frac{\dot{p}(r,t-\frac{r}{c})}{cr^2}\right)\left(\hat{r}2\cos\theta + \hat{\theta}\sin\theta\right) + \hat{\theta}\sin\theta\frac{\ddot{p}(r,t-\frac{r}{c})}{c^2 r}\right] \quad (2.1)$$

where p is the magnitude of the dipole moment ($\vec{p}$), c is the speed of light, ε_0 is the permittivity of vacuum (=8.854 pF/m), $\dot{p}$ is a time derivative of p , and $\ddot{p}$ is a double time derivative of p . Typically, the first term with the $1/r^3$ factor is called the static field, the second term with $1/r^2$ factor is considered as an induction field, and the third term with $1/r$ factor is considered as a radiation field. Similarly, the magnetic field of the dipole moment is given as

$$\vec{H}(r,t) = \hat{\phi}\sin\theta\frac{1}{4\pi}\left(\frac{\dot{p}(r,t-\frac{r}{c})}{r^2} + \frac{\ddot{p}(r,t-\frac{r}{c})}{cr}\right) \quad (2.2)$$

The first and second terms are induction and radiation fields respectively.

The electrostatic field term with a $1/r^3$ term has its basis in the classical Coulomb's law. Thus, the reactive power created from the dipole moment is capacitive. Unlike other terms, the static field term is not associated with time derivatives. In a similar manner, a magnetic dipole moment has a magneto-static field with $1/r^3$ term based on Biot-Savart's law. The induction fields with $1/r^2$ terms originate from the Faraday's law and Ampere's law with Maxwell's correction. Basically, time-varying electric fields induce magnetic fields (Faraday's law) while time-varying magnetic fields induce electric fields (Ampere's law with Maxwell's correction). Thus, these induced field terms are associated with a time derivative, which appear in both (2.1)

and (2.2). The radiation fields with a $1/r$ term are related to a radial power density contributing to a far-field real power transmission (P_{rad}), i.e.

$$\begin{aligned} P_{rad} &= \lim_{r\to\infty}\left[\oint_S \vec{E}\times\vec{H}\cdot d\vec{s}\right] \\ &= \lim_{r\to\infty}\left[\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}(\vec{E}\times\vec{H})\cdot\hat{r}r^2\sin\theta d\theta d\phi\right] \end{aligned} \tag{2.3}$$

Thus, the fields with the $1/r$ term in both the electric and magnetic fields are classically considered as the only terms contributing to antenna far-field radiation, while the other field terms with $1/r^2$ and $1/r^3$ variation are considered as near-field terms, though such a conclusion for the $1/r^2$ terms has only been assumed and never justified. The radiation field terms are associated with the double time derivative of time-dependant charges from a mathematical viewpoint or with equivalent accelerated charges from a physical viewpoint. Therefore, it has been historically believed that accelerated charges are the source of the antenna radiation [*Balanis*, 2005; *R E Collin and Zucker*, 1969; *Elliott*, 2003; *Schelkunoff*, 1952; *Stutzman and Thiele*, 2012]. In addition, it has been known that stationary or moving charges with a constant speed do not contribute to radiation.

The brief review of dipole moment terms helps to identify the cause of antenna radiation (charge acceleration), but it does not clearly show why antenna radiation occurs and how the radiation fields are decoupled from an antenna. In order to understand the physical radiation mechanism between the accelerated charges and radiation fields, the classical single and dual charge (dipole) models are reviewed in the following sub sections.

2.1.1 The Single Charge Model

It appears that Heaviside first considered the single-charge radiation model in his classical electromagnetics book [*Heaviside*, 1894]. Due to the simplicity of the model, the single charge model has been popular for explaining the basic radiation mechanism. It begins by considering a positive charge moving with a constant speed (v_1) in z direction. The electric field lines of the charge are shown in Figure 6.1a for the instant of time when the charge is at the origin (O). Notice that the electric field lines only have a radial vector component. Then, the speed of the charge increases right after passing the origin. After a short time of period, the acceleration ends and the charge starts moving with a new constant speed (v_2) that is faster than the initial constant

speed (v_1) due to the applied acceleration. Because the propagating speed of the electric fields is limited to the speed of light (c), update of the electric field lines, reflecting the new constant speed (v_2) of the charge, is delayed in time. A snapshot of the electric field lines is shown in Figure 6.1b, which is before the spatial update of the electric field lines is completed.

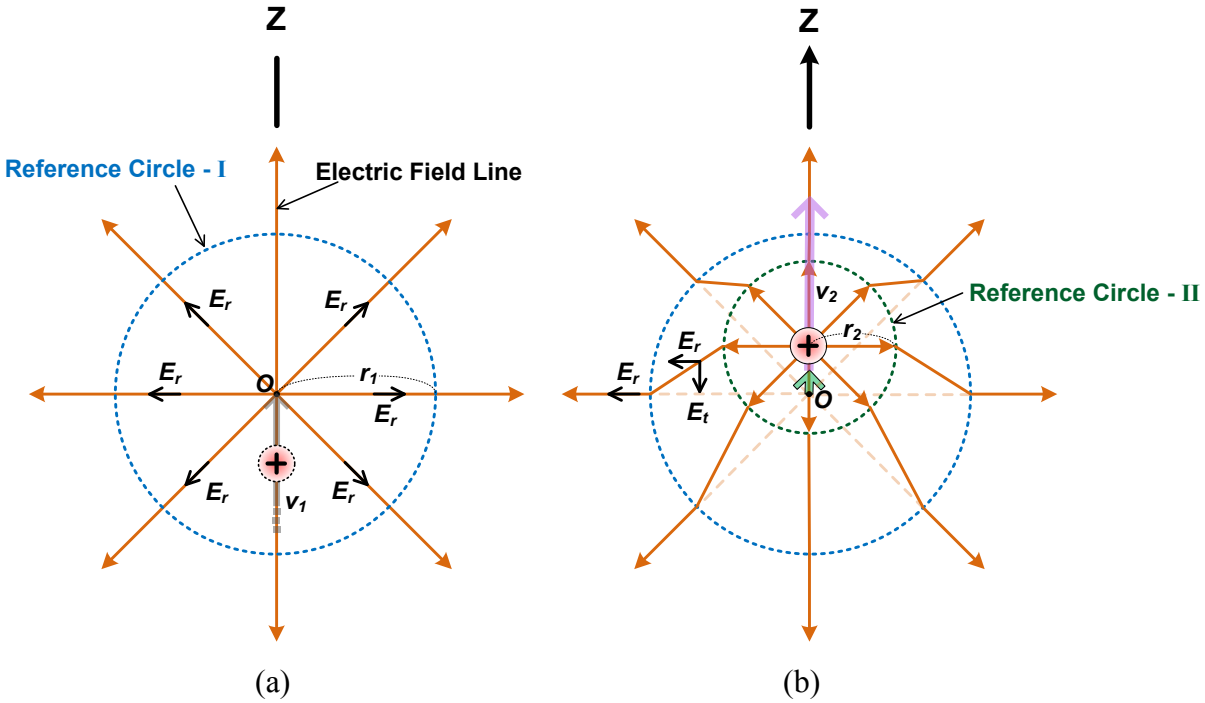


Figure 2.1 Illustration of single-charge radiation model – (a) Electric field lines at the origin for a single charge moving with a constant speed and (b) Electric field lines due to accelerated charge movement before spatial field-line update is completed.

Due to the finite propagating speed of the electric fields, the updated field lines of the charge with the speed of v_2 only appear in a limited circular area with the radius of r_2 (reference circle-II). The electric fields beyond the circular area with the radius of r_1 (reference circle-I) are the same as the original electric fields of the charge moving with the speed of v_1. In other words, the spatial electric-field update that occurred during the charge acceleration time period cannot be observed outside of the reference circle-I. Because field lines are continuous, the straight field lines outside the reference circle-I and inside the reference circle-II are connected to each other. The overall shape of the resulting electric field lines looks like a straight line with a kink

between the reference circles. Thus, the single charge model is sometimes referred to as "kink model" [*E.K. Miller*, 2001].

The kink portion of the electric field lines can be decomposed in to radial (E_r) and tangential vector (E_t) components, as depicted in Figure 6.1b. The radial vector component is the same as the electric field line of charge moving with constant speed. The tangential vector component is added during the time period of charge acceleration. Thus, it appears that the disturbed shape (kink) in the electric field lines is caused from the additional tangential vector (E_t) component due to charge acceleration. As we discussed earlier, radiation occurs from charge acceleration. Thus, it may be logical to assume that the tangential vector component (E_t) of the disturbed field lines contribute to radiation.

The single charge model visually explains how radiation fields can be generated from the accelerated charges, based on two basic physical principles, i.e. the finite propagating speed of the fields and the continuous field lines. However, the single charge model cannot explain how the disturbed field lines contributing to radiation can be decoupled from the charge, even in case of an oscillating single charge.

2.1.2 The Dual Charge Model

In 1880s, Hertz derived field equations for the case of two equal charges with opposite signs and oscillating in a straight-line path, creating a dipole moment. Hertz's field-line drawings for the dipole moment in his first book [*Hertz*, 1893] have provided insight for many researchers and engineers to understand how an antenna radiates. The simplified field-line model has been used in many antenna textbooks [*Balanis*, 2005; *Kraus and Marhefka*, 2002; *Stutzman and Thiele*, 2012]. A similar radiation concept model is presented in Figure 2.2 through Figure 2.6.

Consider two equal charges with opposite signs aligned in a vertical path with a separation distance, as shown in Figure 2.2a. These charges oscillate with varying speed in the vertical direction (the dotted lines in Figure 2.2a), similar to the case of a dipole antenna with a sinusoidal source forcing charges to oscillate. At the initial stage ($t = 0$), both charges are located at the ends of the vertical path with the maximum separation distance. Electric field lines start from the positive charge and terminate at the negative charge. Compared to the simple field lines in Figure 2.2a, the total electric field lines shown in Figure 2.2b including both radial and angular components provide an overall picture of the electric field distribution. The total electric

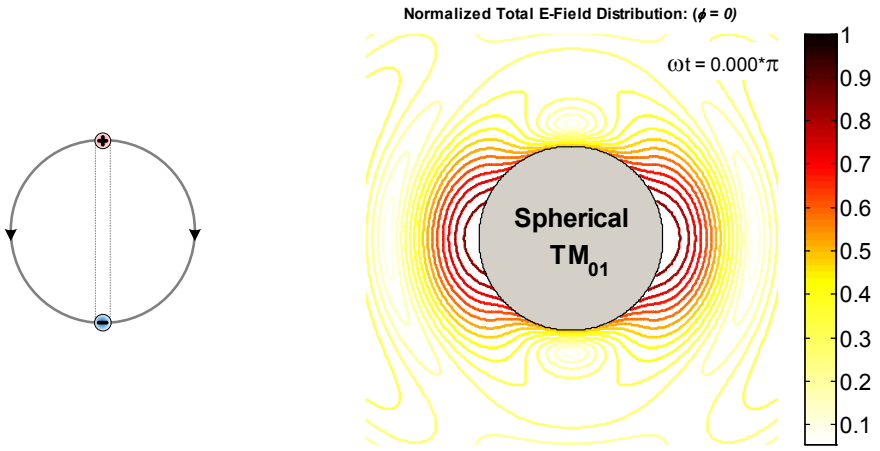
field lines were computed from the exact time-harmonic field solution of the spherical TM_{01} mode (electric dipole mode) [*R F Harrington*, 2001] and normalized to the peak amplitude of the total electric fields in the entire oscillating period. Because the total electric field lines were obtained from a time-harmonic solution, the entire space is filled with electric field lines that were generated in the previous period. Thus, the field lines in Figure 2.2a represent a simplified version of complete the field lines in Figure 2.2b.

After the initial stage, the charges oscillate with varying speed, i.e. moving toward each other. It may be worthwhile to note that the oscillation is caused by an external force, even though there is a force of attraction originated from the charges with opposite signs. The electric field lines at time $t = 1/8\ T$ (where T = the period of oscillation) of the oscillation are depicted in Figure 2.3. We can observe that the electric field lines are deformed from the original circular shape. In addition, strong electric fields start building up at the top and bottom of the vertical path in Figure 2.3b, which is not often pointed out in the antenna textbooks.

At time $t = 2/8\ T$ (Figure 2.4a), the charges pass the center location of the vertical path. New field lines are generated due to the reversed orientation of charges in vertical direction. As a result, the old electric field lines start to decouple from the charges due to the electric force of repulsion. The electric fields at the top and bottom of the vertical path increase, as shown in Figure 2.4b.

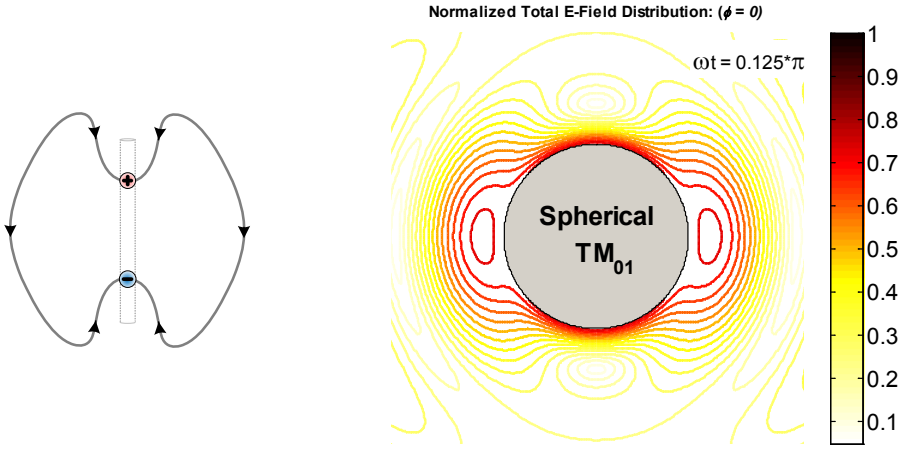
As the charges move further toward the ends of the vertical path (see Figure 2.5 and Figure 2.6), we can clearly observe that the closed loop shape of the electrical lines are completely decoupled and move away from the charges. The decoupled, closed electric field lines contribute to far-field radiation, but not all of them. Some of them will be absorbed into the charges and reradiate later (see Section 2.2.4 for details). In addition, we can observe similar closed electric field lines at the top and bottom of the vertical path, which also contribute to radiation. However, the overall amplitude of the closed electric field lines at the top and bottom is significantly smaller than that of the closed electrical lines to the side of the vertical path. Thus, dominant radiation would be observed at the far-field region from the side of the vertical path. On the other hand, we may think that a dominant energy storage process may occur due to the dominant electric fields at the top and bottom of the vertical path. Thus, a field pattern at near-field region

would be a spheroidal shape with amplitude peaks at the top and bottom direction, compared to the typical far-field radiation pattern of ideal dipole with a rotated figure-8 shape.



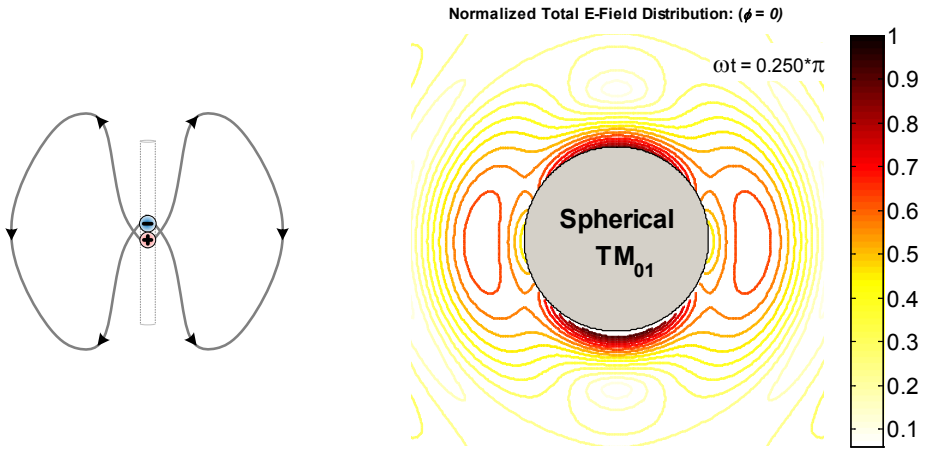
(a) (b)

Figure 2.2 Electric field lines at $t = 0$ – (a) Simplified field model (E_θ) and (b) Computed total electric field lines (E_r and E_θ) from the exact time-harmonic solution of the spherical TM_{01} mode.



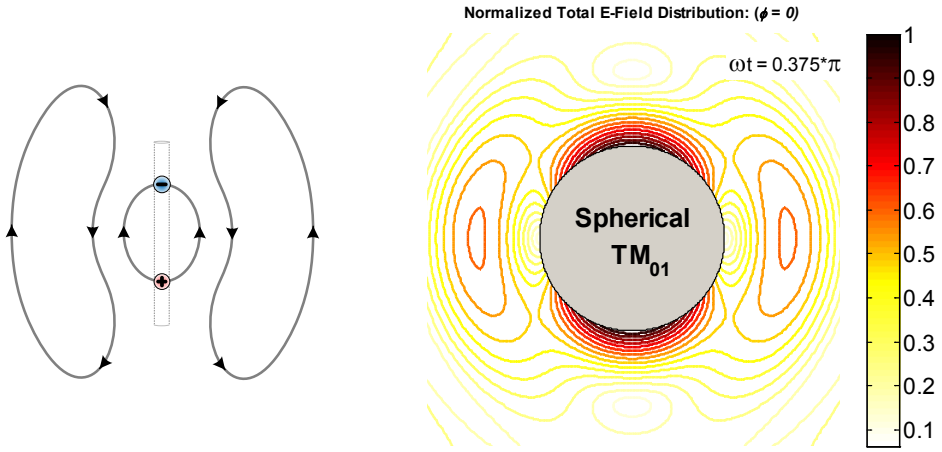
(a) (b)

Figure 2.3 Electric field lines at $t = 1/8\ T$ – (a) Simplified field model and (b) Computed total electric field lines from the exact time-harmonic solution of the spherical TM_{01} mode.



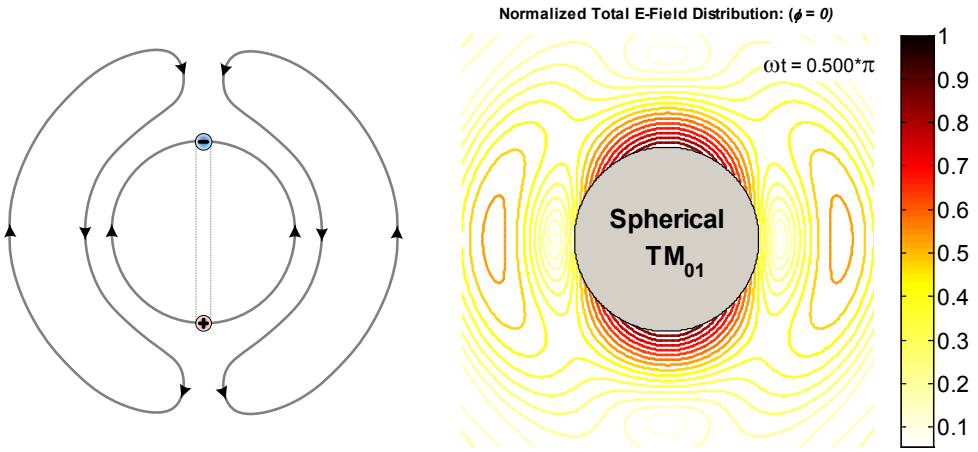
(a) (b)

Figure 2.4 Electric field lines right after $t = 2/8\ T$ – (a) Simplified field model and (b) Computed total electric field lines from the exact time-harmonic solution of the spherical TM_{01} mode.



(a) (b)

Figure 2.5 Electric field lines at $t = 3/8\ T$ – (a) Simplified field model and (b) Computed total electric field lines from the exact time-harmonic solution of the spherical TM_{01} mode.



(a) (b)

Figure 2.6 Electric field lines at $t = 4/8\ T$ – (a) Simplified field model and (b) Computed total electric field lines from the exact time-harmonic solution of the spherical TM_{01} mode.

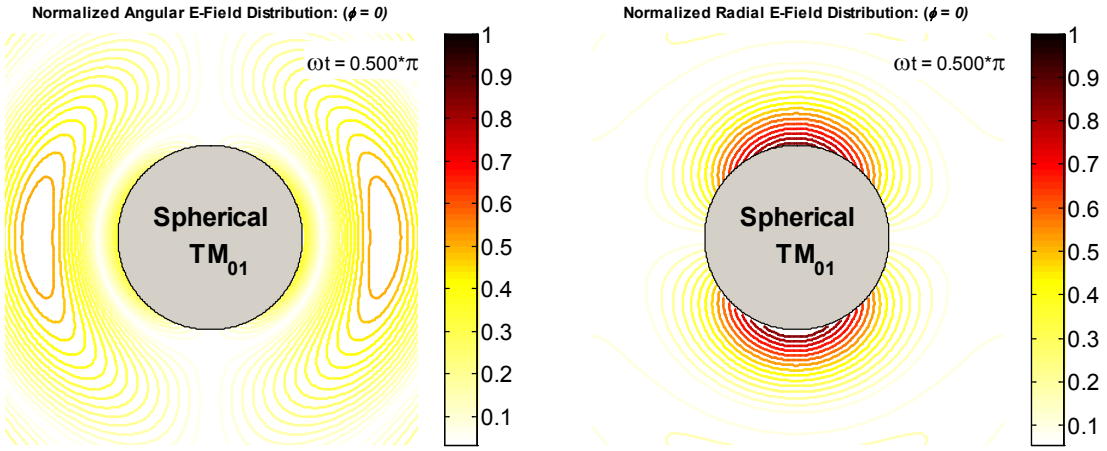
Unlike the complete total electric field distribution presented in this sub section, Hertz's original field-line drawings for the dipole moment [*Hertz*, 1893] are more similar to the amplitude-contour plots that only considers the angular component (E_θ) of total electric fields (see Figure 2.7a). Drawings similar to Hertz's original drawings have been used in some antenna textbooks [*Balanis*, 2005; *Kraus and Marhefka*, 2002]. However, the detailed field-line drawings that do not clearly show the contribution of radial electric component (E_r) may be misleading. This is because the closed loops of the electric field lines are composed from both an angular component (E_θ) and a radial component (E_r) of electric field. Without the contribution of radial electric component (E_r), there would not be a decoupled, closed loop of the electric field lines nor a displacement current in space, i.e. no radiation.

Neither Figure 2.7a (angular electric-field only) nor Figure 2.7b (radial electric-field only) represent actual physical radiation phenomena.

In fact, the radial electric field (E_r) dominates the near fields during the radiation time frame from $t = 0\ T$ to $t = 4/8$ T. This may suggest that energy storage may dominant over radiation phenomena. This aspect will be clarified in the next section.

Compared to the single charge model in the previous section, the simple dual charge model (dipole model) is an excellent tool to explain the field-decoupling process (radiation). However,

the dipole model does not show an energy storage mechanism. Unless the radiation and the energy storage processes are independent physical phenomenon, the conventional simple dipole model does not show a complete picture of the radiation process.



(a) (b)

Figure 2.7 Computed, decomposed electric field lines from the exact time-harmonic solution of the spherical TM_{01} mode at $t = 4/8\ T$ – (a) Angular electric field lines and (b) Radial electric field lines. Magnitude of each field lines in both plots was normalized with the peak magnitude of total electric field in the entire oscillating period.

2.2 Alternative Viewpoints

The reviewed classical single and dual-charge models in the previous section helped us to understand the cause of radiation (accelerated charges) and the decoupling process of radiation fields through illustrations. However, classical models only focus on the electric fields that mainly contribute to far-field radiation. The transition process from the reactive near-field region around an antenna to the far-field region is not included in the classical radiation models. In addition, the energy storage process is not explained. Thus, the classical radiation modes do not provide an overall picture of the antenna radiation mechanism that includes energy storage. In this section, we try to obtain an overall picture of the antenna radiation mechanism by investigating ideal dipoles with some alternative viewpoints.

2.2.1 The Wave Impedance Viewpoint

The exact fields of the ideal electric dipole (spherical TM_{01} mode) in spherical coordinates with $e^{j\omega t}$ time-dependence and a source normalization factor of $Idl = 4\pi / k$ in the frequency domain are [*Stutzman and Thiele*, 2012]

$$\begin{aligned}\vec{E} = \hat{r}\, jk\eta_0 2\cos\theta\left[\frac{1}{(jkr)^2}+\frac{1}{(jkr)^3}\right]e^{-jkr} \\ +\hat{\theta}\, jk\eta_0 \sin\theta\left[\frac{1}{(jkr)}+\frac{1}{(jkr)^2}+\frac{1}{(jkr)^3}\right]e^{-jkr}\end{aligned} \tag{2.4}$$

$$\vec{H} = \hat{\phi}\, jk\sin\theta\left[\frac{1}{(jkr)}+\frac{1}{(jkr)^2}\right]e^{-jkr} \tag{2.5}$$

where k is the wave number and η_0 is the free-space wave impedance (≈ 377 Ω). From the field equations, radial wave-impedances can be defined as

$$Z_r^{TM_{01}} = \frac{E_\theta}{H_\phi} = \eta_0 \frac{1+(jkr)+(jkr)^2}{(jkr)+(jkr)^2} \tag{2.6}$$

Using the duality between the ideal electric and magnetic dipole [*R F Harrington*, 2001], i.e. $\vec{E}_{TM_{01}} = \eta_0 \vec{H}_{TE_{01}}$ and $\vec{H}_{TM_{01}} = -\vec{E}_{TE_{01}} / \eta_0$, the fields of the ideal magnetic dipole (spherical TE_{01} mode) can be easily obtained. The corresponding radial wave impedance for the ideal magnetic dipole is then

$$Z_r^{TE_{01}} = -\frac{E_\phi}{H_\theta} = \eta_0 \frac{(jkr)+(jkr)^2}{1+(jkr)+(jkr)^2} \tag{2.7}$$

As kr increases, both radial wave impedances approach the free-space wave impedance (η_0). On the other hand, the asymptotic behaviors ($kr \ll 1$) of the wave impedances for the electric and magnetic ideal dipoles are, respectively, given as

$$Z_{r,kr\ll 1}^{TM_{01}} = \frac{\eta_0}{jkr} \tag{2.8}$$

$$Z_{r,kr\ll 1}^{TE_{01}} = \eta_0 (jkr) \tag{2.9}$$

The normalized magnitudes of radial wave impedances versus kr are shown in Figure 2.8 for both spherical TM_{01} and TE_{01} modes along with their asymptotic lines for $kr \ll 1$, which is

similar to that presented in [*Capps*, 2001]. When *kr* approaches zero, $\left|Z_r^{TM_{01}}\right|$ of the spherical TM_{01} mode (ideal electric dipole) becomes very high (open circuit) while $\left|Z_r^{TE_{01}}\right|$ of the spherical TE_{01} mode (ideal magnetic dipole) becomes very low (short circuit). The magnitudes of both wave impedances vary rapidly in the range of $kr < 1$. The rapid change in the wave impedance magnitude from very low to high values implies the energy storage process is a dominant phenomenon in the range of $kr < 1$. Above $kr = 10$, the magnitudes of both wave impedances are close to the free-space wave impedance (η_0). The magnitudes of both wave impedances are found to equal η_0 at $kr = 1/\sqrt{2}$. On the other hand, the asymptote lines of both $\left|Z_r^{TM_{01}}\right|$ and $\left|Z_r^{TE_{01}}\right|$ for $kr \ll 1$ cross at $kr = 1$, where the magnitudes of both wave impedances again equal that of free space. The virtual sphere with the radius of $r_0 = 1/k$ is known as *radian sphere* [*Wheeler*, 1959]. The radian sphere has been used as criteria to identify a reactive near-field zone ($r \le r_0$) [*Hansen*, 1988].

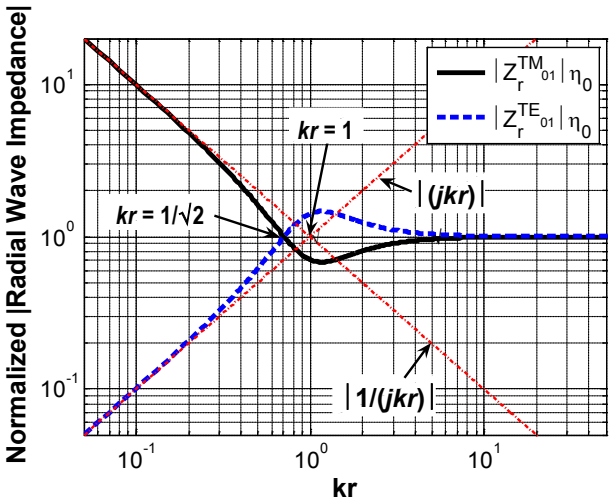


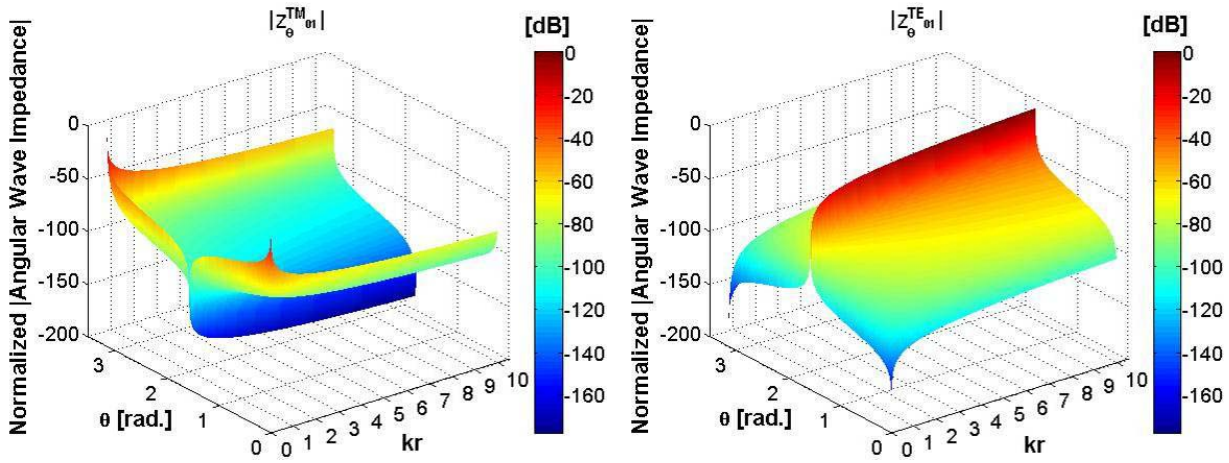
Figure 2.8 Normalized magnitude of radial wave-impedances and asymptotic lines ($kr \ll 1$) versus *kr* for spherical TM_{01} and TE_{01} modes.

Similarly, the angular wave impedances for both spherical modes are

$$Z_\theta^{TM_{01}} = -\frac{E_r}{H_\phi} = -2\eta_0 \frac{\cot\theta}{(jkr)} \tag{2.10}$$

$$Z_\theta^{TE_{01}} = \frac{E_r}{H_\phi} = -\frac{\eta_0}{2} \frac{(jkr)}{\cot\theta} \tag{2.11}$$

The angular wave impedance is capacitive (inversely proportional to *jkr*) for the spherical TM_{01} mode and inductive (proportional to *jkr*) for the spherical TE_{01} mode. As depicted in Figure 2.9a, the magnitude of $Z_{\theta}^{TM_{01}}$ has peaks around $\theta = 0$ and $\theta = \pi$ while the magnitude becomes a minimum at $\theta = \pi/2$ for a given *kr*. Thus, a capacitive energy storage process is dominant around $\theta = 0$ and $\theta = \pi$ for radiation from an ideal electric dipole antenna. For the spherical TE_{01} mode in Figure 2.9b, the maximum and minimum magnitudes of the wave impedance occur at opposite angles, i.e. peaks are around $\theta = \pi/2$ and minima at $\theta = 0$ and $\theta = \pi$. Thus, an inductive energy storage process is dominant around $\theta = \pi/2$ for the case of ideal magnetic dipole antenna.



(a) (b)

Figure 2.9 Normalized magnitude of angular wave-impedances versus *kr* and θ for spherical TM_{01} and TE_{01} modes.

The impedance viewpoint provides some insight into the spatial and angular reactive-field ranges, where a dominant energy storage process occurs during the radiation process. Outside the reactive-field range, radiation dominates. However, these observations are yet to be substantiated. Also, information obtained from the frequency domain is limited. Time-domain wave impedance can be obtained from time-harmonic field solutions, but the meaning of time-domain impedance is not clear. Thus, we investigate power, Poynting vector, and energy velocity viewpoints in the time domain in the following sub sections.

2.2.2 The Power Viewpoint

The time harmonic field solutions can easily be found from frequency domain solutions of (2.4) and (2.5) by taking real parts of the products between fields with time dependence $e^{j\omega t}$ [*R F Harrington*, 2001], i.e.

$$\begin{aligned}\vec{E}(r,t) = \hat{r}2\eta\frac{\cos\theta}{r}\left[\frac{\sin(\omega t-kr)}{kr}-\frac{\cos(\omega t-kr)}{k^2r^2}\right] \\ +\hat{\theta}\eta\frac{\sin\theta}{r}\left[\cos(\omega t-kr)+\frac{\sin(\omega t-kr)}{kr}-\frac{\cos(\omega t-kr)}{k^2r^2}\right]\end{aligned} \tag{2.12}$$

and

$$\vec{H}(r,t) = \hat{\phi}\frac{\sin\theta}{r}\left[\cos(\omega t-kr)+\frac{\sin(\omega t-kr)}{kr}\right] \tag{2.13}$$

Then, the corresponding total power through the spherical surface at kr is found as

$$\begin{aligned}P(r,t) &= \oint_s \vec{E}\times\vec{H}\cdot\hat{r}ds \\ &= P_{av}\left[1+\left(1-\frac{2}{(kr)^2}\right)\cos 2(\omega t-kr)+\left(\frac{2}{kr}-\frac{1}{(kr)^3}\right)\sin 2(\omega t-kr)\right]\end{aligned} \tag{2.14}$$

where P_{av} is time average of the total power ($=\eta_0 4\pi/3$). The time-space plot of the normalized total power is shown in Figure 2.10. The positive sign for the total power represents outgoing propagation toward far-field region, whereas a negative power represents power returning to the source. The energy storage process is clearly observed inside the radian sphere ($kr<1$). There appears to be a periodic energy leak from the radian sphere to produce the radiation. This is the process of converting reactive energy to radiated energy, which is discussed by Schantz [*H G Schantz*, 2001]. It may be worthwhile to emphasize that what we can observe from the power flow plots is a net quantity. In other words, we can only indentify the net behaviors of radiation and energy storage through the time-space plots.

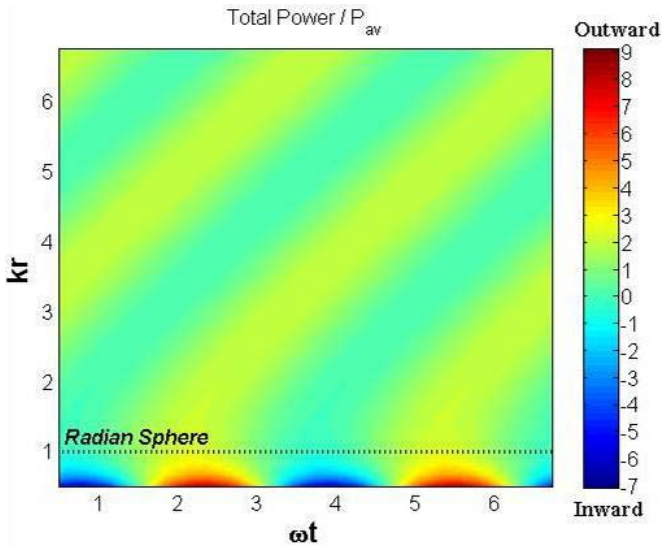


Figure 2.10 Time-space plot of normalized total power of ideal electric dipole (spherical TM_{01} mode) [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

In Figure 2.10, the starting time for energy radiation and storage mechanisms cannot be clearly observed because energy storage dominates inside the radian sphere. In order to separate real and reactive powers, (2.14) can be rearranged to give [*Davis et al.*, 2011]

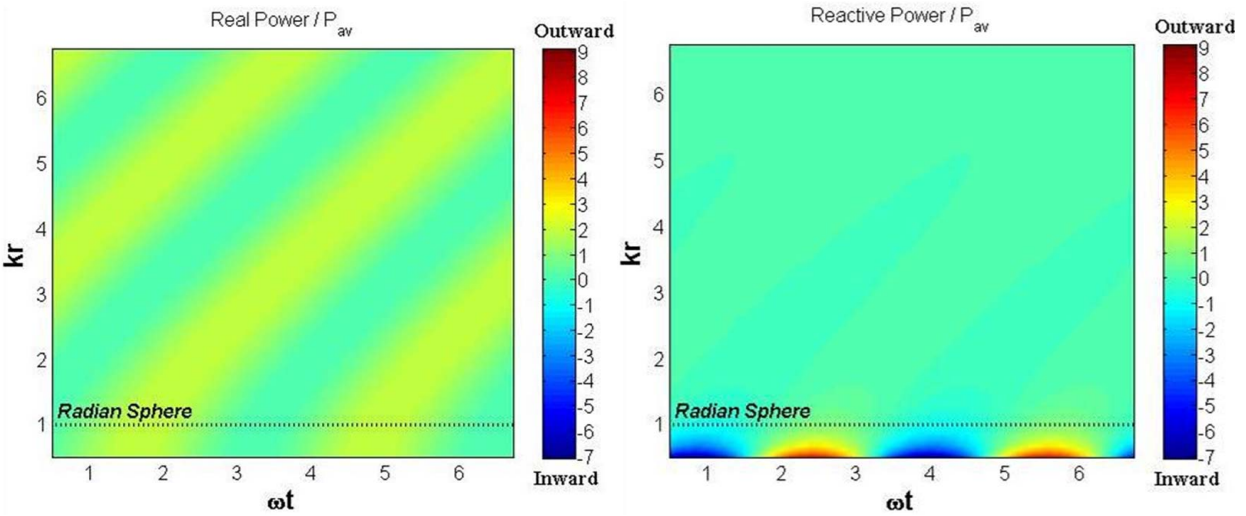
$$P(r,t) = P_{av}\left[\left[1+\cos 2\left(\omega t - kr - \zeta\right)\right] + \frac{1}{(kr)^3}\sin 2\left(\omega t - kr - \zeta\right)\right] \tag{2.15}$$

where $\zeta = \tan^{-1}(1/kr)$ is the excess phase delay of the radiated field. The first and second terms of (2.15) are related to the radiated power [$\cos^2(\cdot)$], and the third term is related to reactive power [$\sin(\cdot)\cos(\cdot)$]. These radiated and reactive powers are 90° out of phase, which is consistent with circuit definitions of real and reactive power and appropriate for the linear polarization of this mode.

Time-space plots for the real and reactive power terms in (2.15) are depicted in Figure 2.11. From Figure 2.11a, it appears that far-field radiation is more like a traveling wave phenomenon originating from somewhere between $\omega t = 2\pi/8$ and $\omega t = 3\pi/8$. The time frame of $\omega t = 2\pi/8$ corresponds to the case that two charges with opposite signs are passing the center of the ideal electrical dipole (see Figure 2.4). The next radiation starts at $\omega t = 2\pi/8 + \pi$. Thus, the period of the radiation process is π, which is also obvious from the factor of two in the sinusoidal functions

of (2.14). Notice that the visually identifiable starting time for the radiation in Figure 2.10 appears to be shifted to a later time, around $\omega t = 4\pi / 8$.

We can also observe the energy storage process through the reactive power flow in Figure 2.11b. The reactive power flow occurs in both outward and inward directions, which is similar to a standing wave phenomenon. In order words, it appears that the reactive power waves are circulating between the antenna terminal and the radian sphere, but not directly contributing to far-field radiation.



(a) (b)

Figure 2.11 Time-space plots of the normalized real and reactive power of ideal electric dipole (spherical TM_{01} mode) – (a) Real power and (b) Reactive power.

It is commonly assumed that radiated energy propagates in free space at the speed of light [*Balanis*, 2005; *R E Collin and Zucker*, 1969; *Elliott*, 2003; *Schelkunoff*, 1952; *Stutzman and Thiele*, 2012]. However, if that is true, the trajectory of the real-power flow, the yellow "band" in Figure 2.11a, would be a straight line, i.e. $k\Delta r / \omega\Delta t = 1$. However, we can clearly observe the trajectory curve of the real power flow in Figure 2.12 is not linear in the near-field region, where the radiation and energy storage processes occur. Compared to the dotted reference line with a 45-degree slope, the slop angle of the real-power flow is higher than 45 degrees in the near-field region. Thus, the phase speed of the power flow in the near-field region appears to be faster than the speed of light, which implies that the energy velocity is slower than speed of light. Thus, Figure 2.12 suggests that real power flow gets started with a slower speed than the speed of light, accelerated in near-field region, and reached to the speed of light in the far-field region. Thus,

the speed of real power flow is not always constant in the antenna radiation process even in free-space environment; this has been overlooked in the literature.

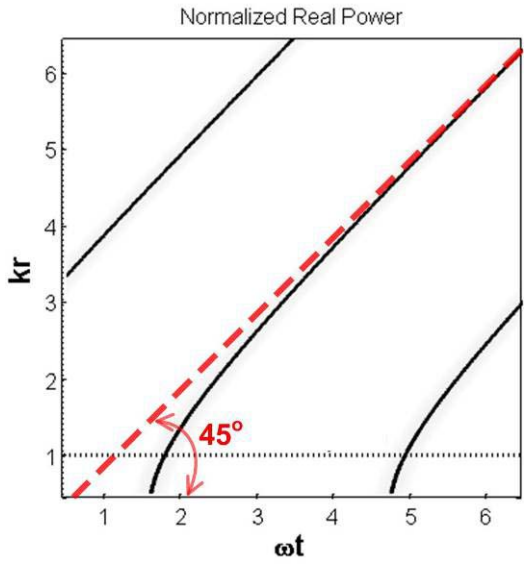


Figure 2.12 Trajectory of real-power radiation in time and space coordinate, compared to the reference linear path. The trajectory of real-power radiation was obtained from the time-space image of Figure 2.11a through a color filtering process.

The increased phase speed of real power flow in the near-field region is related to phase behavior of the radiated fields. The fields of the ideal electric dipole in (2.12) and (2.13) can be rewritten, including the excess phase ($\zeta = \text{atan}(1/kr)$) [*Davis et al.*, 2011] as

$$\begin{aligned}\vec{E}(r,t) = \hat{r}2\eta_0 \frac{\cos\theta}{r}\frac{\sqrt{1+k^2r^2}}{k^2r^2}\sin(\omega t - kr - \zeta) \\ + \hat{\theta}\eta_0 \frac{\sin\theta}{r}\frac{kr}{\sqrt{1+k^2r^2}}\left[\cos(\omega t - kr - \zeta) + \frac{1}{k^3r^3}\sin(\omega t - kr - \zeta)\right]\end{aligned} \tag{2.16}$$

and

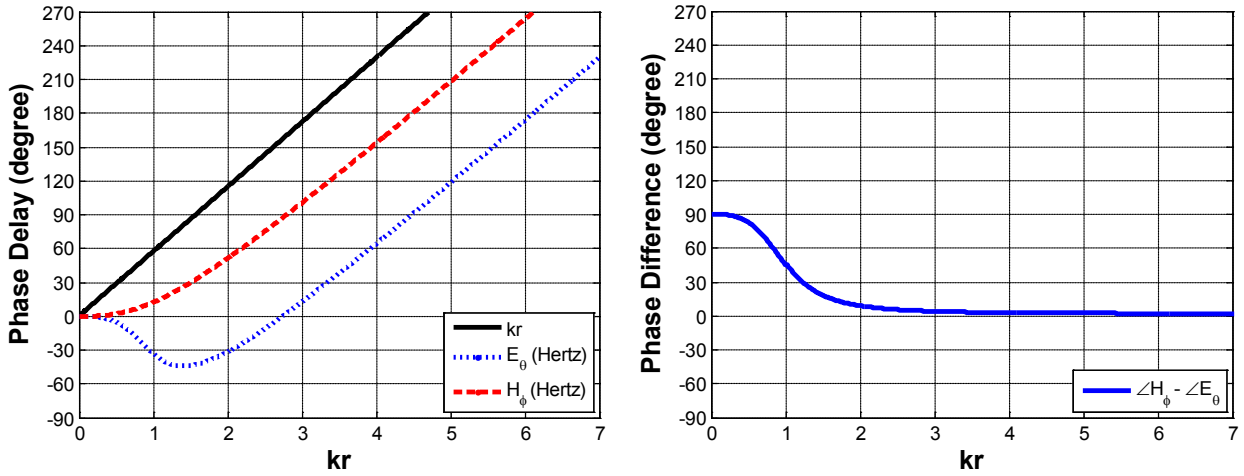
$$\vec{H}(r,t) = \hat{\phi}\frac{\sin\theta}{r}\frac{\sqrt{1+(kr)^2}}{kr}\cos(\omega t - kr - \zeta) \tag{2.17}$$

The θ component of the electric field combining the sinusoidal terms is found as

$$E_\theta = \eta_0 \frac{\sin\theta}{r}\frac{kr}{\sqrt{1+k^2r^2}}\frac{\sqrt{1+k^6r^6}}{k^3r^3}\cos(\omega t - kr - \zeta - \gamma) \tag{2.18}$$

where $\gamma = \tan^{-1}(1/k^3r^3)$. The overall phase delay of E_θ and H_ϕ is shown in Figure 2.13a, compared to the reference phase delay (*kr*). We can clearly observe that E_θ and H_ϕ are involved in the excess phase. Therefore, the excess phase in the fields causes the reduced speed of the real power flow in the near-field region.

The phase difference between E_θ and H_ϕ (obtained from the phase plots of Figure 2.13a) can be determined subject to a phase reference of zero degrees in the far-field region. As shown in Figure 2.13b, the phase difference between E_θ and H_ϕ becomes close to zero degrees at *kr* > 6.



(a) (b)

Figure 2.13 Field phase delay and difference of electric and magnetic fields of the ideal electric dipole (spherical TM_{01} mode) – (a) Phase delay of E_θ and H_ϕ compared to reference phase response without excess delay [*Hertz*, 1962] and (b) Phase difference between E_θ and H_ϕ.

The time-space plots of the power flow allowed us to understand the far-field radiation (traveling wave) and energy-storage (standing wave) mechanisms better, compared to the classical radiation models. In addition, we observed that the power does not propagate with the speed of light in near-field region. However, the power viewpoint has the shortcoming that the identifiable direction of the power flow is limited to either inward or outward directions.

2.2.3 The Poynting Vector Viewpoint

The Poynting vector ($\vec{S} = \vec{E} \times \vec{H}$) describes power density. Due to the vector nature of the Poynting vector, analysis using the Poynting vector provides additional information to help understand radiation and energy storage, compared to the power viewpoint in the previous sub section. Thus, there was some effort to investigate the Poynting vector by decomposing fields of the ideal electric dipole antenna into radiated and non-radiated terms [*Liénard*, 1898; *Wiechert*, 1901]. In order to explain the basic concept, consider the hypothetically-decomposed fields in time domain, i.e.

$$\vec{E}_{total} = \vec{E}_{non-rad} + \vec{E}_{rad} \tag{2.19}$$

$$\vec{H}_{total} = \vec{H}_{non-rad} + \vec{H}_{rad} \tag{2.20}$$

The corresponding Poynting vector may be written as

$$\vec{S}_{total} = \vec{S}_{non-rad} + \vec{S}_{rad} + \left(\vec{E}_{non-rad} \times \vec{H}_{rad}\right) + \left(\vec{E}_{rad} \times \vec{H}_{non-rad}\right) \tag{2.21}$$

Then, two fundamental questions arise for this approach. First, what is the physical meaning of the crossed terms ($\vec{E}_{non-rad} \times \vec{H}_{rad}$ and $\vec{E}_{rad} \times \vec{H}_{non-rad}$)? Secondly, how we can separate field terms into radiated and non-radiated terms? The first question may be hard to answer, but one may say that it is obvious to separate the field terms because we can select fields with $1/r$ terms. As we saw in the classical radiation models, the field terms with a $1/r$ factor were directly related to the accelerated charges. This aspect has been widely adopted in electromagnetics community. It also has been presented in antenna textbooks [*Balanis*, 2005; *Kraus and Marhefka*, 2002; *Stutzman and Thiele*, 2012]. For the ideal electrical dipole antenna case, only the θ component in (2.12) has a $1/r$ term in the electric field:

$$\vec{E}_{rad}(r,t) = \hat{\theta}\eta_0 \sin\theta \frac{\cos(\omega t - kr)}{r} \tag{2.22}$$

However, we observed that the loops of electric fields, which decouple from the antenna, have both angular (θ) and radial components. In addition, $\vec{E}_{rad}(r,t)$ does not satisfy Maxwell's equations. In order to demonstrate it, consider the divergence of the radiation electric field:

$$\nabla \cdot \vec{E}_{rad}(r,t) = 2\eta_0 \cos\theta \frac{\cos(\omega t - kr)}{r^2} \tag{2.23}$$

If we recall Gauss' theorem [*R F Harrington*, 2001], the divergence of any electric field in source-free region should be zero. Apparently (2.23) is not zero except at $\theta = \pi / 2$. Thus, only the electric field term with $1/r$ does not satisfy Maxwell's divergence equation.

Then, the next question is what term(s) would be missed. Recognizing that we may need a radial term and observing (2.23) has a $1/r^2$ factor, consider the additional term from a radial component in (2.12), i.e.

$$\vec{E}_{rad,add}(r,t) = \hat{r} 2\eta_0 \cos\theta \frac{\sin(\omega t - kr)}{kr^2} \tag{2.24}$$

The divergence of the additional electric field term is found as

$$\nabla \cdot \vec{E}_{rad,add}(r,t) = -2\eta_0 \cos\theta \frac{\cos(\omega t - kr)}{r^2} \tag{2.25}$$

which is the same as (2.24) except for the opposite sign. Thus, we have

$$\nabla \cdot \left(\vec{E}_{rad}(r,t) + \vec{E}_{rad,add}(r,t) \right) = 0 \tag{2.26}$$

suggesting that the radiated electric field does not travel radially at the speed of light.

On the other hand, divergence of any magnetic field term in (2.13) becomes zero because (2.13) does not have a ϕ dependence. Thus we would find the radiating magnetic fields by taking the curl ($\nabla \times$) of the radiating electric fields. However, the resulting radiating magnetic fields are not similar to the original magnetic fields in (2.13). Thus, the physical meaning of the resulting radiating magnetic fields is not clear.

We also attempted to separate the fields of the dipole into waves with the forward ($+\hat{r}$) and backward ($-\hat{r}$) directions. Four vectors were defined, i.e. forward radial and angular vectors and backward radial and angular vectors. Various approaches were considered, but we could not separate the fields. Especially, there were some coupled terms between forward and backward waves in the process of the separation. Basically, the radial and theta field components were correlated. Even breaking up the spherical Hankel function into two Bessel functions (two different standing waves with a phase difference) was not successful. This result indicates that the radiation phenomenon cannot be easily separated into forward and backward traveling waves.

In fact, Hertz also tried to separate the field terms into radiating and non-radiating terms after he found the unexpected phase responses of the electromagnetic fields for the ideal electric dipole antenna [*Hertz*, 1893]. However, Hertz also could not separate the fields. He indicated

that the dipole fields may not be separated into radiating and non-radiating terms. To date, there has been no known method to separate the dipole fields. Considering that Maxwell's equations are a macroscopic model [*Heaviside*, 1925], the field separating approach may violate the applicability of Maxwell's equations. We may need to consider quantum electrodynamics [*Feynman*, 1961] to achieve the field separation, but this is beyond the scope of this study.

From the power viewpoint in the previous sub section, we observed that real power flow looks like a traveling wave phenomena while the reactive power looked like a standing wave phenomena. This observation encourages us to separate Poynting vector terms into traveling and standing wave terms. Consider the radial Poynting vector ($S_r = E_\theta H_\phi$) of the ideal electric dipole, including the excess phase (ζ), i.e.

$$\begin{aligned} S_r = \frac{\eta_0}{2r^2}\Bigg\{\sin^2\theta\Bigg[\ 1\ +\ \left(1-\frac{1}{k^3r^3}\right)\ \cos 2(\omega t-(kr+\zeta)) \\ +\frac{1}{k^3r^3}\big[\cos 2\omega t\cos 2(kr+\zeta)-\cos 2\omega t\sin 2(kr+\zeta) \\ +\sin 2\omega t\sin 2(kr+\zeta)+\sin 2\omega t\cos 2(kr+\zeta)\big]\Bigg]\Bigg\} \end{aligned} \tag{2.27}$$

The term with $\sin^2\theta$ represents the radiation pattern of the ideal electric dipole. The constant number term would be related to the time-average radiated power. The second term has the form of a traveling wave, which would be related to the real power radiation. The plot of the second term versus *kr* in Figure 2.14 clearly shows a traveling-wave phenomenon. For $kr < 1$ region, the second term has a negative value at certain time period, even when we include the first term representing the time-average radiated power. The negative value indicates that power returns back to the antenna. Because the second term has the form of traveling wave, the inward power-density wave needs to be absorbed at the source. In this sense, the traveling inward power-density wave may be seen as a part of energy storage mechanism, but the inward power-density wave would be a reflected wave from the radian sphere boundary ($kr = 1$). Thus, there exist both outward and inward traveling power-density waves in the $kr < 1$ region. The amplitude of the power-density waves is damped due to the $1/k^3r^3$ factor in the second term.

Above $kr = 1$, the second term becomes a traveling wave without an amplitude damping. When the first term is included with the second term for $kr > 1$, the minimum value of the resulting sinusoidal waveform is zero during one period. Thus, there is only an outward power-

density wave above kr = 1, which represents far-field radiation. These observations for the traveling-wave second term suggest that some of radiation energy passes through the radian sphere boundary (kr = 1). The rest of the radiation energy returns to the antenna and the absorbed radiating energy at the antenna source radiates again. This radiation cycle repeats. Thus, it would appear that far-field radiation energy periodically leaks out of the radian sphere. In fact, the additional $1/r^2$ factor in (2.27) is considered to the second term, the overall power-density wave in Figure 2.15 looks like the impulse response of typical dipole antenna with late-time ringing (see Chapter 3 for details).

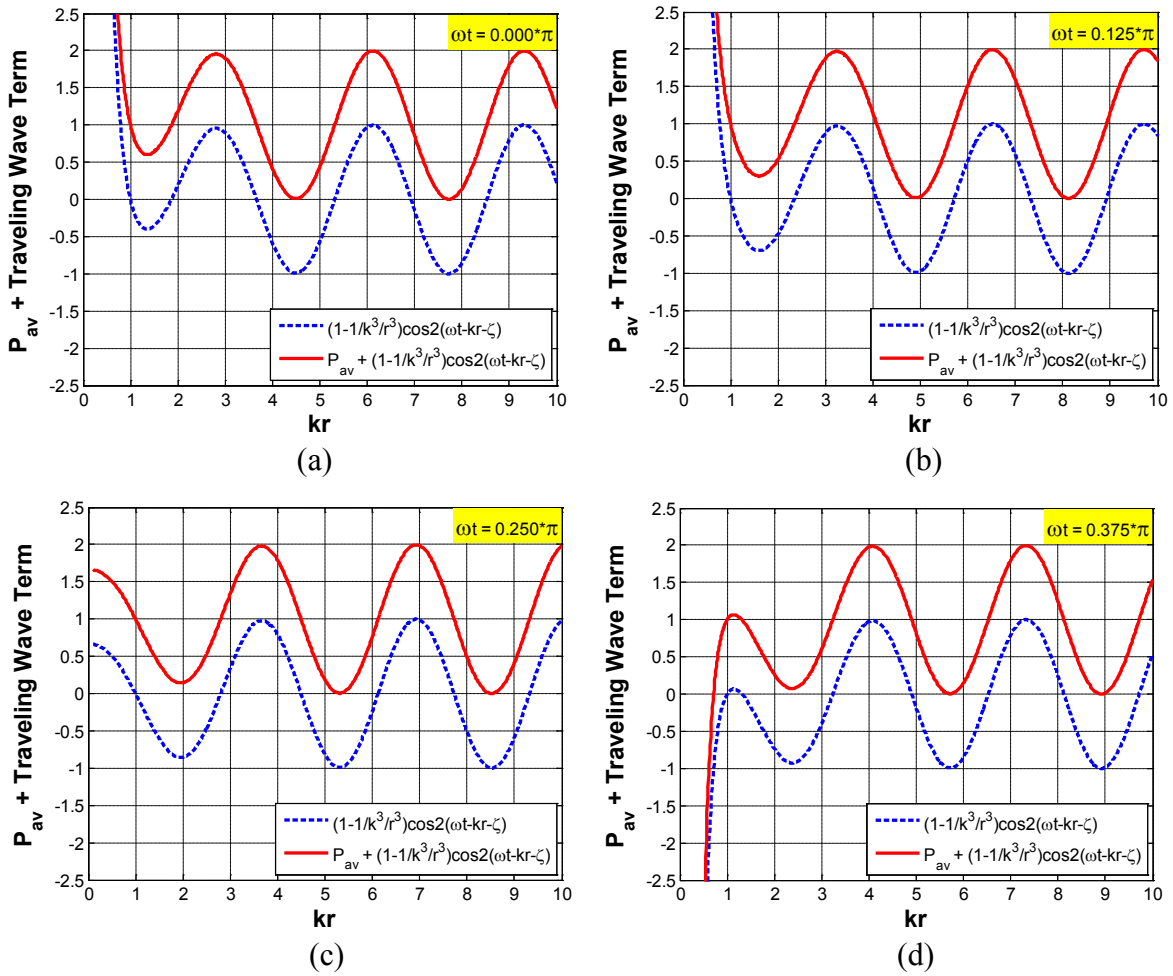


Figure 2.14 Traveling-wave term versus kr for the Poynting vector of the ideal electric dipole – (a) At t = 0, (b) At t = 1/8 T, (c) At t = 2/8 T, and (d) At t = 3/8 T.

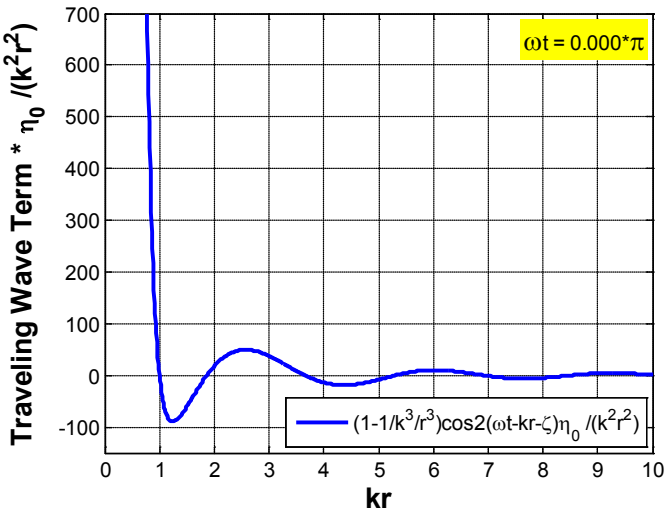


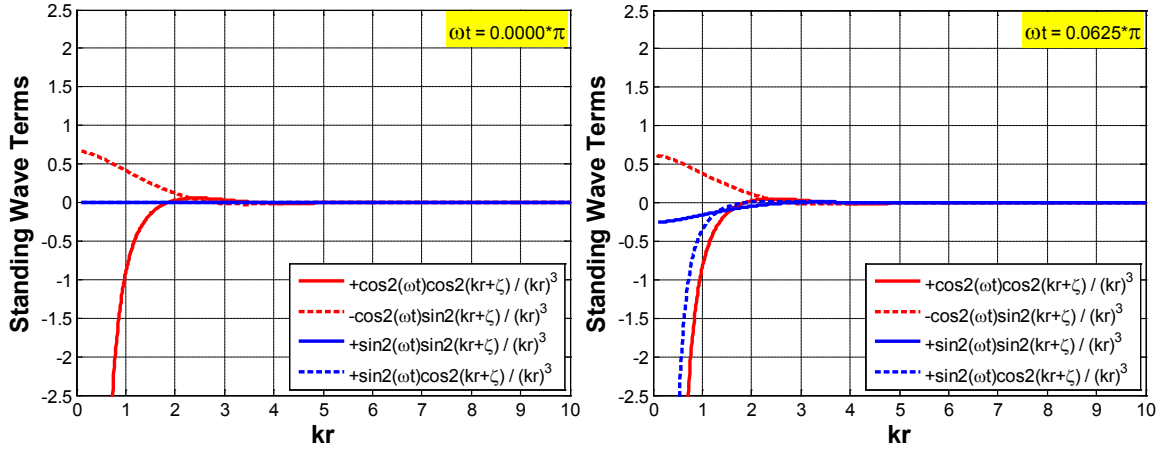
Figure 2.15 Traveling-wave term including $\eta_0 / k^2 r^2$ factor versus *kr* for the Poynting vector of the ideal electric dipole.

On the other hand, the rest of terms with a $1/k^3 r^3$ factor in (2.27) have a form of standing wave, which might be related to the reactive-power radiation. The time-average of these standing-wave terms is zero. The amplitude behaviors of these terms are compared in Figure 2.16. Due to the product between sine and cosine functions, each standing-wave term shows a different effective time-delay in oscillation. Unlike the traveling wave of the second term, the standing-wave terms almost vanish after *kr* = 5. Thus, these standing wave terms represent a reactive power circulation between antenna source and some point around the radian sphere (*kr* = 1).

The angular Poynting vector ($S_\theta = E_r H_\phi$) can be written as

$$S_\theta = -\frac{\eta_0}{2r^2}\frac{k^2r^2+1}{k^3r^3}\sin 2\theta\left[\sin 2\omega t\cos 2(kr+\zeta) - \cos 2\omega t\sin 2(kr+\zeta)\right] \qquad (2.28)$$

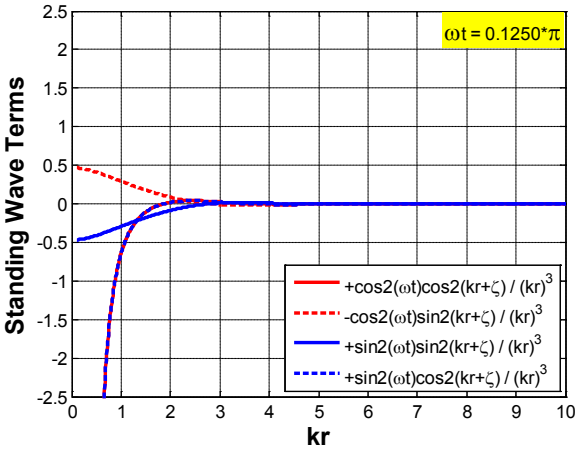
The first and the second terms also have standing wave form. Interesting enough, both standing-wave terms are exactly same as the standing-wave terms in (2.27), i.e. fourth and sixth terms in (2.27). Thus, it appears that both standing-wave terms in the angular Poynting vector might be related to some of the standing-wave terms in the radial Poynting vector.


ωt = 0.0000*π
Standing Wave Terms
kr
+cos2(ωt)cos2(kr+ζ) / (kr)3
-cos2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)cos2(kr+ζ) / (kr)3

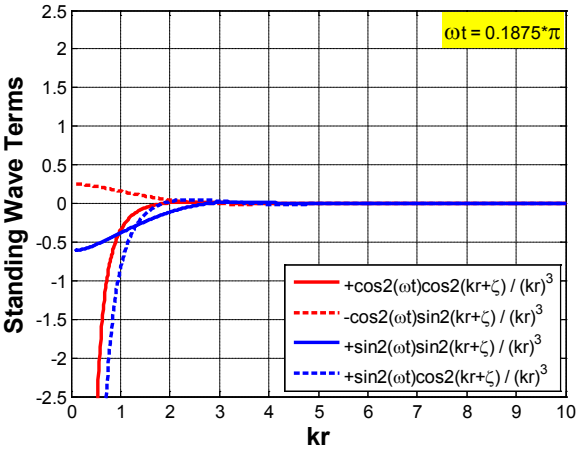
(a)

ωt = 0.0625*π
Standing Wave Terms
kr
+cos2(ωt)cos2(kr+ζ) / (kr)3
-cos2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)cos2(kr+ζ) / (kr)3

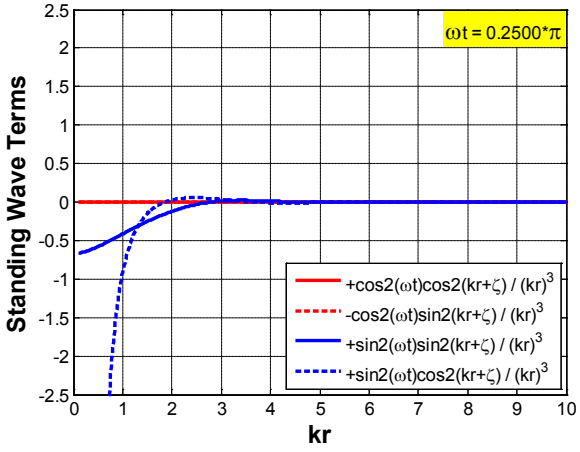
(b)


ωt = 0.1250*π
Standing Wave Terms
kr
+cos2(ωt)cos2(kr+ζ) / (kr)3
-cos2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)cos2(kr+ζ) / (kr)3

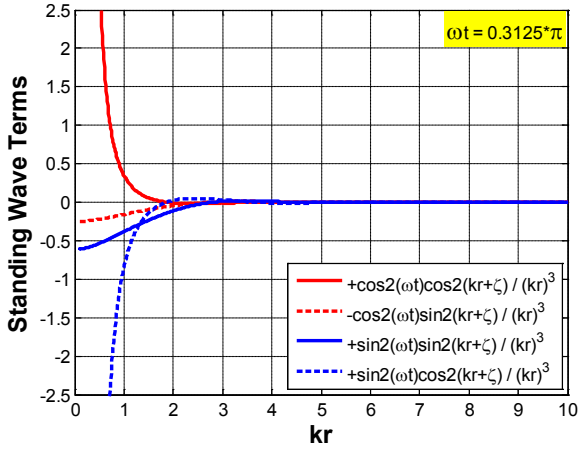
(c)


ωt = 0.1875*π
Standing Wave Terms
kr
+cos2(ωt)cos2(kr+ζ) / (kr)3
-cos2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)cos2(kr+ζ) / (kr)3

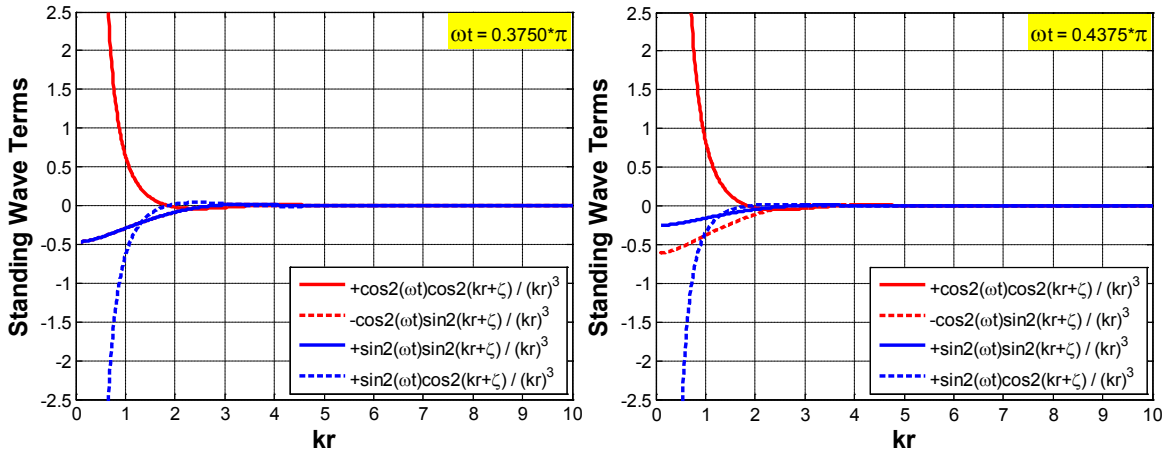
(d)


ωt = 0.2500*π
Standing Wave Terms
kr
+cos2(ωt)cos2(kr+ζ) / (kr)3
-cos2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)cos2(kr+ζ) / (kr)3

(e)


ωt = 0.3125*π
Standing Wave Terms
kr
+cos2(ωt)cos2(kr+ζ) / (kr)3
-cos2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)sin2(kr+ζ) / (kr)3
+sin2(ωt)cos2(kr+ζ) / (kr)3

(f)



(g) (h)

Figure 2.16 Standing-wave terms versus *kr* for the Poynting vector of the ideal electric dipole – (a) At $t =$, (b) At $t = 1/16\ T$, (c) At $t = 2/16\ T$, (d) At $t = 3/16\ T$, (e) At $t = 4/16\ T$, (f) At $t = 5/16\ T$, (g) At $t = 6/16\ T$, and (h) At $t = 7/16\ T$.

The viewpoint separating the Poynting vector of the ideal electric dipole into traveling and standing-wave terms helps us understand some details of real and reactive power in near-field region due to the vector nature of the Poynting vector. However, it seems to be very challenging to understand the different behavior of each standing-wave term. If we try to associate both traveling and standing-wave terms with the original the field terms, it appears that all of the field terms are related to both real and reactive power phenomena. Again, it seems that field terms may not be separated to radiation and non-radiation terms.

Combination of two standing waves with tapered amplitude can be decomposed into either two standing waves or one standing wave pulse one traveling wave. Thus, grouping traveling and standing-wave terms is an arbitrary choice in terms of mathematical viewpoint, although we try to be careful with the physical meaning of the decomposed each term. However, at least, it was shown that the real and reactive power-density terms can be represented as traveling (far-field radiation and source absorption) and standing-wave (energy storage) terms. Thus, the Poynting vector viewpoint with term separation supports the power viewpoint in the previous subsection.

2.2.4 The Energy Velocity Viewpoint

Recently, Bedko presented a time-space plot for measured radiating signal of dipole antenna in near-field region [*Budko*, 2009; 2010]. Bedko's space-time plot is very similar to Figure 2.10. He claimed that he observed a superluminal phenomenon that is electromagnetic wave propagates

faster than the speed of light [*Mugnai et al.*, 2000; *Wynne*, 2002] as well as a backward-wave phenomenon. The backward-wave phenomenon may be obvious because of the energy storage process in near-field region, as we discussed in the power viewpoint of the radiation mechanism. However, the superluminal phenomenon is highly questionable and it is also contradictory to our previous observation, i.e. that energy in the near-field region travels slower than the speed of light. Thus, we discuss the energy velocity viewpoint of antenna radiation in this section.

A space-average, localized total energy velocity for the ideal electrical dipole may be defined as

$$v_{e,spatial\ average}^{Total} = \frac{P_{Total}}{U_{Total}} = \frac{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}(\vec{E}\times\vec{H})\cdot\hat{r}r^2\sin\theta d\theta d\phi}{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}(w_e+w_m)r^2\sin\theta d\theta d\phi} \tag{2.29}$$

where P_{Total} is the total power integrated over spherical surface in (2.15), w_e and w_m are electric and magnetic energy densities, and U_{Total} is the total energy density per radial distance. The total energy density per radial distance is found as

$$U_{Total} = \frac{4\pi}{3}\frac{\eta_0}{2c}\left\{2+\frac{2}{k^2r^2}+\frac{3}{k^4r^4}+\left[2-\frac{2}{k^2r^2+1}-\frac{3}{k^4r^4}\right]\cos 2\left(\omega t-kr-\zeta\right) + \frac{2}{kr\left(k^2r^2+1\right)}\sin 2\left(\omega t-kr-\zeta\right)\right\} \tag{2.30}$$

where μ_0 is the permeability of free space, c is the speed of light, and ζ is the excess delay. The defined space-average, localized total energy velocity is not in vector format. Thus, it may be more appropriate to refer it as speed. However, at least we can identify the direction of the energy velocity from its sign, i.e. inward or outward direction. In a time-space plot, we can use the direction of the velocity vector for showing whether the energy velocity exceeds the speed of light or not. The definitions for inward and outward directions and the speed clarifications are shown in Figure 2.17. If the velocity vector is parallel to ωt axis, it means that the speed of the energy is zero. If the energy velocity slope is less than 45 degrees from the ωt axis, the speed of the energy is slower than the speed of light.

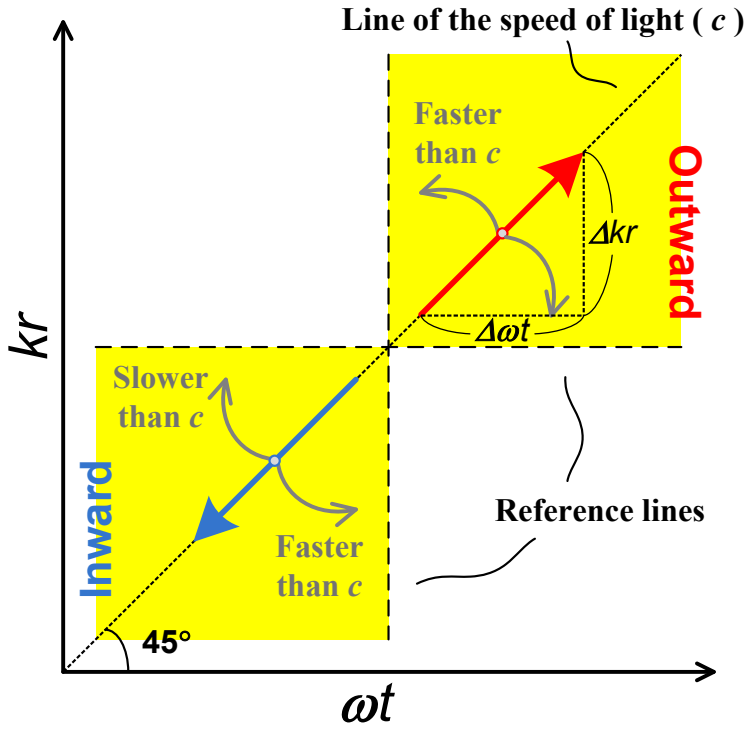
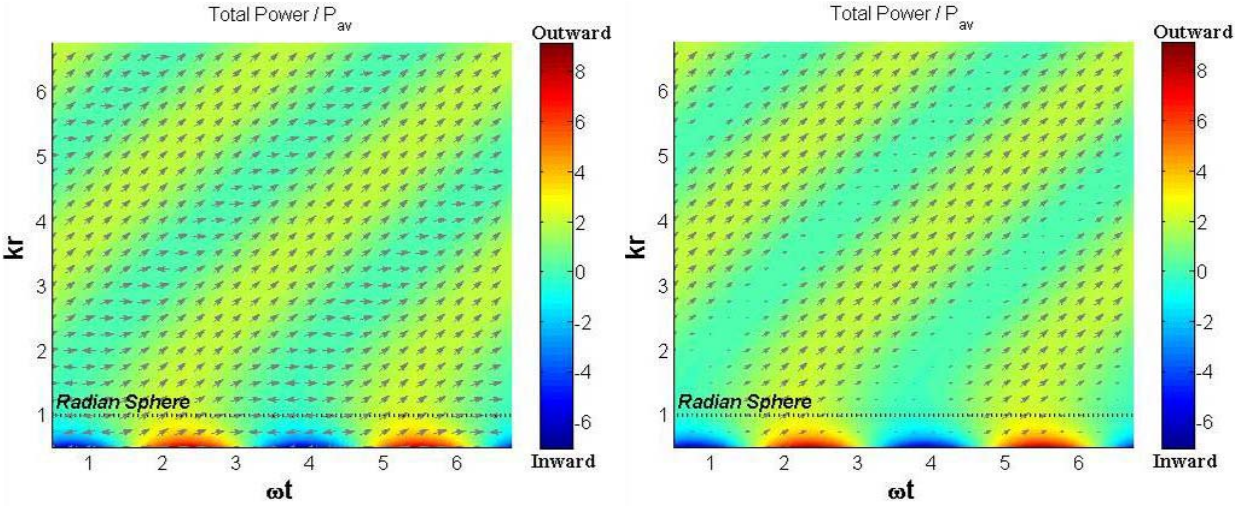


Figure 2.17 Reference directions and speed clarifications of speed for the space-average localized energy velocity.

The spatial-average, local energy velocity is shown in Figure 2.18 on top of the total power plot for the ideal electric dipole antenna. In Figure 2.18a, the arrow direction of the energy velocity represents the relative speed compared to the speed of light as defined in Figure 2.17. In the light yellow-color region indicating the real-power flow, the energy-velocity arrows are 45 degrees relative to the ωt axis. Thus, the energy contributing to real-power flow propagates with the speed of light beyond the radiation sphere boundary ($kr = 1$). However, in the light blue-color region (indicating reactive-power flow or backward-power flow), the speed of the corresponding energy is slower than the speed of light or close to zero. If the arrow size is proportion to the magnitude of the energy velocity (see Figure 2.18b), the locations having zero energy velocity becomes clear. Therefore, the radiating energy does not propagate faster than the speed of light and also does not propagate with a constant speed.



(a) (b)

Figure 2.18 Time-space plot of total power with spatial-average localized energy velocity for ideal electric dipole – (a) Energy-velocity vectors with fixed arrow size and (b) Energy-velocity vectors with scaled arrow size corresponding to the magnitude of the velocity [*Davis et al.*, 2011], copyright © 2011 IET.

In order to include the complete vector information, a localized total energy velocity may be defined as

$$\vec{v}_e^{\,Total} = \frac{\vec{S}}{w_{Total}} = \frac{\vec{E}\times\vec{H}}{w_e + w_m} \tag{2.31}$$

where $\vec{S}$ is the Poynting vector composed of (2.27) and (2.28) and w_{Total} is the total energy density. The total energy density is found as

$$\begin{aligned} w_{Total} = \frac{\mu_0}{2}k^2\Bigg\{ & 4\cos^2\theta\frac{k^2r^2+1}{k^6r^6}\sin^2\left(\omega t-kr-\zeta\right) \\ & +\sin^2\theta\frac{1}{k^2r^2+1}\left[\cos^2\left(\omega t-kr-\zeta\right)+\frac{\sin 2\left(\alpha-\zeta\right)}{k^3r^3}+\frac{1}{k^6r^6}\sin^2\left(\omega t-kr-\zeta\right)\right] \\ & +\sin^2\theta\frac{k^2r^2+1}{k^4r^4}\cos^2\left(\omega t-kr-\zeta\right)\Bigg\} \end{aligned} \tag{2.32}$$

In order to identify the bound on the localized total energy velocity, we take the square of the total energy velocity and divide it by square of the speed of light:

$$\frac{\vec{v}_e^{\,Total}\cdot\vec{v}_e^{\,Total}}{c^2} = 4c^2{\mu_0}^2\frac{\left|\vec{E}\right|^2\left|\vec{H}\right|^2-\left(\vec{E}\cdot\vec{H}\right)^2}{\left(\left|\vec{E}\right|^2+c^2{\mu_0}^2\left|\vec{H}\right|^2\right)^2} \tag{2.33}$$

Where c is the speed of light and μ_0 is the permeability of free space. In order to compare with the speed of light, consider the following.

$$1-\frac{\left|\vec{v}_e^{\,Total}\right|^2}{c^2}=\frac{\left(\left|\vec{E}\right|^2-c^2\mu_0{}^2\left|\vec{H}\right|^2\right)^2+4c^2\mu_0{}^2\left(\vec{E}\cdot\left|\vec{H}\right|\right)^2}{\left(\left|\vec{E}\right|^2+c^2\mu_0{}^2\left|\vec{H}\right|^2\right)^2}\ \geq\ 0 \tag{2.34}$$

Because the magnitude of the localized total energy velocity should be greater than or equal to zero, the bound of the localized total energy velocity is given as

$$0\leq\ \left|\vec{v}_e^{\,Total}\right|\ \leq c \tag{2.35}$$

The bound of the energy velocity is valid for any electromagnetic wave even for the cases of time and/or spatial average.

The computed localized total energy velocity for the ideal electric dipole (spherical TM_{01} mode) is shown in Figure 2.19 at some selected times. The dotted white circle line indicates the boundary of the radian sphere (kr = 1). Until the quarter of radiation process period, we can clearly observe the dominant energy storage phenomenon inside and around the radian sphere (see Figure 2.19a and Figure 2.19b). As discussed in the previous section, the quarter period of the radiation corresponds to the time that charges are located at the center of the dipole antenna. After the quarter period, the electric fields start to decouple from the antenna (see Figure 2.4). Thus, the far-field radiation process starts becoming the dominant phenomenon after the quarter period, as shown in Figure 2.19c and Figure 2.19d. Both the radiation process and the energy storage process continue during the entire radiation period. However, energy storage process dominates during the first quarter period while far-field radiation dominates after a half period of the radiation.

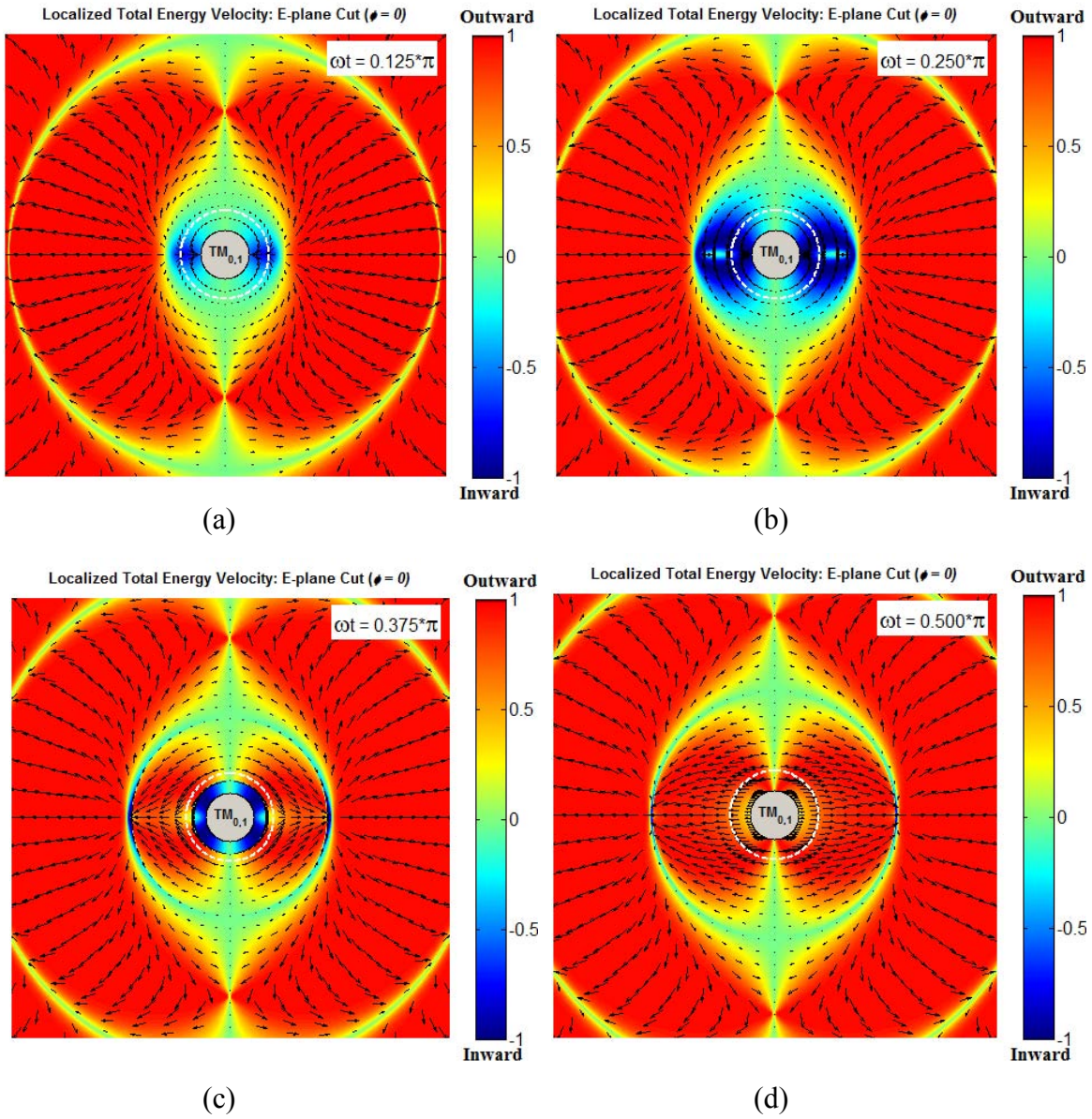
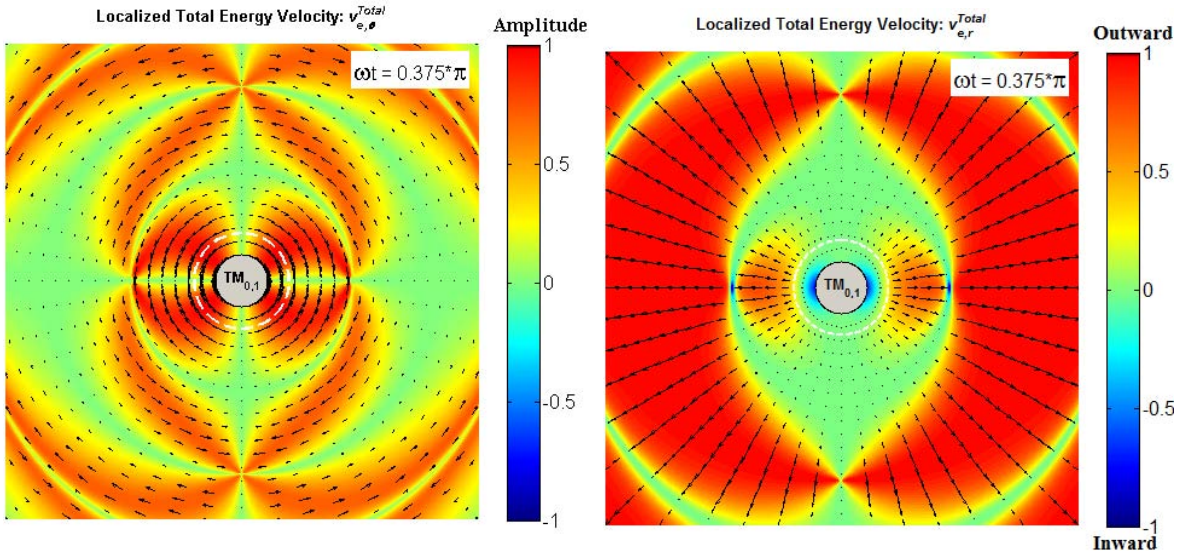


Figure 2.19 Localized total energy velocity for ideal electric dipole electric field lines – (a) At t = 1/8 period, (b) At t = 2/8 period, (c) At t = 3/8 period, and (d) At t = 4/8 period. The dotted circle represents the radian sphere criteria (kr = 1).

The localized total energy velocity, decomposed into angular and radial vectors, is shown in Figure 2.20 at t = 3/8 T. The angular component of the energy velocity is due to E_r and H_ϕ fields while the radial vector is due to E_θ and H_ϕ fields. Since the angular localized total energy velocity is circulating, it is neither forward nor backward. Thus, the color in Figure 2.20a only indicates the magnitude of the angular localized total energy velocity. An interesting observation

is that the angular localized total energy velocity exists in entire space, indicating that E_θ is not the only field component contributing to far-field radiation. On the other hand, the radial component of the energy velocity in Figure 2.20a shows an energy storage process inside the radian sphere and far-field radiation process outside the radian sphere. Notice that there exist backward waves around approximately double the distance of the radian sphere (kr ≈ 2) Figure 2.20b, which is a very interesting phenomenon that the classical radiation models cannot easily explain. This phenomenon can be explained with the causal surfaces suppressing the net energy flow [*H G Schantz*, 2001].



(a) (b)

Figure 2.20 Decomposed localized total energy velocity for ideal electric dipole electric field lines at $t = 3/8\ T$ – (a) Angular localized total energy velocity and (b) Radial localized total energy velocity.

Recall the Poynting vector for the ideal electric dipole antenna, i.e.

$$\vec{S} = \vec{E} \times \vec{H} = \hat{r}S_r + \hat{\theta}S_\theta = \hat{r}E_\theta H_\phi - \hat{\theta}E_r H_\phi \tag{2.36}$$

There is no net energy flow if the angular magnetic field of the ideal electric dipole (H_ϕ) is zero because Poynting vector becomes zero. The time when there is zero energy velocity is found as

$$\omega t\big|_{\vec{v}_e^{Total}=0} = kr + + \frac{\pi}{2} + \tan^{-1}\left(\frac{1}{kr}\right) + n\pi \tag{2.37}$$

where n is integer. Zero energy velocity does not mean that the radiation is blocked at the boundary. Since the localized total energy velocity is a net quantity, zero energy velocity means

that both outward and inward energies have the same amplitude but with opposite directions. As an example, the boundary of the zero energy velocity at t = 0.26 T is shown in Figure 2.21. We observe that the boundary for zero energy velocity is located at the interface between the outward and inward energy velocities. Far-field radiation is dominant outside the boundary circle while energy-storage process is dominant inside the boundary circle. Thus, it appears that the boundary pushes the radiating energies outward in the radiation direction as time increases.

On the other hand, the radial vector component of the localized total energy velocity ($v_{e,r}^{Total}$) becomes zero if the angular electric field of the ideal electric dipole (E_θ) is zero. The time frame giving the zero radial energy flow is found as

$$\omega t\big|_{v_{e,r}^{Total}=0} = kr + \frac{\pi}{2} + \tan^{-1}\left(\frac{kr}{k^2r^2-1}\right) + n\pi \tag{2.38}$$

where n is integer. The boundary circle for $v_{e,r}^{Total} = 0$ is located inside the boundary circle for $\vec{v}_e^{Total} = 0$, as shown in Figure 2.21, meaning that the boundary circle for $v_{e,r}^{Total} = 0$ is advanced in time. These radiation boundaries are moving in time with different speeds. In the far-field, the arctangent terms in (2.37) and (2.38) become insignificant and both boundaries overlap, which is a similar phenomenon with what we observed in the phase response of H_ϕ and E_θ in the previous subsection. However, there is an additional boundary in between the boundaries of $\vec{v}_e^{Total} = 0$ and $v_{e,r}^{Total} = 0$. The additional boundary seems to be caused by reflected energy at the boundaries of $\vec{v}_e^{Total} = 0$ and $v_{e,r}^{Total} = 0$. Schantz presented these boundaries first and named them as causal surfaces [*H G Schantz*, 2001].

The energy velocity viewpoint of antenna radiation enabled us to observe two additional radiation boundaries (causal surfaces), partially explaining the energy conversions and interactions between radiation and non-radiating energies in near-field region. A transient analysis [*G.S. Smith*, 2002; *G.S. Smith and Hertel*, 2001] may provide additional useful observations in order to understand the near-field energy radiation and storage mechanism better.

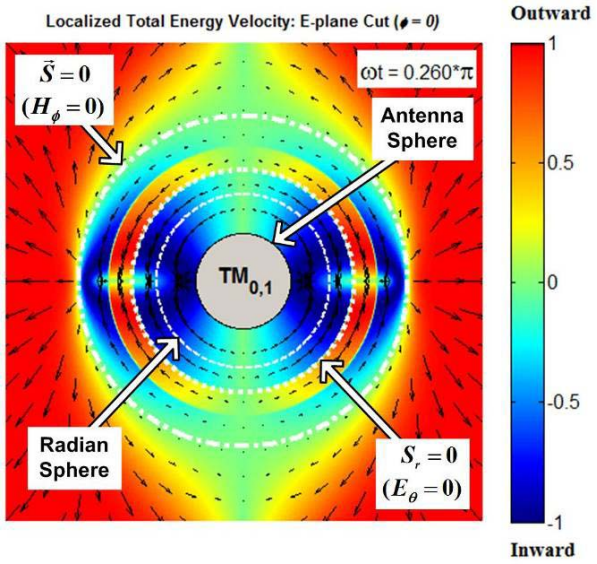


Figure 2.21 Illustration of radiation boundaries for ideal electric dipole antenna at t = 0.26 period.

2.3 Chapter Summary

Two classical models for antenna radiation were reviewed, the single and dual charge models. The single charge model visually explains how radiation fields are generated from accelerated charges without using electromagnetic equations, based on two fundamental physical principles, i.e. the finite propagation speed of the fields and the continuity of field lines. However, the single charge model does not explain how the radiating fields can be decoupled from antenna. Compared to the single charge model, the dual charge model (dipole model) provides an explanation for the field-decoupling process. However, an energy storage mechanism was not included in the dual charge model. These classical viewpoints over-emphasize accelerated charges as the cause of the antenna radiation and missed the overall picture of the antenna radiation mechanism including energy storage process.

In order to understand the antenna radiation phenomenon better, alternative viewpoints were discussed by investigating wave impedance, power, Poynting vector, and energy velocity of ideal dipole antenna. The investigation suggested that all of the antenna fields contribute to both radiation and energy storage. In addition, the energy velocity of radiating energy is not constant. These observations are contradictory to the well-known facts, i.e. only angular fields with $1/r$ variation contribute to far-field radiation and the propagation speed of fields in free space is

constant. However, we provided various evidences and proofs that these widely-accepted facts are not always true.

Although we could understand the radiation and energy storage mechanism better with the alternative viewpoints, it turns out that the radiation process in the near fields is very complicated due to the virtual boundaries, i.e. radian sphere and causal surfaces. These boundaries appear to cause energy reflections and transmissions. In particular, the causal surfaces move over time, which creates another layer of difficulty in understanding radiation process details. Quantum electrodynamics [*Feynman*, 1961] is an additional tool to investigate the antenna radiation mechanism in depth, but it is out of our study scope.

A simple, high-level illustration depicted in Figure 2.22 sums up the observations from the classical and alternative viewpoints on antenna radiation. The *antenna sphere* is the smallest sphere which encloses the antenna structure. Some of the energy provided by the antenna structure remains in the vicinity of the antenna as non-radiating energy, not contributing directly to the radiation process. This non-radiating stored energy dominates inside the radian sphere of radius $1/k$, where k is wave number. The non-radiating stored energy is trapped and circulates between the antenna and approximately the radian sphere. The circulation of the non-radiating energy occurs in both radial and angular directions. On the other hand, the radiation process dominants outside the radian sphere. A periodic energy leak occurs from the radian sphere to produce the far-field radiation. Some of radiating energy comes back to antenna source, gets absorbed in antenna source, and reradiate later. This is the process of converting reactive energy to radiated energy.

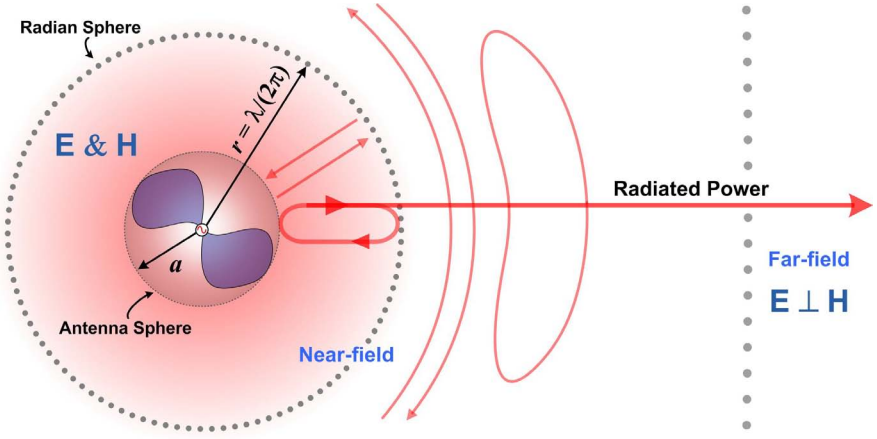


Figure 2.22 High-level illustration of antenna radiation and energy-storage mechanism.

Chapter 3

Antenna Physical Structure and Transfer Function

The investigation of antenna radiation and the energy storage mechanism in the previous chapter provided fundamental physical insights about how energy decouples from an antenna and contributes to far-field radiation through multiple virtual boundaries including the radian sphere and causal surfaces. The fundamental physical insights are useful for understanding some challenging antenna interaction and communication link problems. However, the investigated viewpoints of the antenna radiation mechanism do not directly provide information about designing antennas to meet specifications for real-world antenna design problems.

In this chapter, an antenna transfer-function model, based on both the singularity expansion method (SEM) and a spherical wave expansion (SWE), is introduced as a tool to understand the correlation between physical parameters and antenna performance.

3.1 Antenna Transfer Function Modeling

In practice, an antenna design is obtained from computational or experimental iterative steps. During the iterative design process, a typical objective is to determine how much power is delivered to the antenna (reduced by return loss) and to the far-field region (maximized range or radiation efficiency) over a specified operational frequency range (bandwidth) and angular range (radiation pattern). Thus, in the practical design process, the antenna has been treated as a black box located between a radio-frequency (RF) block and the far field. The black box can be seen as

an antenna transfer function. For transmission, the input of the antenna transfer function is an input signal that comes from the RF circuit in the radio. The output can be radiating fields in the frequency domain or radiating pulses in the time domain. Alternatively, the output can target performance characterization parameters, including total realized gain and radiation pattern. If the antenna transfer function can be treated as a linear, time-invariant system, the output can be easily evaluated from the frequency-domain product or the time-domain convolution of the input signal and the antenna transfer function [*Oppenheim and Schafer*, 2010; *Oppenheim et al.*, 1997]. Thus, the transfer function can be obtained by matrix inversion or a de-convolution process if the input and output functions are known.

If the antenna transfer function is treated as a true black box, the ultimate interest for the transfer function may be how accurately it describes the output. In this sense, the antenna transfer function can be modeled through curve-fitting techniques [*Arlinghaus*, 1994; *Daniel et al.*, 1999; *Tarter and Lock*, 1993]. However, this type of transfer function model often involves higher-order mathematical functions. In addition, the transfer function model is not necessary to provide physical meaning. On the other hand, if the transfer function model is based on first principles in electromagnetics, the model itself offers some insight into how the physical antenna parameters relate to the performance parameters, i.e. correlation between antenna physical structure and performance change. Thus, the models based on first principles in electromagnetics provide insight on how to modify the antenna structure to accommodate changes in the original antenna platform. The models based on first principles in electromagnetics are described with reduced-order mathematical expression, compared to the blind curve fitting method. Thus, the reduced-order model is also useful for predicting overall system performance, including realistic antenna characteristics without heavy computation.

There are various ways to obtain the transfer-function model based on first principles in electromagnetics [*E. K. Miller*, 1998a; b; c], but one popular method is based on the singularity expansion method (SEM) [*C.E. Baum*, 1971; 1976]. SEM-based models have been successfully used in modeling ultra-wideband antenna transfer functions [*Licul*, 2004] and channel characteristics [*Joshi*, 2006]. Thus, we review the SEM-based antenna transfer-function model in the following subsections.

3.1.1 SEM-Based Modeling Approach

The essence of the singularity expansion method (SEM) [*C.E. Baum*, 1971] is that a characteristic waveform can be modeled with a sum of damped sinusoidal waves corresponding to complex-frequency resonances (poles) and the relative angular contributions to the waveform (residues), i.e.

$$x(t) = \sum_{i=1}^{M} R_i e^{s_i t} \Leftrightarrow X(\omega) = \sum_{i=1}^{M} \frac{R_i}{s - s_i} \tag{3.1}$$

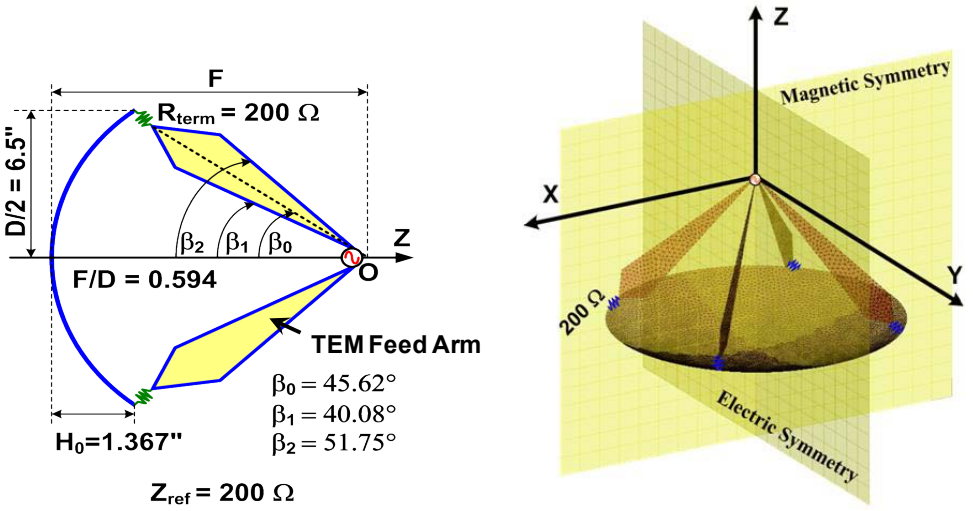
where $s_i = \sigma_i + j\omega_i$ represents the complex poles, σ_i is damping factor, ω_i is angular frequency, and R_i is the angle-dependant residue. For an exact representation of the response, an infinite sum of the complex exponentials is required. In practice, however, the response can be modeled accurately using only a few dominant complex poles due to the band-limited nature of typical sources, i.e. not all the poles of the antenna are excited strongly. The extra poles at some point become modeling poles for all the information that has been truncated.

Various methods are available to obtain the complex poles and the corresponding residues in a SEM-based model, such as frequency-domain search [*Tesche*, 1973], Prony [*Prony*, 1795; *Van Blaricum and Mittra*, 1975; *Weiss and McDonough*, 1963], and matrix pencil [*Sarkar and Pereira*, 1995] methods. Among these methods, the matrix pencil method has been known to be computationally efficient and effective even in the presence of noise [*Hua and Sarkar*, 1990; *Sarkar and Pereira*, 1995].

In order to demonstrate the concept of the SEM-based model, consider the impulse radiating antenna (IRA) [*C.E. Baum et al.*, 1999; *Kangwook and Scott*, 2005; *Manteghi and Rahmat-Samii*, 2006] in Figure 3.1. Typically, simulation of an IRA is very challenging because IRA is electrically large. In order to reduce the mesh and computer memory required in simulation, we applied both electric and magnetic symmetries (see Figure 3.1b) so that we can simulate a quarter of the original antenna size.

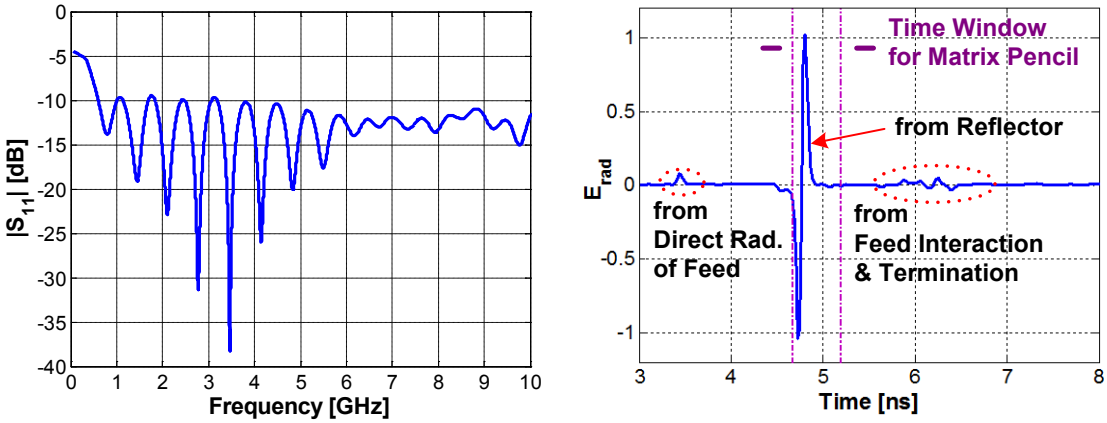
Using a commercial method of moment code [*FEKO*, 2007], the return loss (Figure 3.2a) and the radiated electric pulse (Figure 3.2b) were calculated. The IRA design provides an impedance match from 500 MHz to 10 GHz and shows some mismatch below 500 MHz, even with the resistive loads at the end of feeds. The minor impedance mismatch could be improved with a different value of termination resistor or surface resistance that provides a smooth resistance

change. On the other hand, the radiated electric pulse has glitches around 3.4 ns and 6.3 ns. These are caused from the direct radiation of feed arms and/or reflections of mismatched termination resistors.



(a) (b)

Figure 3.1 An example design of impulse radiating antenna (IRA) – (a) Antenna geometry and (b) Simulation model using commercial method of moments (MoM) code [*FEKO*, 2007].

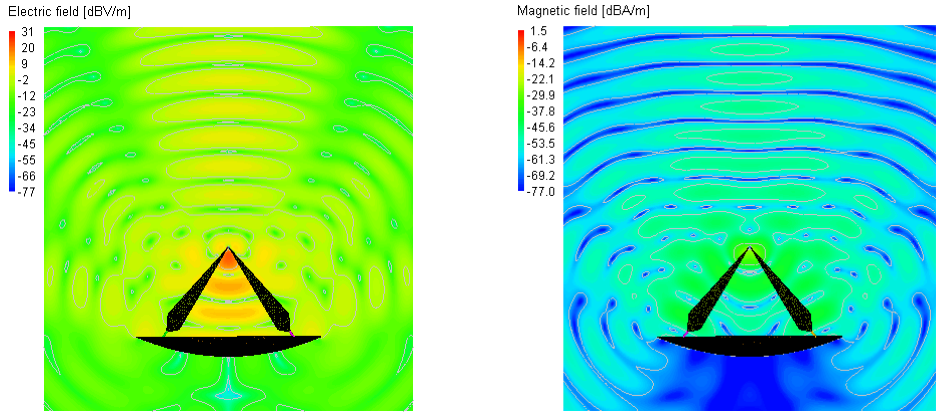


(a) (b)

Figure 3.2 Simulated performance of designed impulse radiating antenna (IRA) – (a) Return loss versus frequency and (b) Radiated electric pulse at boresight versus time. Simulation results were obtained with a commercial method of moments (MoM) code [FEKO, 2007].

In order to observe whether the IRA generates plane waves beyond the focal point, the instantaneous electric and magnetic fields were calculated at 3 GHz. As shown in Figure 3.3, it is

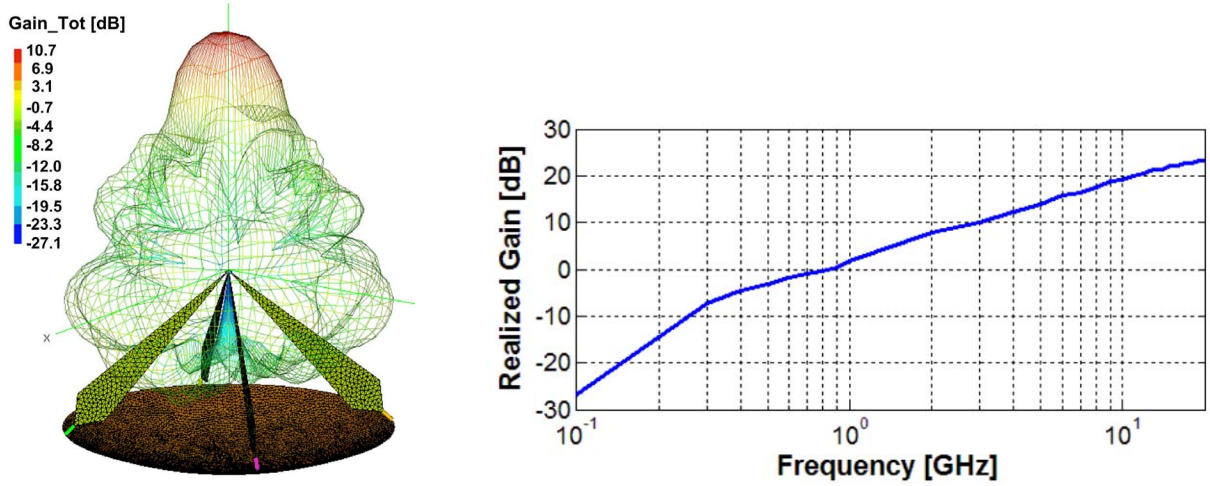
apparent that the reflected waves become planar, but there is also some interaction caused from the feed region.



(a) (b)

Figure 3.3 Simulated instantaneous fields of impulse radiating antenna (IRA) at 3 GHz with ωt = 30, obtained by using commercial method of moments (MoM) code [*FEKO*, 2007] – (a) Electric-field distribution and (b) Magnetic-field distribution.

The radiation gain pattern was also calculated and is depicted in Figure 3.4. Basically, the IRA has maximum gain at boresight with some side lobes caused from the feed reaction. Because the IRA is a constant aperture antenna, the calculated realized gain increases as frequency increases. Thus, the designed IRA shows excellent performance in both frequency and time domains.



(a) (b)

Figure 3.4 Simulated radiation performance of the design impulse radiating antenna (IRA) obtained by using commercial method of moments (MoM) code [*FEKO*, 2007] – (a) Gain pattern at 3 GHz and (b) Realized gain versus frequency.

The radiated electric-pulse data with sampling time (Δ_t) and sample index (k) in Figure 3.2b can be modeled with in SEM representation, i.e.

$$E_{rad}(\theta,\phi,k\Delta_t) \simeq \sum_{i=1}^{M} R_i(\theta,\phi) z_i^k \tag{3.2}$$

where M is a maximum number of needed poles (a priori information), and $z_i^k = \exp(s_i k\Delta_t)$. Note that each data sample at a given angle is a summation of all the poles and residues. Therefore, one can obtain a model of the radiating electric pulse after finding the poles and residues from the time-windowed signal. Because noise is always present whether the sampled data is obtained from simulations or measurements, we consider the matrix pencil method [*Sarkar and Pereira*, 1995] for obtaining the poles and residues. Typically, noise energy is almost constant in the sampled data while the signal energy is reduced as time increases. Thus, we can consider a time-gated signal (see the time windows in Figure 3.2b) as a pre-filtering in order to reduce the noise.

We select N as the total number of data samples in the time window and L is the pencil parameter, which is related to an estimation of the maximum number of poles (M). The maximum number of poles (M) and the matrix pencil parameter (L) are typically estimated from an experience-based or trial-and-error approach. From the truncated radiated electric pulse data, we can define $(N\text{-}L) \times L$ time delayed data matrix as the following:

$$[Y]_q = \begin{bmatrix} \sum_{i=1}^{M} R_i z_i^{q} & \sum_{i=1}^{M} R_i z_i^{(q+1)} & \cdots & \sum_{i=1}^{M} R_i z_i^{(q+L-1)} \\ \sum_{i=1}^{M} R_i z_i^{(q+1)} & \sum_{i=1}^{M} R_i z_i^{(q+2)} & \cdots & \sum_{i=1}^{M} R_i z_i^{(q+L)} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{M} R_i z_i^{(q+N-L-1)} & \sum_{i=1}^{M} R_i z_i^{(q+N-L)} & \cdots & \sum_{i=1}^{M} R_i z_i^{(q+N-2)} \end{bmatrix}_{(N-L)\times L} \tag{3.3}$$

where q is a time delay index. Because eigenvalues (λ_i) of the matrix pencil ($[Y]_{q=1} - \lambda [Y]_{q=0}$) correspond to z_i of the original data, we can obtain poles easily from the eigenvalues, allowing evaluation of residues using (3.2). Generally, we have the following constraints on M, L and N:

$$M \le L \le (N - M) \tag{3.4}$$

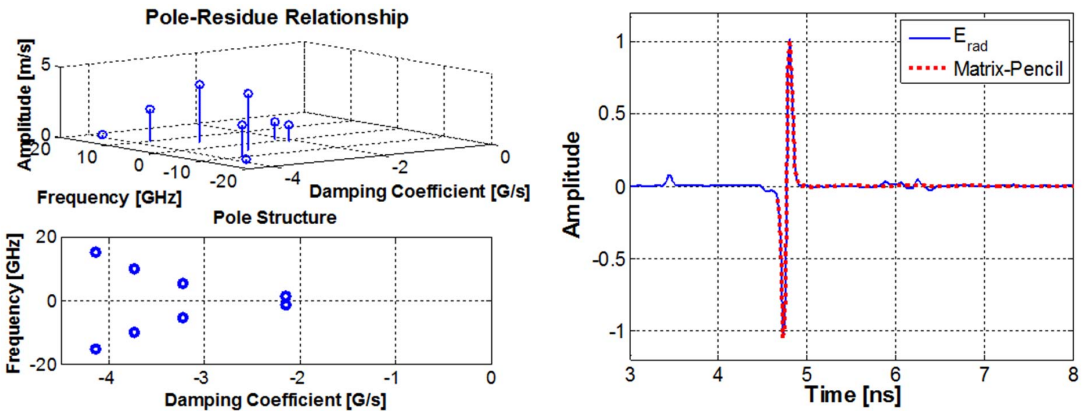
which suggests that it may not be possible to extract enough poles to reconstruct the original signal effectively unless there is a sufficient number of data samples in the truncated window. In practice, we have more data than needed, which makes the matrix pencil matrix an over-determined system. Thus, the rank of the matrix pencil matrix can be reduced using singular value decomposition (SVD) [*Golub and Van Loan*, 1996], i.e.

$$\left[E_{rad,truncated}(\theta,\phi,t)\right]_{(N-L+1)\times L} = [U]_{(N-L+1)\times M} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_M \end{bmatrix}_{M\times M} [V]^H_{M\times L} \tag{3.5}$$

where H means a conjugate transpose of a matrix. Because the small singular values might be caused from either numerical or measurement noises. We can reduce the rank of the matrix pencil matrix by eliminating the small singular values. In addition, this approach can reduce the noise effects too. Alternatively, we can eliminate poles with low energy. The energy of pole pair is proportional to the square of pole magnitude and inversely proportional to the damping factor:

$$Energy = \int_0^\infty \left(R\cdot e^{-st} + R^*\cdot e^{-s^*t}\right)^2 dt\Big|_{\omega \gg 0} \approx \frac{|R|^2}{\sigma} \tag{3.6}$$

The pole-residue relationship of the SEM model of the IRA is shown in Figure 3.5a. The reconstructed radiating electric pulse using the obtained SEM-based model is compared to the original pulse in Figure 3.5b. We can clearly observe that the restored pulse almost overlaps the original pulse. The glitches due to direct feed radiation and feed interaction are not seen in the restored pulse because of the time gating, which demonstrates the capability of noise reduction.



(a) (b)

Figure 3.5 SEM-based model of the designed impulse radiating antenna (IRA) and its performance – (a) Pole-residue relationship of the SEM-based IRA model and (b) Comparison of radiated E-field response and the restored response from matrix pencil method.

3.1.2 SEM-Based Model of Transfer Function for Canonical Antennas

Licul [*Licul*, 2004] showed that a SEM-based model of antenna effective height provides a clear and unified antenna description to characterize the antenna properties in both the frequency and time domains, as summarized in Figure 3.6.

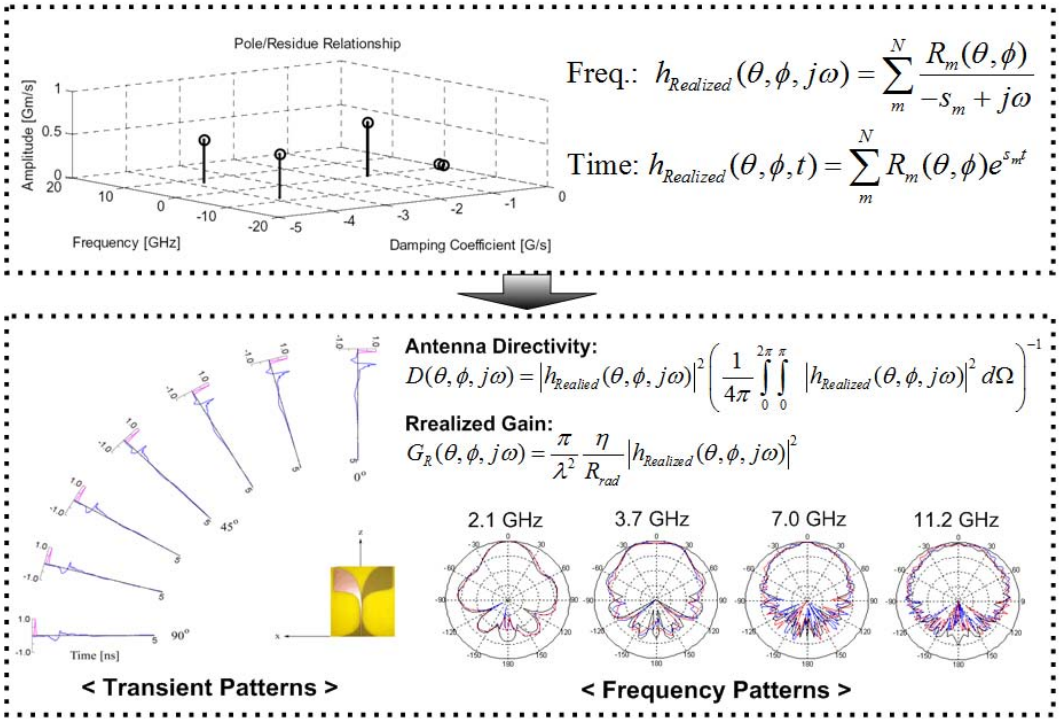


Figure 3.6 Antenna transfer function based on singularity expansion method [*Licul*, 2004].

Traditionally, *effective height* ($\vec{h}$) of an antenna is defined in terms of a receiving system, i.e. [*Stutzman and Thiele*, 2012]

$$V_{oc} = \vec{h}^{*}(\theta,\phi)\cdot\vec{E}_{inc}(\theta,\phi) \tag{3.7}$$

where $\vec{E}_{inc}$ is the incoming electric field and V_{oc} is the open circuit voltage in the receiving system. Alternatively, the effective height of a transmitting antenna can be related to the radiating electric field ($\vec{E}_{rad}$) as the following [*Licul*, 2004]:

$$\vec{E}_{rad}(\theta,\phi) = j\omega\mu_0\Phi_{FF}I_t\vec{h}^{*}(\theta,\phi) \tag{3.8}$$

where μ_0 is the permeability of free space, Φ is the free-space green's function in far-field form and I_t is the input current at the transmitting antenna terminal. Thus, the effective height of the antenna can be considered as an antenna transfer function giving the output of the radiating electric field along with the far-field free-space Green's function ($e^{-jkr}/4\pi r$). A similar approach was presented by Shlivinski (see Chapter 6 for details) [*Shlivinski et al.*, 1997].

The far-field realized effective height of antenna ($\vec{h}_{realized,FF}$) including a return loss can be measured with a communication link. The circuit description of the communication link is depicted in Figure 3.7, where V_s is a source voltage, Z_0 is a reference impedance, Z_A is an antenna input impedance, and I_t is the current into the antenna. The open circuit voltage V_{oc} is excited on the receiving antenna terminal from the radiated electric field of the transmitting system. If two identical antennas are used for both transmitting and receiving systems, the far-field realized effective height of antenna ($h_{realized,FF}$) is found from [*Licul*, 2004]

$$h_{realized,FF}(j\omega) = \sqrt{\frac{\sqrt{R_t R_r}}{j\omega\mu_0}2\pi r e^{jkr}s_{21}(j\omega)} \tag{3.9}$$

where s_{21} is the scattering parameter [*Pozar*, 2012] representing a link response.

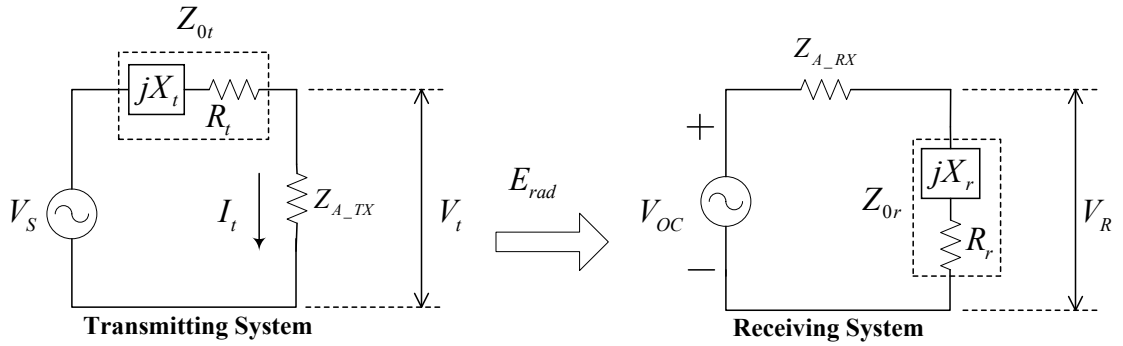


Figure 3.7 Circuit description of communication link in a line of sight.

As an example, the link response (s_{21}) using two identical planar half-disk antennas [*Yang and Davis*, 2004] was measured using an HP8150 vector network analyzer and the data are shown in Figure 3.8a. It appears that the link response curve is contaminated by noticeable noise above approximately 14 GHz. The frequency link response of the half-disk antenna is relatively flat over 3.1 – 10.6 GHz, suggesting that the half-disk antenna is a good candidate as an ultra-wideband antenna. The dipole radiation mode is dominant at the lowest operating frequency while dual-slot radiation mode becomes dominant as operating frequency increases. The magnitude of the evaluated far-field realized effective height using (3.9) is shown in Figure 3.8a in the frequency domain. The time-domain far-field realized effective height in Figure 3.8b was obtained from the inverse Fourier transform of evaluated effective height in frequency domain.

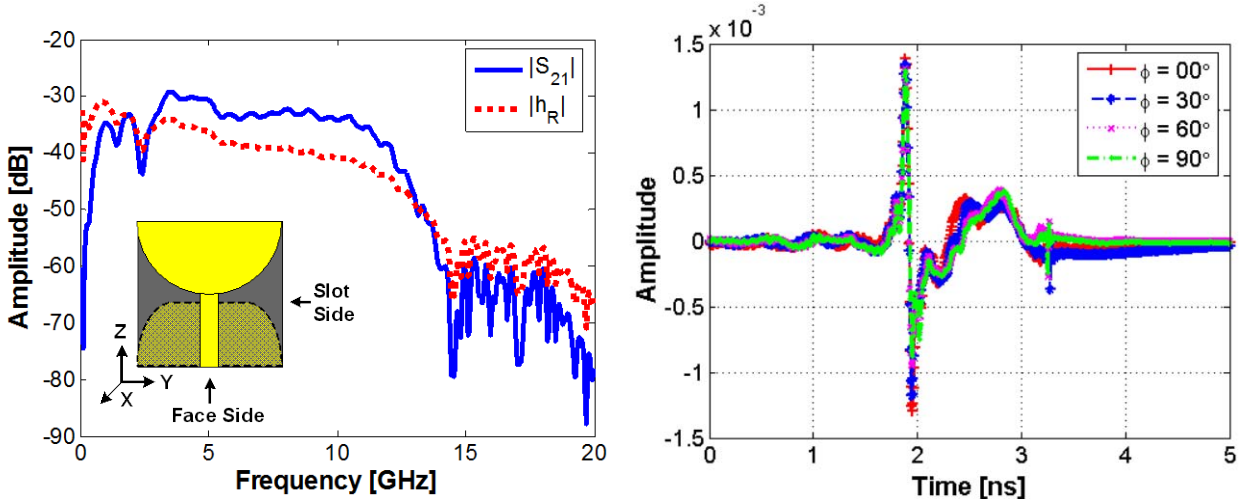


Figure 3.8 Evaluated far-field realized effective height of the planar half-disk antenna – (a) Link response and far-field realized effective height at $\phi = 0°$ in frequency domain and (b) Far-field realized effective height at selected azimuth angles in time domain.

The SEM-based model of the far-field realized effective height for the planar half-disk antenna was obtained from the matrix pencil approach. The pole-residue relationship of the SEM-based model is shown in Figure 3.9a. The comparisons between the original and reconstructed $h_{realized,FF}$ in time (Figure 3.9b) and frequency (Figure 3.10) show the validity of the model. If we replace the noisy data above 14 GHz with asymptote-curve-fitted data, we might obtain a SEM-based model with a reduced number of poles.

One pair of poles located around 1 GHz seems to dominate the late response of minor ringing. Poles located around 3 GHz may be related to the dipole-mode radiation operation of the planar half-disk. Other high-energy poles may be related to the slot-mode radiation operation. In fact, we can investigate the SEM-based antenna transfer function and corresponding radiation performance by adding extra poles and/or removing existing poles in the model. This allows us to identify correlations between the model, the physical structure, and the performance. The investigation to obtain the correlation employs an iterative process, but certainly helps to understand antenna design.

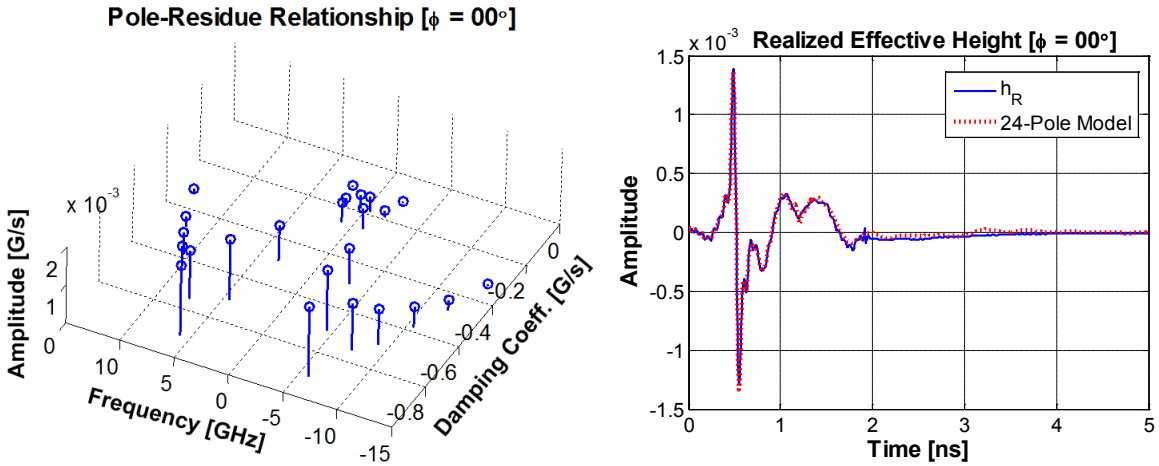
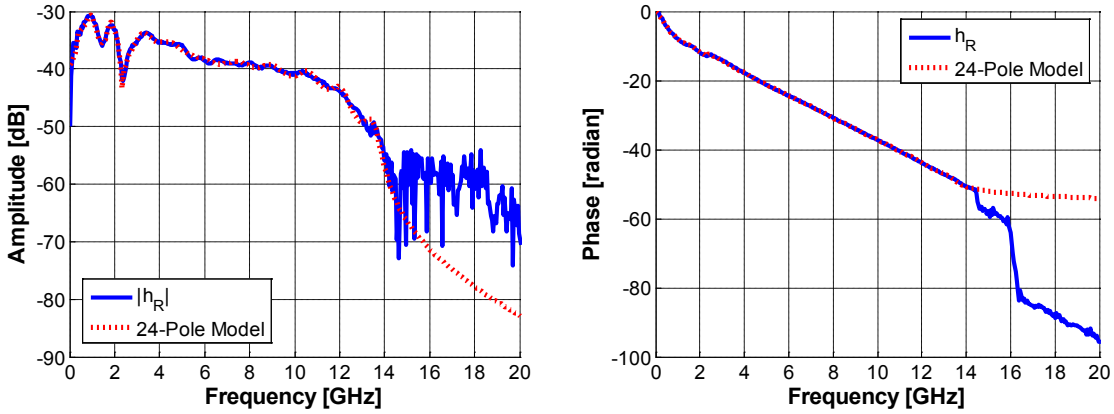


Figure 3.9 SEM-based model of planar half-disk antenna (a) Pole-residue relationship of SEM-based model of far-field realized effective height and (b) Comparison between the original and reconstructed far-field realized effective height in time domain.



(a) (b)

Figure 3.10 Far-field realized effective height of planar half-disk antenna in frequency domain (a) Amplitude response comparison between original and reconstructed $h_{realized,FF}$ and (b) Phase response comparison between original and reconstructed $h_{realized,FF}$.

Pole-residue representations of SEM-based transfer function models for various antenna types are shown in Figure 3.11. The resonant antenna in Figure 3.11a has a pair of poles with a lowest damping coefficient and a highest residue at a low frequency among the three dominant pairs of poles, which indicates that the poles have the highest energy and dominate the time-domain response, leading to a ringing phenomenon. We also note that there are many poles other than the major three pairs of poles. These pairs might come from higher-order spherical mode excitation due to the antenna excitation source with extremely wide frequency range, but it is not straightforward to identify exactly where these poles come from. Some of these poles cannot be easily filtered with energy criteria or time gating because their energy is not negligible. The frequency-independent cavity-backed spiral antenna in Figure 3.11b has a series of poles over a wide frequency range. Each pole pair is probably related to the resonant frequency of the active region on the spiral antenna. The ultra-wideband Vivaldi (exponentially-tapered slot) in Figure 3.11c and the planar half-disk antenna in Figure 3.11d have two distinct groups of poles. One of the groups clusters at low operational frequencies and has low residues and damping coefficients, which may partially explain why the Vivaldi and the planar half-disk provide a radiating pulse with a narrow pulse width and reduced ringing (see Chapter 6 for time-domain response). Thus, different types of antennas show a different pattern of pole-residue relationship. Thus, the SEM-

based model is useful not only for antenna characterization, but also for target identification [*C.E. Baum*, 2012; *C. E. Baum et al.*, 1991].

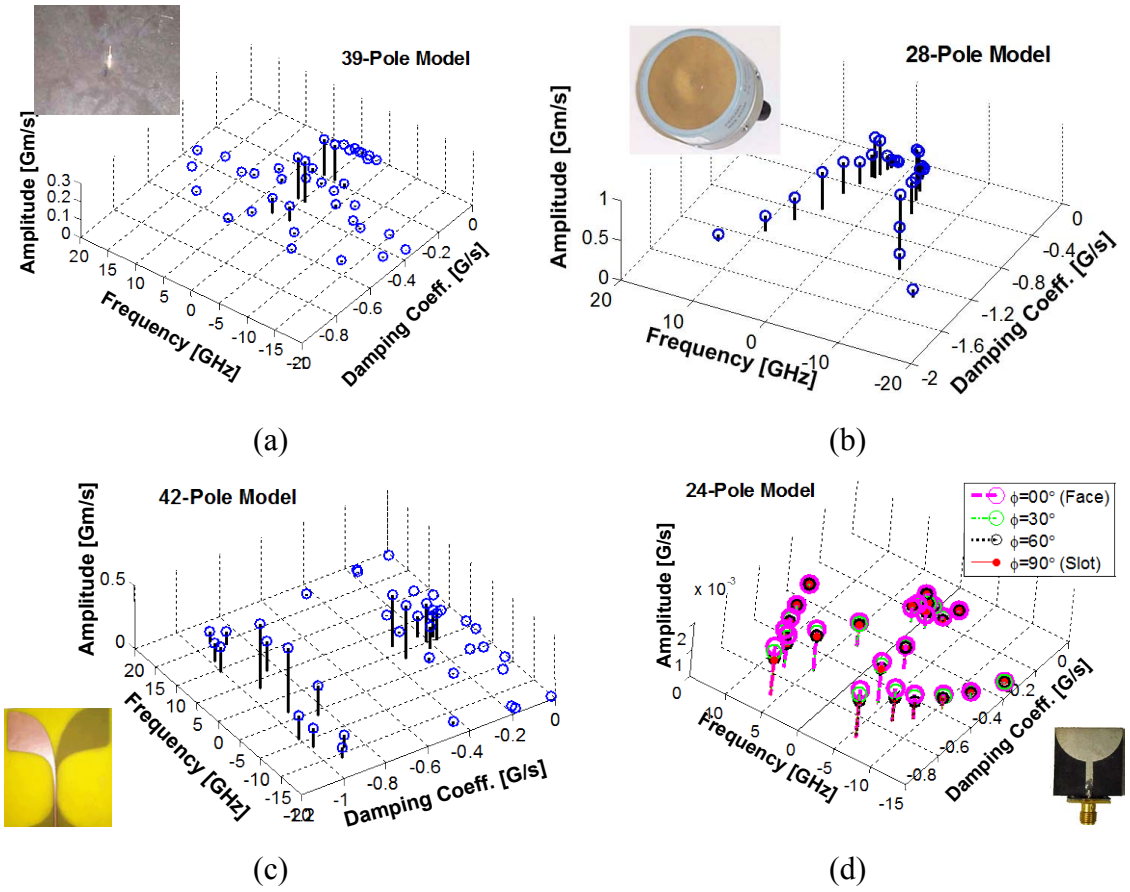


Figure 3.11 Pole-residue relationships of SEM-based transfer function for various antennas at boresight – (a) 1-GHz resonant monopole [*Licul*, 2004], (b) Frequency-independent cavity-backed spiral [*Licul*, 2004], (c) Ultra-wideband Vivaldi [*Licul*, 2004], and (c) Ultra-wideband Half disk [*Yang and Davis*, 2004].

The SEM-based transfer model using antenna effective height was found to be very useful to describe far-field antenna characteristics, as demonstrated in [*Licul*, 2004]. However, there are some practical difficulties for obtaining a model representing both near and far-field antenna characteristics if a measurement approach is used to obtain the effective height in near-field range. This is because there are potential interactions between the antennas that are not part of the antenna transfer function. In addition, observable poles through the line-of-sight measurement approach may be limited. Technically, sampled time domain data at any given

angle is a sum of all the poles. However, some poles might not be excited strongly at a given angle (i.e. small residue), as demonstrated in [*Taeyoung Yang et al.*, 2009]. This issue becomes more challenging in practice due to noisy data if the model needs to be valid for all angles, which suggests that measurement data should be collected over various angles.

3.2 The Decomposed SEM-based Antenna Transfer Function

With this study, we would like to understand the radiation mechanism, investigate theoretical limits on antenna size and performance, and find potential solutions for near-field interaction problems and challenging compact antenna design problems, as outlined in Chapter 1. If there is an antenna transfer function model to support the investigation process, the model can be a good tool to obtain integrated knowledge and backgrounds for what we observe in the study. Thus, preferred features for the model of an antenna transfer function are summarized as:

1. Easy modeling process through either simulations or measurements
2. Identifiable correlation between antenna physical structure and performance parameters
3. Compact description in both frequency and time domains
4. Clean connection to the fundamental-limit theory on antenna size
5. Capability to explain both near and far-field characteristics of antenna

The SEM-based transfer function model using antenna effective height in the previous section would be a good candidate for our purposes. However, the model using an antenna effective height has some practical limitations in explaining near-field phenomena. In addition, it is not also clear how the model is used to investigate the fundamental-limit theory on antenna size.

Because the classical fundamental-limit theory on antenna size [*Chu*, 1948] is based on the fundamental spherical mode, it seems that the spherical wave expansion (SWE) [*Stratton*, 1941] is a good candidate tool to characterize an antenna in both near and far-field regions. The SWE technique is used to represent arbitrary fields with a summation of free-space spherical modes. The mode coefficient of each mode can be found from known fields on a spherical surface with a radius of r_0 ($r = r_0$). After the mode coefficients are known, fields beyond the surface ($r > r_0$) can be evaluated from the summation of the modes. From the known fields, antenna characterization factors including gain and radiation pattern can be easily computed. The SWE method has been successfully used in many near and far-field applications including near to far-field transformation [*Hansen*, 1988] and reflector antenna design [*Potter*, 1967].

However, the SWE method requires spherical near-field scan data to obtain the mode coupling coefficients. Although it depends on the electrical size and directivity of the antenna, a spherical near-field scan often needs extensive computing resources and long measurement time. However, the spherical near-field scan be achieved almost in real time due to the recent advances in computer technology and multi-probe techniques [*Bolomey et al.*, 1988]. The technical advance leads us to consider combining the SEM and the SWE techniques in order to meet the listed preferred features of an antenna transfer function modeling.

In the next subsections, we review the essence of the SWE method and present SEM-based antenna transfer functions that are decomposed with SWE for various antenna types.

3.2.1 The Spherical Wave Expansion (SWE)

Several spherical wave expansion methods are popular, including integral, matrix, and Fourier transform approaches [*Hansen*, 1988]. We will consider in the following subsections the integration and matrix approaches in order to demonstrate the concept of SWE.

3.2.1.1 The Integration Approach

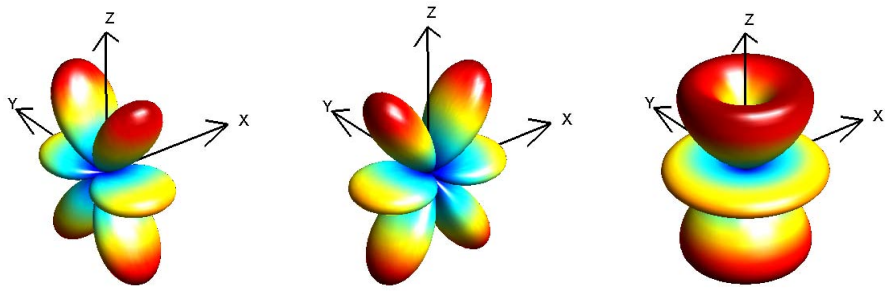
The integration approach requires two orthogonal fields (a set of E_θ and E_ϕ or a set of H_θ and H_ϕ) that are measured on a spherical surface with radius of r_0.

With $e^{j\omega t}$ time dependence, vector potentials for outward waves from an antenna can be written as [*R F Harrington*, 2001]

$$\vec{A} = \hat{r}\sum_{n=1}^{\infty}\sum_{m=-n}^{n} A_{m,n}\hat{H}_n^{(2)}(kr)P_n^{|m|}(\cos\theta)e^{jm\phi} \tag{3.10}$$

$$\vec{F} = \hat{r}\sum_{n=1}^{\infty}\sum_{m=-n}^{n} F_{m,n}\hat{H}_n^{(2)}(kr)P_n^{|m|}(\cos\theta)e^{jm\phi} \tag{3.11}$$

where $\hat{H}_n^{(2)}(kr)$ is the alternative form of the spherical Hankel function of the 2nd kind, $P_n^{|m|}(\cos\theta)$ is the associated Legendre function of the 1st kind, and A_{mn} and F_{mn} are the spherical coefficients of the expansion. Notice that we choose $e^{jm\phi}$ as a harmonic function in (3.10) and (3.11), which implies that we are interested in an antenna with omni-directional radiation pattern in azimuth. In Figure 3.12, θ and ϕ behavior for the product of $P_3^1(\cos\theta)$ and typical harmonic functions are shown.



(a) (b) (c)

Figure 3.12 Plots for product of $P_3^1(\cos\theta)$ and various harmonic functions – (a) $\left|P_3^1(\cos\theta)\sin\phi\right|$, (b) $\left|P_3^1(\cos\theta)\cos\phi\right|$ and (c) $\left|P_3^1(\cos\theta)e^{j\phi}\right|$.

From Maxwell's equations, we have [*R F Harrington*, 2001]

$$E_r = \frac{1}{j\omega\varepsilon_0}\left(\frac{\partial^2}{\partial r^2}+k^2\right)A_r\bigg|_{r=r_0} \tag{3.12}$$

$$E_\theta = \frac{-1}{r\sin\theta}\frac{\partial}{\partial\phi}F_r + \frac{1}{j\omega\varepsilon_0 r}\frac{\partial^2}{\partial r\partial\theta}A_r\bigg|_{r=r_0} \tag{3.13}$$

$$E_\phi = \frac{1}{r}\frac{\partial}{\partial\theta}F_r + \frac{1}{j\omega\varepsilon_0 r\sin\theta}\frac{\partial^2}{\partial r\partial\phi}A_r\bigg|_{r=r_0} \tag{3.14}$$

where ε_0 is the permeability of free space. Substituting the vector potentials in (3.10) and (3.11) into the above field equations gives

$$E_r(r_0,\theta,\phi) = \frac{k^2}{j\omega\varepsilon_0}\sum_{m,n}[\hat{H}_n^{(2)}(ka)+\hat{H}_n^{(2)\prime\prime}(ka)]A_{m,n}P_n^{|m|}(\cos\theta)e^{jm\phi} \tag{3.15}$$

$$\begin{aligned}E_\theta(r_0,\theta,\phi) = \sum_{m,n}[&\frac{-jm}{r_0\sin\theta}F_{m,n}\hat{H}_n^{(2)}(kr_0)P_n^{|m|}(\cos\theta)\\ &+\frac{k}{j\omega\varepsilon_0 r_0}A_{m,n}\hat{H}_n^{(2)\prime}(kr_0)\frac{d}{d\theta}P_n^{|m|}(\cos\theta)]e^{jm\phi}\end{aligned} \tag{3.16}$$

$$\begin{aligned}E_\phi(r_0,\theta,\phi) = \sum_{m,n}[&\frac{1}{r_0}F_{mn}\hat{H}_n^{(2)}(kr_0)\frac{d}{d\theta}P_n^{|m|}(\cos\theta)\\ &+\frac{jmk}{j\omega\varepsilon_0 r_0\sin\theta}A_{mn}\hat{H}_n^{(2)\prime}(kr_0)P_n^{|m|}(\cos\theta)]e^{jm\phi}\end{aligned} \tag{3.17}$$

The simplest way to obtain the mode coefficients is to use the radial field components ($A_{m,n}$ from E_r and $F_{m,n}$ from H_r). However, because the radial components of the fields decay fast, the radial components need to be measured at very near to the antenna. Thus, it is not practical, considering the measurement approach to model an antenna. The alternative approach is to use the transversal set (θ and ϕ components) of electric or magnetic fields.

In order to utilize the orthogonality properties of the associated Legendre functions [*Hansen*, 1988; *R F Harrington*, 2001], we multiply E_θ with $\frac{jp}{\sin\theta} P_q^{|m|}(\cos\theta)e^{-jm\phi}$ and integrate it over spherical surface to give

$$\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left(E_\theta \frac{jp}{\sin\theta} P_q^{|m|} e^{-jm\phi}\right)\sin\theta d\theta d\phi$$
$$= \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left\{\left(\sum_{m,n}\frac{m^2}{r_0\sin^2\theta}F_{m,n}\hat{H}_n^{(2)}P_n^{|m|}P_q^{|m|}\right)+\sum_{m,n}\left(\frac{jmk}{j\omega\varepsilon_0 r_0\sin\theta}A_{m,n}\hat{H}_n^{(2)\prime}\frac{d}{d\theta}P_n^{|m|}P_q^{|m|}\right)\right\}\sin\theta d\theta d\phi \tag{3.18}$$

Similarly, multiply E_ϕ with $\frac{d}{d\theta} P_q^{|m|}(\cos\theta)e^{-jm\phi}$ and integrate it over spherical surface, i.e.

$$\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left(E_\phi \frac{d}{d\theta} P_q^{|m|} e^{-jm\phi}\right)\sin\theta d\theta d\phi$$
$$= \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left\{\sum_{m,n}\left(\frac{1}{r_0}F_{m,n}\hat{H}_n^{(2)}\frac{d}{d\theta}P_n^{|m|}\frac{d}{d\theta}P_q^{|m|}\right)+\sum_{m,n}\left(\frac{jmk}{j\omega\varepsilon_0 r_0\sin\theta}A_{m,n}\hat{H}_n^{(2)\prime}P_n^{|m|}\frac{d}{d\theta}P_q^{|m|}\right)\right\}\sin\theta d\theta d\phi \tag{3.19}$$

Adding the above two equations, we have

$$\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left(\left[E_\theta\frac{jm}{\sin\theta}P_q^{|m|}+E_\phi\frac{d}{d\theta}P_q^{|m|}\right]e^{-jm\phi}\right)\sin\theta d\theta d\phi$$
$$= \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\sum_{m,n}\left(\frac{F_{mn}}{r_0}\hat{H}_n^{(2)}\left[\frac{m^2}{\sin^2\theta}P_n^{|m|}P_q^{|m|}+\frac{d}{d\theta}P_n^{|m|}\frac{d}{d\theta}P_q^{|m|}\right]\right)\sin\theta d\theta d\phi \tag{3.20}$$
$$+\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\sum_{m,n}\left(\frac{jmk}{j\omega\varepsilon_0 r_0}A_{mn}\hat{H}_n^{(2)\prime}\left[\frac{d}{d\theta}P_n^{|m|}\frac{P_q^{|m|}}{\sin\theta}+\frac{P_n^{|m|}}{\sin\theta}\frac{d}{d\theta}P_q^{|m|}\right]\right)\sin\theta d\theta d\phi$$

The second integral term in (3.20) vanishes for any combination of *n* and *q*, since

$$m\int_0^{\pi}\left[\frac{d}{d\theta}P_n^{\,m}(\cos\theta)\frac{P_q^{\,m}(\cos\theta)}{\sin\theta}+\frac{P_n^{\,m}(\cos\theta)}{\sin\theta}\frac{d}{d\theta}P_q^{\,m}(\cos\theta)\right]\sin\theta d\theta=0 \tag{3.21}$$

Furthermore, the first integral term (3.20) can be evaluated using the following equation utilizing the orthogonality property of the associated Legendre functions:

$$\begin{aligned}&\int_0^{\pi}\left[\frac{m^2}{\sin^2\theta}P_n^{\,m}(\cos\theta)P_q^{\,m}(\cos\theta)+\frac{d}{d\theta}P_n^{\,m}(\cos\theta)\frac{d}{d\theta}P_q^{\,m}(\cos\theta)\right]\sin\theta d\theta\\&=\begin{cases}\dfrac{2}{2n+1}\dfrac{(n+m)!}{(n-m)!}n(n+1), & n=q\\ 0, \quad otherwise\end{cases}\end{aligned} \tag{3.22}$$

Proof of (3.22) is provided in the Appendix A. Finally, we have

$$\begin{aligned}&\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left(\left[E_\theta\frac{jm}{\sin\theta}P_q^{|m|}+E_\phi\frac{d}{d\theta}P_q^{|m|}\right]e^{-jm\phi}\right)\sin\theta d\theta d\phi\\&=2\pi\frac{F_{m,q}}{r_0}\hat{H}_q^{\,(2)}\frac{2}{2q+1}\frac{(q+|m|)!}{(q-|m|)!}q(q+1)\end{aligned} \tag{3.23}$$

or

$$F_{m,n}=\frac{\dfrac{r_0}{2\pi}\displaystyle\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left(\left[E_\theta\frac{jm}{\sin\theta}P_n^{|m|}+E_\phi\frac{d}{d\theta}P_n^{|m|}\right]e^{-jm\phi}\right)\sin\theta d\theta d\phi}{\hat{H}_n^{\,(2)}\dfrac{2}{2n+1}\dfrac{(n+|m|)!}{(n-|m|)!}n(n+1)} \tag{3.24}$$

In a similar manner, $A_{m,n}$ is found as

$$A_{m,n}=\frac{\dfrac{j\omega\varepsilon_0 r_0}{2\pi k}\displaystyle\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left(\left[E_\theta\frac{d}{d\theta}P_n^{|m|}+\frac{-jm}{\sin\theta}E_\phi P_n^{|m|}\right]e^{-jm\phi}\right)\sin\theta d\theta d\phi}{\hat{H}_n^{\,(2)\prime}\dfrac{2}{2n+1}\dfrac{(n+|m|)!}{(n-|m|)!}n(n+1)} \tag{3.25}$$

As soon as the mode coefficients are known, fields beyond the surface ($r > r_0$) can be evaluated from the summation of the modes.

The explained process of obtaining the coupling coefficients is similar to the method of moments (MoM) expanding currents with basis and weighting functions [*RF Harrington*, 1993]. Consider sample data (f) that is expanded with a basis $u_{m,n}$, i.e.

$$f = \sum_{m,n} \alpha_{m,n} u_{m,n} \tag{3.26}$$

Then, we apply an inner product ($<\cdot>$) to both sides in (3.26) with a weighting function $w_{m,n}$ to have

$$\left\langle f, w_{m,n} \right\rangle = \sum_{m,n} \alpha_{m,n} \left\langle u_{m,n}, w_{m,n} \right\rangle \tag{3.27}$$

Thus, the coefficient $\alpha_{m,n}$ can be found through the residue theorem [*RF Harrington*, 1993], however, the process of obtaining coefficients $\alpha_{m,n}$ can be straightforward if we apply orthogonal properties in evaluating (3.27), as we demonstrated with the case of the associated Legendre functions. This MoM-based concept was applied to a near-to-far-field transform in [*Sarkar and Taaghol*, 1999].

Modal radiated powers of the each spherical TE and TM modes are found respectively as

$$P_{rad}^{TE_{m,n}} = \pi \frac{2n(n+1)}{2n+1} \frac{(n+|m|)!}{(n-|m|)!} \left(\frac{1}{\eta_0} \left| F_{m,n} \right|^2 \right) \tag{3.28}$$

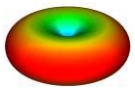
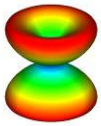
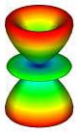
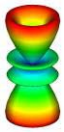
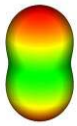
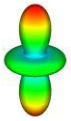
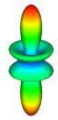
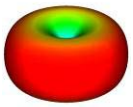
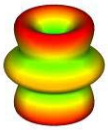
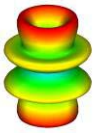
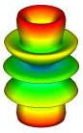
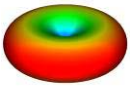
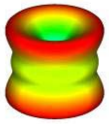
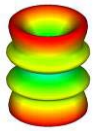
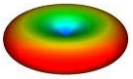
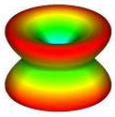
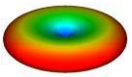
$$P_{rad}^{TM_{m,n}} = \pi \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{2n(n+1)}{2n+1} \frac{(n+|m|)!}{(n-|m|)!} \left(\eta_0 \left| A_{m,n} \right|^2 \right) \tag{3.29}$$

Therefore, the square of coupling coefficient magnitude is proportional to the modal power of the spherical mode. The total power of each spherical TE and TM mode is simply given as a sum of the modal powers:

$$P_{rad}^{TE} = \sum_{m,n} P_{rad}^{TE_{m,n}} \text{ and } P_{rad}^{TM} = \sum_{m,n} P_{rad}^{TM_{m,n}} \tag{3.30}$$

Radiation patterns of various spherical modes are shown in Table 3-1 with the assumption of an exponential harmonic function, giving an omni-directional pattern. Because the associated Legendre function $P_n^m(\cos\theta)$ becomes zero if $m > n$, an antenna cannot excite the spherical modes with $m > n$.

Table 3-1 Radiation patterns of various spherical modes in linear scale with the assumption of exponential harmonic function

	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
$m = 0$	none					
$m = 1$	none					
$m = 2$	none	none				
$m = 3$	none	none	none			
$m = 4$	none	none	none	none		
$m = 5$	none	none	none	none	none	

The spherical wave expansion of integration approach was implemented with Matlab® [*Matlab*, 2007] in order to understand SWE further. Consider an electrically-small electric dipole exciting the fundamental and two harmonic spherical TM modes (TM_{01}, TM_{03}, and TM_{05}). From (3.10), the magnetic vector potential can be written as

$$A_r = \sum_{n=1,2,3} A_{0,n} \hat{H}_n^{(2)}(kr) P_n^0(\cos\theta) e^{j0\phi} \tag{3.31}$$

where n = 1, 3, 5 and $A_{0,n}$ are the mode coupling coefficients. The electric fields of the summed spherical modes are found as

$$E_r = \frac{1}{j\omega\varepsilon_0}\left(\frac{\partial^2}{\partial r^2} + k^2\right)\left[\sum_{n=1,2,3} A_{0,n} \hat{H}_n^{(2)}(kr) P_n^0(\cos\theta)\right] \tag{3.32}$$

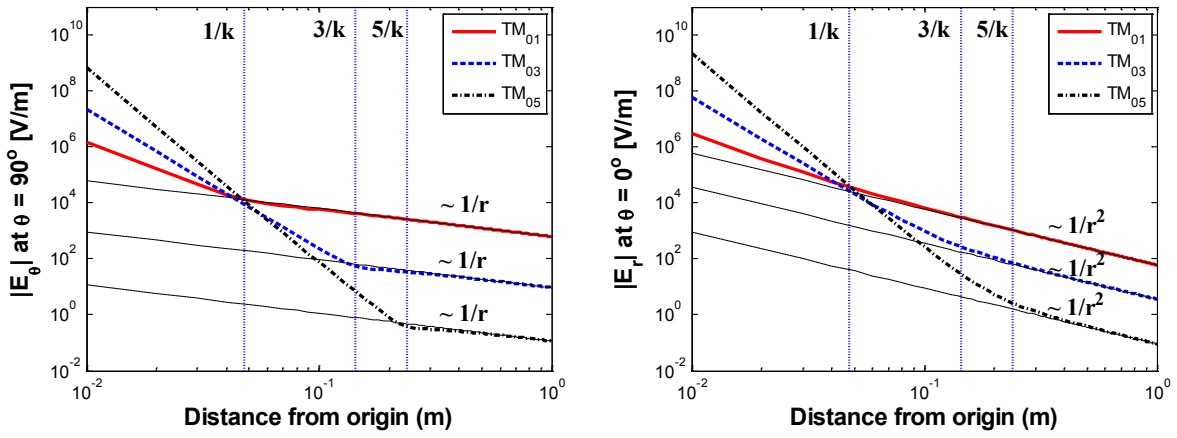
$$E_\theta = \frac{1}{j\omega\varepsilon_0 r}\frac{\partial^2}{\partial r \partial\theta}\left[\sum_{n=1,2,3} A_{0,n}\hat{H}_n^{(2)}(kr)P_n^0(\cos\theta)\right] \tag{3.33}$$

Then, the electric fields for each spherical TM_{0n} mode are evaluated from

$$E_r^{TM_{0,n}}(r,\theta,\phi) = -jk\eta_0 A_{0,n}P_n(\cos\theta)\left[\frac{d^2}{d(kr)^2}\hat{H}_n^{(2)}(kr) + \hat{H}_n^{(2)}(kr)\right] \tag{3.34}$$

$$E_\theta^{TM_{0,n}}(r,\theta,\phi) = -j\eta_0 A_{0,n}\frac{d}{d(kr)}\hat{H}_n^{(2)}(kr)\frac{d}{d\theta}P_n(\cos\theta) \tag{3.35}$$

Comparison of the electric field strength for $\left|E_\theta^{TM_{0,n}}\right|$ and $\left|E_r^{TM_{0,n}}\right|$ is shown in Figure 3.13 for the spherical TM modes. The angular electric fields $\left|E_\theta^{TM_{0,n}}\right|$ at $\theta = \pi/2$ decay with $1/r$ dependence above the radian sphere, corresponding to each mode, while the radial electric fields $\left|E_r^{TM_{0,n}}\right|$ at $\theta = 0$ decay with $1/r^2$ dependence. Therefore, angular electric fields decay in the far-field region. However, the near-field pattern will not have a null at the poles like the well-known, donut-shape far-field pattern. The field strength with coupling coefficients $A_{0n} = 1$ can have an approximate order of 10^{10} in the vicinity of the antenna. This suggests that spherical coefficients with a fairly low value also need to be considered to evaluate the near-field strength.



(a) (b)

Figure 3.13 Comparison of electric field strength versus distance for various spherical TM modes at 1 GHz – (a) At $\theta = 0$ and (b) At $\theta = \pi/2$.

Considering a source normalization factor ($Idl = 4\pi$) for the example dipole, we set the spherical coefficients as the followings as

$$A_{0,1} = -\frac{jk}{4\pi} = -j1.6667 \tag{3.36}$$

$$A_{0,3} = A_{0,1} * 10^{-2} = -j1.6667\text{e-}2 \tag{3.37}$$

$$A_{0,5} = A_{0,1} * 10^{-4} = -j1.6667\text{e-}4 \tag{3.38}$$

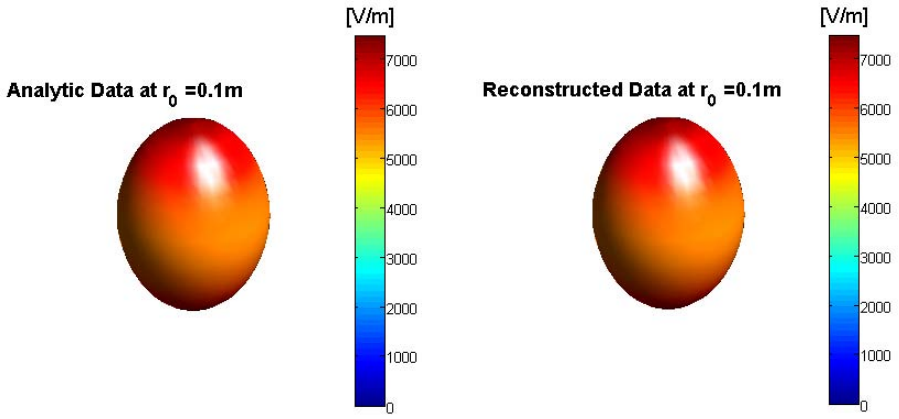
In order to test the developed code, near-field electric-field data sampled every 10 degrees in θ and ϕ angles were obtained at r = 0.05 m at 1 GHz from the exact field equations (3.32) and (3.33). The sampled data was used as input to the developed code to evaluate the mode coupling coefficients.

The screen capture of the analysis result is shown in Figure 3.14. It is observed that there are real parts in the computed mode coupling coefficients, which should be zero. However, these real numbers are significantly small, compared to the imaginary part of each mode coupling coefficient. Thus, this small error will not affect the accuracy of the field reconstruction from the spherical-wave expansion results. In order to demonstrate it, the field patterns of the test dipole antenna were evaluated at near (r_0 = 0.1 m) and far distance (r_0 = 10 m) for both analytic and reconstructed cases. As shown in Figure 3.15 and Figure 3.16, both original and reconstructed field patterns match very well.

For error analysis, the index number for field samples in (θ, ϕ) angles are as defined in Table 3-2. The amplitude differences between the original and the reconstructed field data are plotted in Figure 3.17 and Figure 3.18 for the near and far field case, respectively. As we can see in the figures, the differences between the original and the reconstructed data are negligible, compared to the peak magnitude of the fields. The developed SWE code was also tested and verified through simulated and measured data of realistic antennas, but these results are not presented here.

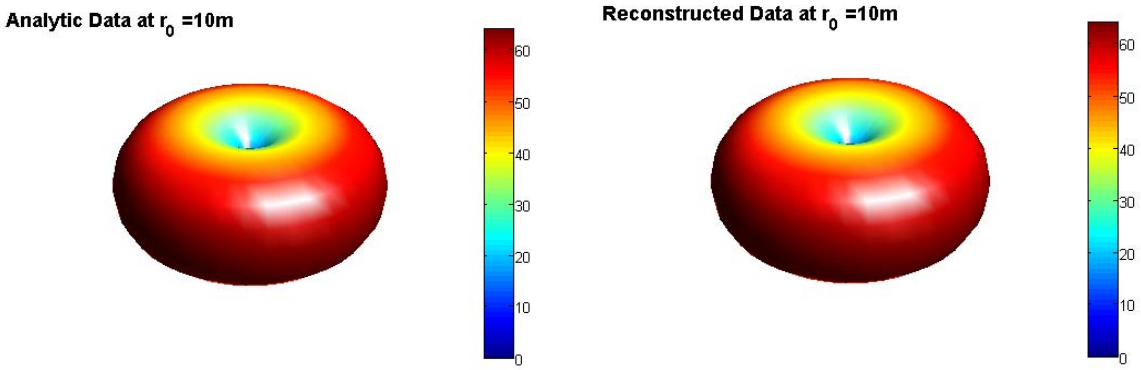
```
TM Mode: Amn( 0,  1) = (-1.153947e-006) +j(-1.666667e+000)
TM Mode: Amn( 0,  3) = (-1.448028e-007) +j(-1.666630e-002)
TM Mode: Amn( 0,  5) = (-1.441978e-010) +j(-1.665010e-004)

Total elapsed time for computation: 1.88 min
```

Figure 3.14 Captured result screen of the SWE code for the test dipole case.



(a) (b)

Figure 3.15 Comparison between analytic and reconstructed near-field pattern in 3-D at $r_0 = 0.1$ to be in near-field region– (a) From analytic exact data and (b) From reconstructed data through the developed SWE code.

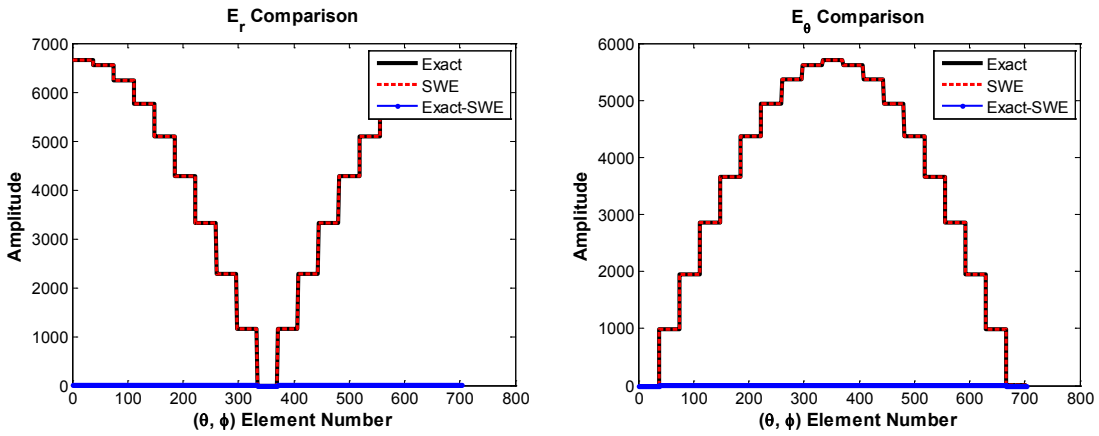


(a) (b)

Figure 3.16 Comparison between analytic and reconstructed near-field pattern in 3-D at $r_0 = 10$ to be in far-field region– (a) From analytic exact data and (b) From reconstructed data through the developed SWE code.

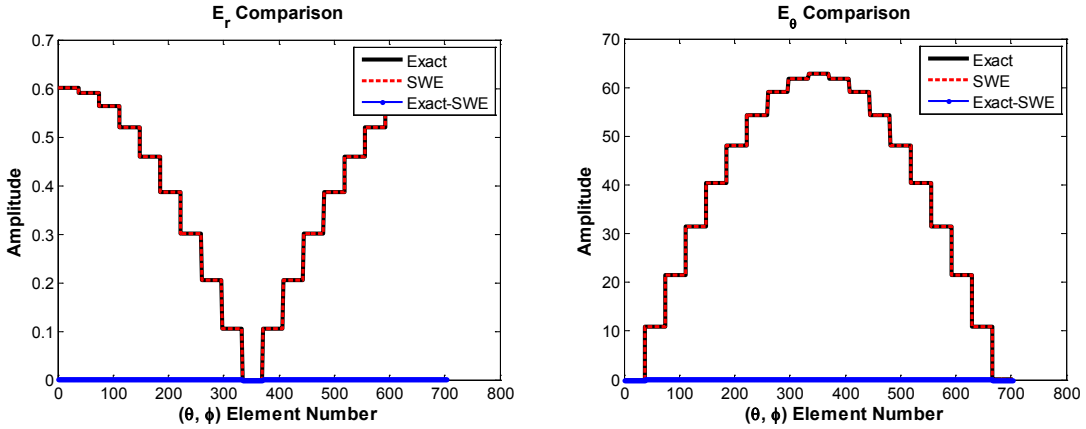
Table 3-2 Reference for index number of data in (θ, ϕ) angle.

	$\phi = 0^\circ$	$\phi = 10^\circ$	$\phi = 20^\circ$	⋯	$\phi = 360^\circ$
$\theta = 0^\circ$	1	20	39	⋯	684
$\theta = 10^\circ$	2	21	40	⋯	685
$\theta = 20^\circ$	3	22	41	⋯	686
⋮	⋮	⋮	⋮	⋮	⋮
$\theta = 180^\circ$	19	38	57	⋯	19* 37 =703



(a) (b)

Figure 3.17 Error analysis through the difference between analytic and reconstructed field amplitudes at $r_0 = 0.1$ – (a) Error in reconstructed radial electric field and (b) Error in reconstructed angular electric field.



(a) (b)

Figure 3.18 Error analysis through the difference between analytic and reconstructed field amplitudes at $r_0 = 10$ – (a) Error in reconstructed radial electric field and (b) Error in reconstructed angular electric field.

3.2.1.2 The Matrix Approach

Observing that the SWE is based on linear operations for a finite number of sampled data, a direct matrix approach [*James and Longdon*, 1969] can be considered as an alternative to the integration approach. If the maximum degree of the excited spherical mode for an antenna is N (*a priori* information), (3.16) can be written as in a matrix form:

$$
\begin{aligned}
jE_\theta(r_0,\theta,\phi) &= \sum_{n=1}^{N}\sum_{m=-n}^{n}\left(\frac{m}{\sin\theta} f_{mn} P_n^{|m|}(\cos\theta) e^{jm\phi}\right) + \sum_{n=1}^{N}\sum_{m=-n}^{n}\left(a_{mn}\frac{d}{d\theta} P_n^{|m|}(\cos\theta) e^{jm\phi}\right) \\
&= \begin{bmatrix} S_{-n,1}^{(1)}(r_0,\theta,\phi) & \cdots & S_{n,N}^{(1)}(r_0,\theta,\phi) & S_{-n,1}^{(2)}(r_0,\theta,\phi) & \cdots & S_{n,N}^{(2)}(r_0,\theta,\phi) \end{bmatrix} \begin{bmatrix} f_{-n,1} \\ \vdots \\ f_{n,N} \\ a_{-n,1} \\ \vdots \\ a_{n,N} \end{bmatrix}
\end{aligned} \tag{3.39}
$$

where $a_{mn} = \eta_0 \hat{H}_n^{(2)\prime}(ka) A_{mn} / r_0$, $f_{mn} = \hat{H}_n^{(2)}(kr_0) F_{mn} / r_0$, $S_{m,n}^{(1)}(r_0,\theta,\phi) = \dfrac{m}{\sin\theta} P_n^{|m|}(\cos\theta) e^{jm\phi}$, and $S_{m,n}^{(2)}(r_0,\theta,\phi) = \dfrac{d}{d\theta} P_n^{|m|}(\cos\theta) e^{jm\phi}$.

Finding a matrix form for E_ϕ in a similar manner and forming the combined matrix:

$$
\begin{aligned}
&\begin{bmatrix} -jE_\theta(r_0,\theta,\phi) \\ E_\phi(r,\theta,\phi) \end{bmatrix} \\
&= \begin{bmatrix} \displaystyle\sum_{n=1}^{N}\sum_{m=-n}^{n}\left(\frac{m}{\sin\theta} f_{mn} P_n^{|m|}(\cos\theta) e^{jm\phi}\right) & \displaystyle\sum_{n=1}^{N}\sum_{m=-n}^{n}\left(a_{mn}\frac{d}{d\theta} P_n^{|m|}(\cos\theta) e^{jm\phi}\right) \\ \displaystyle\sum_{n=1}^{N}\sum_{m=-n}^{n}\left(f_{mn}\frac{d}{d\theta} P_n^{|m|}(\cos\theta) e^{jm\phi}\right) & \displaystyle\sum_{n=1}^{N}\sum_{m=-n}^{n}\left(\frac{m}{\sin\theta} a_{mn} P_n^{|m|}(\cos\theta) e^{jm\phi}\right) \end{bmatrix} \\
&= \begin{bmatrix} S_{-n,1}^{(1)}(r_0,\theta,\phi) & \cdots & S_{n,N}^{(1)}(r_0,\theta,\phi) & S_{-n,1}^{(2)}(r_0,\theta,\phi) & \cdots & S_{n,N}^{(2)}(r_0,\theta,\phi) \\ S_{-n,1}^{(2)}(r_0,\theta,\phi) & \cdots & S_{n,N}^{(2)}(r_0,\theta,\phi) & S_{-n,1}^{(1)}(r_0,\theta,\phi) & \cdots & S_{n,N}^{(1)}(r_0,\theta,\phi) \end{bmatrix} \begin{bmatrix} f_{-n,1} \\ \vdots \\ f_{n,N} \\ a_{-n,1} \\ \vdots \\ a_{n,N} \end{bmatrix}
\end{aligned} \tag{3.40}
$$

In this combined form, we need $2N(N+2)$ linearly-independent sets of sampled data to obtain an exact solution. However, typical measurement systems require uniform sampling points in both θ and ϕ directions in practice. These uniformly-sampled data are not necessarily independent each other. Thus, it has been common to obtain oversampled data. As a result, the uniformly-sampled data matrix typically has a form of a non-square matrix giving an over-determined system. In this sense, the pseudo-inverse can be used to find the mode coupling

coefficients. If K number of sampled data is available for each transverse electric component, the coupling coefficients can be found from

$$\underbrace{\begin{bmatrix} \underbrace{f_{mn}}_{[N(N+2)\times 1]} \\ \underbrace{a_{mn}}_{[N(N+2)\times 1]} \end{bmatrix}}_{[2N(N+2)\times 1]} = \underbrace{\begin{bmatrix} \underbrace{S^{(1)}_{m,n}(r_0,\theta,\phi)}_{[K\times N(N+2)]} & \underbrace{S^{(2)}_{m,n}(r_0,\theta,\phi)}_{[K\times N(N+2)]} \\ \underbrace{S^{(2)}_{m,n}(r_0,\theta,\phi)}_{[K\times N(N+2)]} & \underbrace{S^{(1)}_{m,n}(r_0,\theta,\phi)}_{[K\times N(N+2)]} \end{bmatrix}^{\dagger}}_{[2N(N+2)\times 2K]} \underbrace{\begin{bmatrix} \underbrace{jE_{\theta}(r_0,\theta,\phi)}_{[K\times 1]} \\ \underbrace{E_{\phi}(r_0,\theta,\phi)}_{[K\times 1]} \end{bmatrix}}_{[2K\times 1]} \tag{3.41}$$

where † is Moore-Penrose pseudo-inverse [*Golub and Van Loan*, 1996].

Compared to the integration approach, the direct matrix approach may not be computationally efficient, but the matrix approach is simple and straight forward for obtaining the coupling coefficients. In addition, the noise reduction techniques using the singular value decomposition (SVD) can be applied in (3.41) as a free filtering process. As demonstrated in [*Yang and Davis*, 2011], the direct matrix approach can also be applied to partially-scanned or incompletely-scanned data. In this case, the obtained coupling coefficients correspond to a least-square solution for under-determined system.

Thus, the advance in computing technology makes the direct matrix method more attractable due to the additional advantages of the matrix approach over the integration approach.

3.2.2 The Decomposed Transfer Function of Canonical Antennas

Considering how much power is delivered to the far-field region, the spherical modal power expressions of (3.28) and (3.29) suggest an antenna transfer function model in Figure 3.19 that is similar to Hansen's network description used for a near to far-field transformation [*Hansen*, 1988] . In the model, the relationship between an antenna input port and spherical wave mode is described with scattering parameters representing the antenna transfer function for each spherical mode.

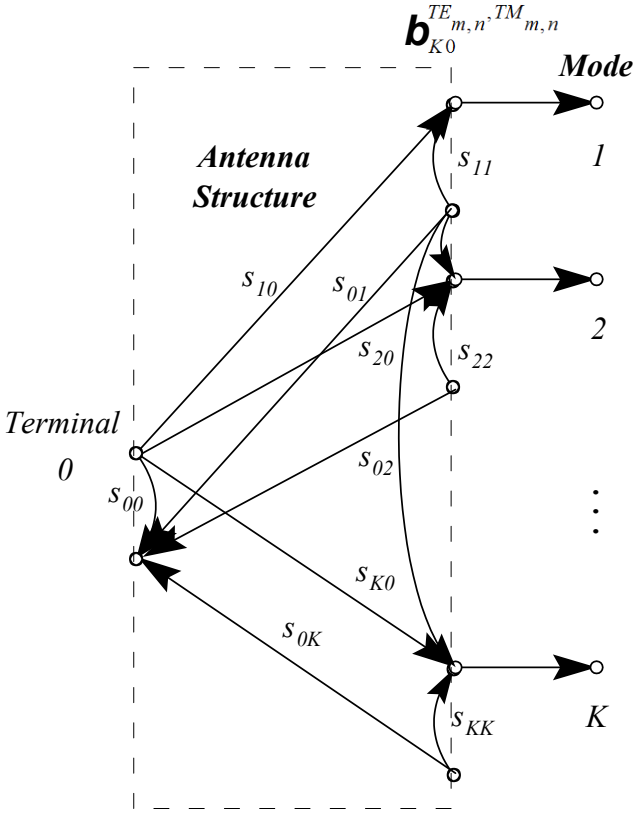


Figure 3.19 Antenna transfer function model with S-parameter representation of radiation coupling to the spherical modes.

The outgoing power wave for each spherical TE and TM mode can be written as

$$b_{K0}^{TE_{m,n}} = 2F_{m,n}\sqrt{\frac{\pi}{\eta}\frac{2n(n+1)}{2(n+1)}\frac{(n+|m|)!}{(n-|m|)!}} \tag{3.42}$$

$$b_{K0}^{TM_{m,n}} = 2A_{m,n}\sqrt{\pi\eta\frac{2n(n+1)}{2(n+1)}\frac{(n+|m|)!}{(n-|m|)!}} \tag{3.43}$$

which can be found from SWE method. Thus, if input to the transfer function is characterized, we can obtain the transfer function.

For example, consider the ideal dipole exciting only the fundamental spherical mode. Then, the model becomes a two-port network system. The corresponding scattering matrix can be defined in terms of incident (a_0 and a_1) and reflected waves (b_0 and b_1):

$$\begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} s_{00} & s_{01} \\ s_{10} & s_{11} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} \tag{3.44}$$

Thus, $|s_{00}|^2$ represents return loss and $|s_{10}|^2$ is proportional to the power delivered to the far-field region. Since poles are related to the characteristic solution of the system, we set both a_0 and a_1 equal to zero, i.e. no input to the system. Then, the scattering matrix should be singular in order to have non-zero b_0 and b_1, i.e. det[s] = 0. If the system is reciprocal (or passive network), we have

$$s_{00}s_{11} = (s_{10})^2 = (s_{01})^2 \tag{3.45}$$

Because the poles are unique, (3.45) implies

$$poles\ of\ s_{00} = poles\ of\ s_{11} = poles\ of\ s_{10} = poles\ of\ s_{01} \tag{3.46}$$

Thus, all scattering parameters have the same poles for a passive system or a passive antenna. In other words, the poles of a return-loss-versus-frequency curve are the same as the poles in power-versus-frequency curve. In addition, the pole of return-loss-versus-frequency curve is the result of a resonance phenomenon at the corresponding frequency of the pole.

This observation leads us to an idea for modeling the coupling coefficient variation in the frequency or the time domain using the singularity expansion method (SEM). For the example of spherical TM mode case, the mode coupling coefficient can be represented as

$$s_{K,0}^{TM_{m,n}}(t) = \sum_i R_i e^{p_i t} \Leftrightarrow s_{K,0}^{TM_{m,n}}(\omega) = \sum_i \frac{R_i}{p - p_i} \tag{3.47}$$

where $p_i = \sigma_i + j\omega_i$ represents the complex poles, σ_i is damping factor, ω_i is angular frequency, and R_i is the angle-dependant residue providing a pattern. Compared to the SEM-based model using the antenna effective height in previous section, the new SEM-base model using SWE offers an additional pre-filtering process that removes non-dominant spherical modes. In addition, the new model can also model near fields.

In the following subsections, we present examples of the new model applied to a resonant, an ultra-wideband, and a frequency-independent antenna in order to demonstrate the usefulness of the new model.

3.2.2.1 Application of Decomposed SEM model to A Resonant Antenna

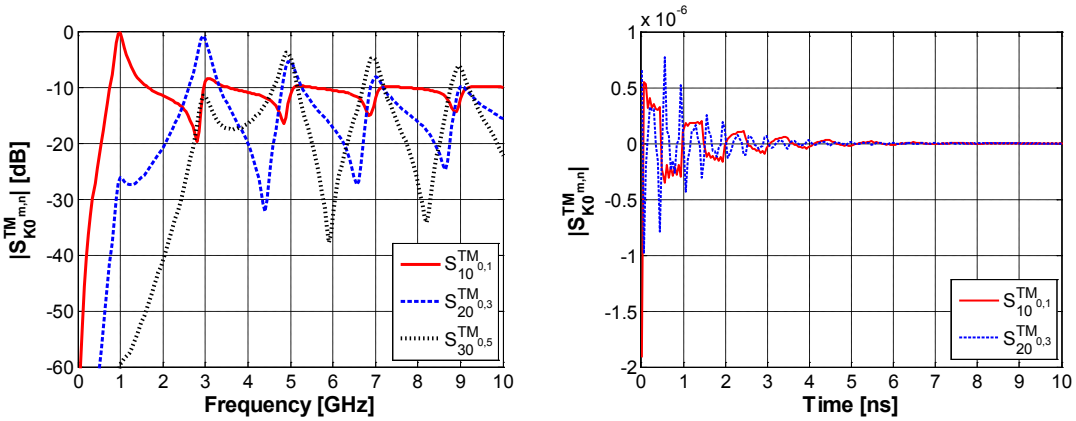
As a resonant antenna example, consider a thin-wire dipole with a wire radius of 0.01 mm. The dipole is a half-wave resonant at 1 GHz. Due to the symmetry in the half-wavelength dipole, spherical modes with an odd numbered order can be excited, i.e. spherical $TM_{0,n}$ mode (n = 1, 3, 5, …). The thin-wire dipole was simulated using a commercial method of moments (MoM) code [*FEKO*, 2007]. The simulation frequency was swept from 50 MHz to 10 GHz in 50 MHz steps. At each simulation frequency, near-field data on the enclosing sphere surface (r_0 = 15 cm) were computed. With the computed near-field data, mode coupling coefficients were computed using the developed SWE code at each sampled frequency.

The computed transfer functions for some selected spherical modes are compared in Figure 3.20. As expected, spherical modes with an odds order number are excited. The fundamental spherical mode (TM_{01}) is strong at 1 GHz while the third harmonic spherical mode (TM_{03}) is strong at 3 GHz. As frequency increases, the number of excited spherical modes with higher degree increases. The comparison for the transfer functions in the time domain shows that the pulse width of the transfer function for the third harmonic spherical mode (TM_{03}) is narrower than that of the fundamental spherical mode (TM_{01}). The radiation pattern at 10 GHz in Figure 3.21 shows many ripples in elevation angle, compared to the pattern at the frequency of the fundamental spherical mode excitation at 1 GHz. That is due to the excitation of spherical modes with a high degree.

The pole-residue relationship of the antenna transfer function for the dominant spherical TM_{01} mode was also found using the matrix pencil method and is shown in Figure 3.22. A pole is located on the damping coefficient axis with a high damping value. This pole seems to represent the source effect dominating the early-time response. Other poles are located at the odd harmonic frequencies. Compared to the pole-residue relationship of the effective height in Figure 3.11a, the pole-residue relationship of the new model in Figure 3.22 does not have many other poles with low residues. This may indicate that the poles with low residues in Figure 3.11a came from the spherical mode excitation with higher degree.

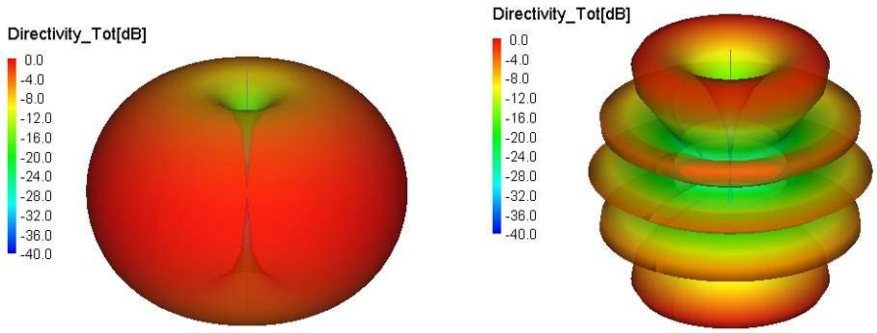
The pole-residue relationship of the transfer function for the fundamental mode is compared for the case of thin (radius = 0.01 mm) and fat-wire (radius = 0.25 mm) dipoles in Figure 3.23. One may think that the radius of 0.25 mm is not fat, but it is relatively fat when it is compared to

the thin-wire case. As is well known, a fatter dipole antenna has a wider bandwidth. In terms of time-domain viewpoint, the wider bandwidth means a pulse with a narrow pulse width and/or a pulse with a high damping factor. In Figure 3.23, we can observe that the damping coefficients of the fat-wire dipole have overall higher damping coefficients than the case of the thin-wire dipole while the pole locations remain at the same frequencies.



(a) (b)

Figure 3.20 Transfer functions for 1-GHz resonant thin-wire dipole antenna with a wire radius of 0.01 mm – (a) Frequency-domain response and (b) Time domain response.



(a) (b)

Figure 3.21 Simulated radiation pattern of thin-wire dipole antenna with a wire radius of 0.01 mm – (a) 1 GHz and (b) 10 GHz. A commercial method of moments (MoM) code [*FEKO*, 2007] was used for the simulation.

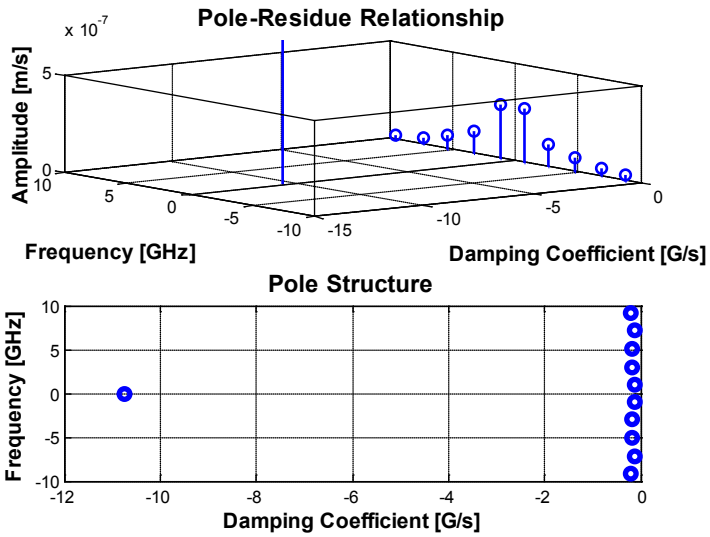


Figure 3.22 Pole-residue relationship of the decomposed transfer function (spherical TM_{01} mode) for 1-GHz thin-wire dipole antenna with a wire radius of 0.01 mm.

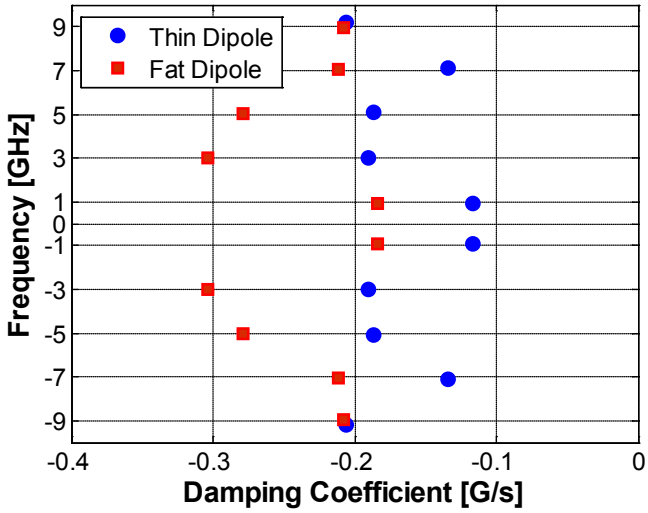


Figure 3.23 Pole location comparison of decomposed transfer functions (spherical TM_{01} mode) between thin (radius of 0.01 mm) and fat-wire (radius of 0.25 mm) dipole antennas.

3.2.2.2 Application of Decomposed SEM model to An Ultra-Wideband Antenna

For an ultra-wideband antenna example, consider a bicone antenna with cone angle of 45 degrees. The antenna height is a half-wavelength at 750 MHz. The overall geometry is shown in Figure 3.24. Similar to the dipole example in the previous subsection, the near-field data on a spherical

surface (r_0 = 30 cm) were obtained through simulations using a commercial method of moments (MoM) code [*FEKO*, 2007]. The frequency range and step are the same as the dipole case.

In a sense, the bicone antenna is a much fatter version of the dipole with a tapered dual slot. Unlike the resonant dipole case, the bicone antenna excites a fundamental spherical mode that is dominant over the entire operational frequency range, as shown in Figure 3.25a. The resulting pulse shape in Figure 3.25b is very narrow with a minor ringing at the late-time response. Thus, the radiation pattern at 10 GHz (Figure 3.26b) is close to the pattern of the fundamental spherical TM_{01} mode at 1 GHz in Figure 3.26a. There are still some ripples in elevation angle for the 10-GHz pattern, but these ripples are minor, compared to the dipole case in Figure 3.21b.

The pole-residue relationship of the bicone transfer function for the dominant spherical TM_{01} mode is shown in Figure 3.27. Except for the pole on the damping coefficient axis caused from the source effects in early-time response, only four poles can represent the radiation characteristics of bicone antenna from 50 MHz to 10 GHz, which is a similar result that Licul had with effective height analysis in [*Licul*, 2004].

The pole-residue relationships of the fundamental spherical-mode transfer functions with cone angles of 45 and 30 degrees are compared in Figure 3.28. The antenna height for both cases remained the same. The bicone antenna with a wider cone angle has lower characteristic impedance in terms of transmission line viewpoint. From the bandwidth viewpoint, the fatter bicone antenna (cone angle of 45 degrees) has wider bandwidth. Since the antenna height is fixed, this increased bandwidth is due to the increased highest operational frequency. As a result, we can observe that one of the pole pairs for the fatter bicone is located around 6.5 GHz while that of the other bicone with cone angle of 30 degree is located around 6 GHz. In addition, the damping coefficients of the fatter bicone around 6.5 GHz is higher than that of the other bicone with cone angle of 30 degree. These pole changes are due to the increased highest operating frequency. On the other hand, the poles located around 350 MHz almost did not change with cone angle change, which implies that these poles are related to the lowest operating frequency, the resonance of overall bicone antenna structure, and/or the minor ringing in late-time response. This observation suggests that antenna height increase will be more effective to lower the lowest operating frequency of bicone antenna than changing cone angle.

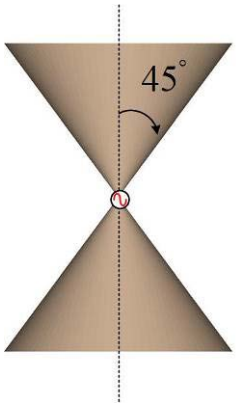
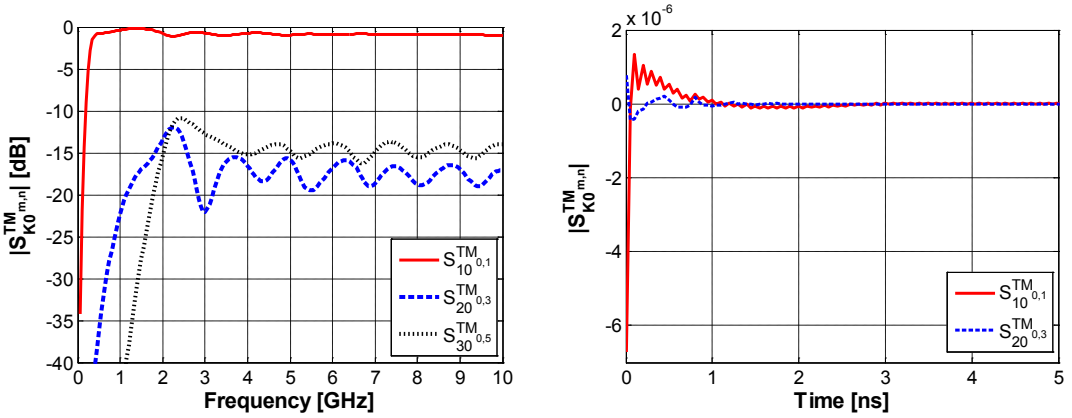
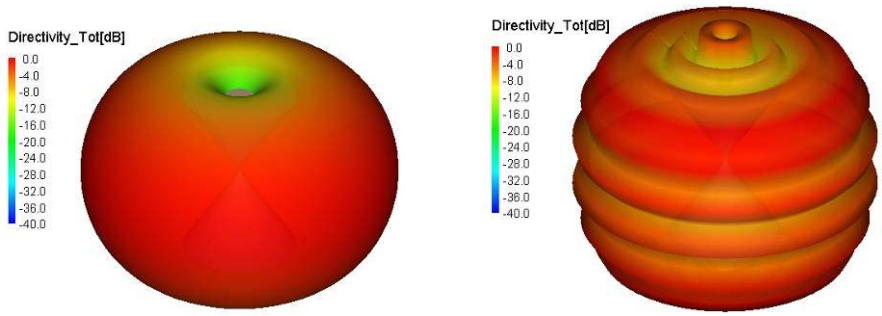


Figure 3.24 Geometry of bicone antenna. Antenna height is a half wavelength at 750 MHz (= 20 cm).



(a) (b)

Figure 3.25 Transfer functions for bicone antenna with cone angle of 45 degrees – (a) Frequency-domain response and (b) Time domain response.



(a) (b)

Figure 3.26 Simulated radiation pattern of bicone antenna – (a) 1 GHz and (b) 10 GHz. A commercial method of moments (MoM) code [*FEKO*, 2007]was used for the simulation.

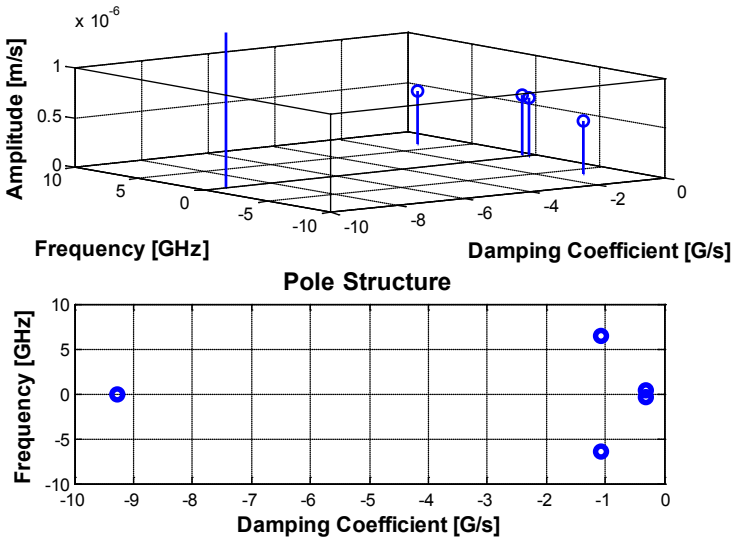


Figure 3.27 Pole-residue relationship of the decomposed transfer function (spherical TM_{01} mode) for bicone antenna with cone angle of 45 degrees.

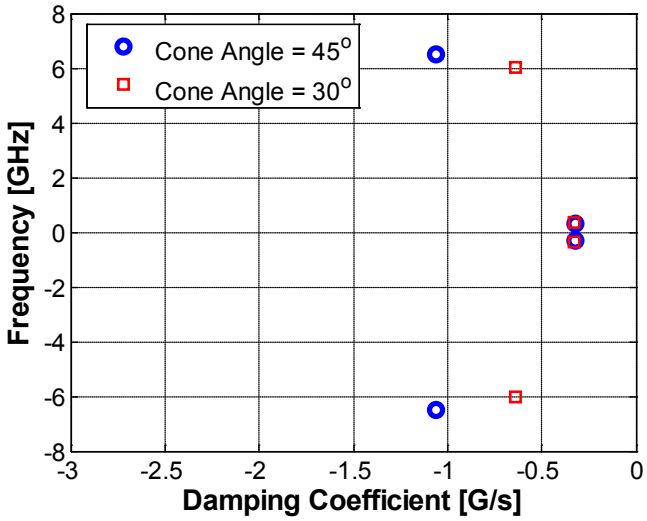


Figure 3.28 Pole location comparison of decomposed transfer functions (spherical TM_{01} mode) between bicone antennas with cone angles of 45 and 30 degrees.

3.2.2.3 Application of Decomposed SEM model to A Frequency-Independent Antenna

For a frequency-independent antenna example, consider the Archimedean spiral antenna in Figure 3.29. The maximum outer radius of the spiral antenna is 3.18 cm which is related to the lowest operational frequency. Similar to previous cases, the near-field data on a spherical surface (r_0 = 6 cm) were obtained through simulations using a commercial method of moments (MoM) code [*FEKO*, 2007]. The frequency range and step are the same as previous examples. In order

to identify the dominant spherical mode, modal power analysis was carried out at 3 GHz. The result of modal power analysis in Figure 3.30 shows that the dominant spherical mode of the Archimedean spiral antenna is $TM_{-1,1}$ mode, which provides a circular polarization. Unlike previous two examples, the Archimedean spiral antenna excites spherical modes with both even and odd orders. The Archimedean spiral antenna has a bi-directional pattern, as shown in Figure 3.31. The pattern at 10 GHz is more directive than the pattern at 1 GHz, but there are no ripples in elevation angle.

The computed transfer functions for the Archimedean spiral antenna are shown in Figure 3.32. Overall, the spherical $TM_{-1,1}$ mode is dominant in the entire frequency range under investigation. This result corresponds to the active-region explanation [*Stutzman and Thiele*, 2012] of the spiral antenna. The Archimedean spiral antenna has many resonances, providing continuous wideband performance. Each resonance excites the same spherical $TM_{-1,1}$ mode at a different frequency. Because each active region is excited with different time delay, the resulting time-domain response of the spiral antenna typically shows a chirp (see Chapter 6 for details). However, we cannot observe an obvious chirp in Figure 3.32b. That might be because the particular spiral antenna example does not have many turns [*Yang and Davis*, 2006].

The pole-residue relationships of the transfer function for spherical $TM_{-1,1}$ mode are shown in Figure 3.33. Both bicone and spiral antennas excite a single spherical mode over a wide frequency range, but the spiral antenna has noticeably more poles that correspond to the multiple resonances of the spiral antenna.

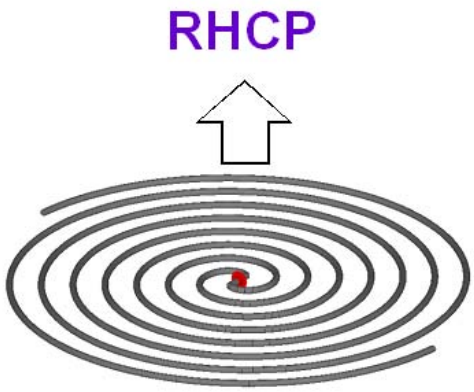


Figure 3.29 Geometry of Archimedean spiral antenna supporting right-hand circular polarization. The outer circumference corresponds to one wavelength at 1.5 GHz (maximum outer radius = 3.18 cm).

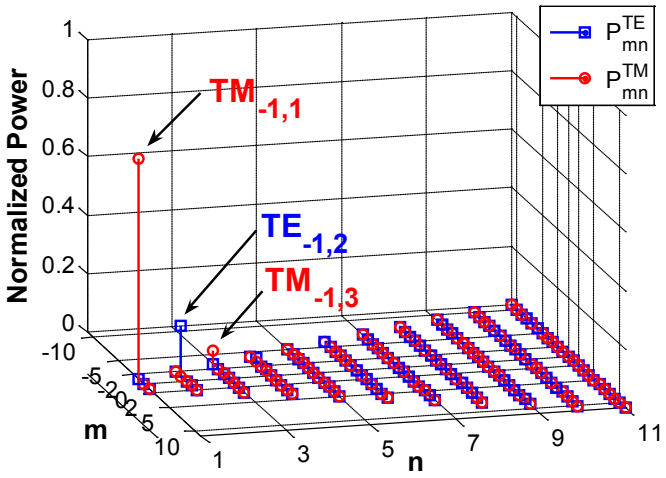
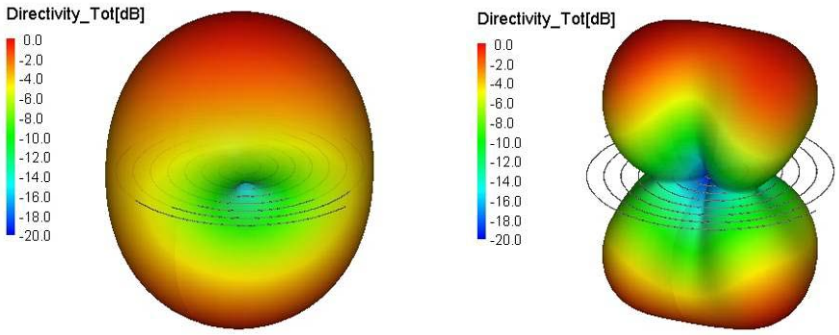
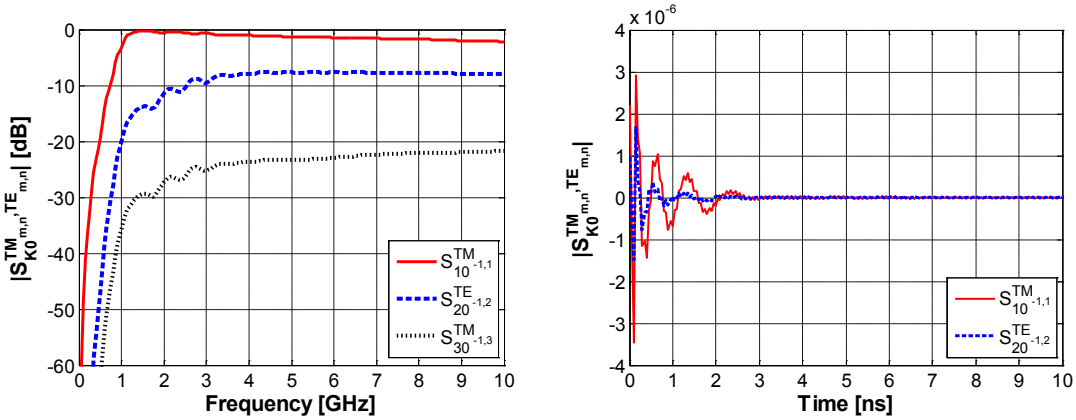


Figure 3.30 Modal power analysis for Archimedean spiral antenna at 3 GHz.



(a) (b)

Figure 3.31 Simulated radiation pattern of spiral antenna – (a) 1 GHz and (b) 10 GHz. A commercial method of moments (MoM) code [*FEKO*, 2007]was used for the simulation.



(a) (b)

Figure 3.32 Transfer functions for frequency-independent spiral antenna – (a) Frequency-domain response and (b) Time domain response.

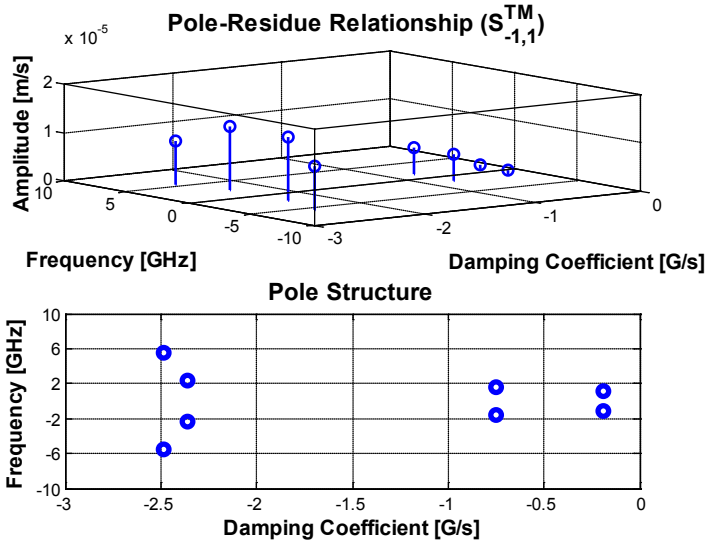


Figure 3.33 Pole-residue relationship of the decomposed transfer functions (spherical $TM_{-1,1}$ mode) for Archimedean spiral antenna.

3.3 Chapter Summary

In this chapter, antenna transfer function modeling approaches were investigated as a tool for our main study in this dissertation, based on the singularity expansion method (SEM) [*C.E. Baum*, 1971] and the spherical wave expansion (SWE) method [*Hansen*, 1988; *Stratton*, 1941].

The singularity expansion method (SEM) was reviewed along with the impulse radiating antenna (IRA) design. SEM-based antenna transfer functions using effective height were explained with examples including resonant, frequency-independent, and ultra-wideband antennas. The matrix pencil method [*Sarkar and Pereira*, 1995] was useful to obtain the poles and residues for SEM-based antenna transfer function, as Licul demonstrated in [*Licul*, 2004]. Particularly, the matrix pencil method offered opportunities to mitigate the noise issues by using a time gating approach and a pre-filtering process with singular value decomposition. Thus, the SEM-based transfer function model using antenna effective height provides a clear and unified antenna description to characterize the antenna properties in both the frequency and time domains. However, it was pointed out that Licul's approach based on the antenna effective height may encounter some practical difficulties in measurements due to the potential direct interactions between antennas when near-field characteristics of the antenna need to be included in the antenna transfer function that is described with the effective height. In addition, it was explained that observable poles can be limited if the poles are searched through measured data.

In order to overcome the practical limitation of Licul's approach, an alternative SEM-based antenna transfer function model was introduced, based on the spherical mode expansion (SWE). In order to understand the SWE in depth, the essential concept of SWE was explained using integration and direct matrix approaches. A Matlab® code was developed and used to demonstrate how SWE works for an example using ideal dipole antenna. In general, the direct matrix approach is not computationally efficient, compared to the integral approach. However, the direct matrix approach can be used even for partial or incomplete near-field scan data. Thus, advances in computing technology make the direct matrix approach more attractive than the integration approach.

The scattering-parameter network description of the decomposed spherical modes was considered as alternative antenna transfer function. It was shown that the transfer function for each spherical mode can be modeled with SEM approach. Using a simplified two-port model, the obtained poles from the transfer function are identical to the poles representing resonances that can be seen antenna terminal. The usefulness of the alternative model was demonstrated using resonant, frequency-independent, and ultra-wideband antennas example with simulated near-field scan data.

The alternative SEM-based antenna transfer function model using the SWE approach can have the same advantages that the previous SEM-based antenna transfer function model using the effective height. Since the alternative SEM-based antenna transfer function approach can include near-field characteristics of an antenna as well, the alternative SEM-based antenna transfer function approach can be considered as an extended approach to the previous SEM-based model using the antenna effective height.

Further investigation on the alternative SEM-based antenna transfer function model is necessary with measured near-field scan data to characterize the capability of the alternative approach compared to the previous approach using the effective height. However, it appears that the alternative SEM-based antenna transfer function model is adequate as a tool to investigate antenna near-field aspects that are included in this dissertation.

Chapter 4

Fundamental Limit Theory on Antenna Size and Performance

As wireless device applications expand, the demand for antenna miniaturization and interference reduction increases. However, antenna size and performance are limited by radiation physics, not technology. Not all of the energy around an antenna contributes to far-field radiation. Physics requires a minimum amount of non-radiating energy in the antenna radiation process, which affects antenna operational bandwidth and radiation efficiency. Non-radiating energy is also responsible for the near-field interference issues. On the other hand, some wireless applications can have an extended near-field communication range if the non-radiating energy is maximized.

The fundamental-limit theory on antenna size and performance not only provides a figure of merit to quantify the amount of non-radiating energy created from antennas, but also suggests how to address the near-field interference and communication range issues.

In this chapter, we investigate the fundamental-limit theory. Assumptions and errors in the classic fundamental-limit are explained based on the observation of varying energy velocity in Chapter 3. Also the classical limit theory is corrected and extended for antennas with an arbitrary polarization. The singularity/spherical-wave expansion antenna models developed in Chapter 3 is used to consider the theoretical size limit of frequency-independent and ultra-wideband antennas.

4.1 Classical Fundamental-Limit Theory

Wheeler first introduced the concept of fundamental-limit theory in 1947 for an electrically small antenna [*Wheeler*, 1947]. Since the antenna input impedance is highly capacitive (dipole antenna case) or inductive (loop antenna case) for the electrically-small antennas, Wheeler used the simple antenna model that is represented with a capacitor in Figure 4.1 and a dual inductor occupying a cylindrical volume.

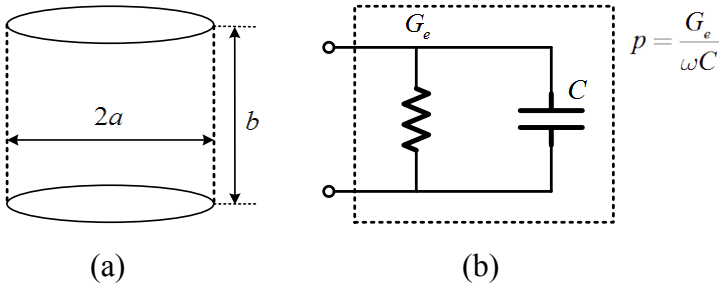


Figure 4.1 Wheeler's antenna model [*Wheeler*, 1947] – (a) Capacitor occupying a cylindrical volume and (b) Circuit model of capacitive electrically small antenna.

Wheeler considered the radiation power factor (p) as a figure of merit to describe antenna size and performance limit. Assuming that there are no fringing fields from the antenna, Wheeler found the radiation power factor for the $a \gg b$ case as

$$p = \frac{G_e}{\omega C} = \frac{k^3 a^2 b}{6} \tag{4.1}$$

where G_e is radiation shunt conductance, C is capacitance of antenna, ω is angular frequency, and k is wave number ($2\pi/\lambda$). If the capacitive electrically small antenna is tuned, the radiation power factor corresponds to the 3-dB bandwidth of the tuned circuit. Therefore, the radiation power factor is related to both antenna radiated power and bandwidth. Even though the radiation power factor provides a useful gauge for relative comparison of antenna performances, it is a rough approximation.

Since Wheeler introduced the concept of the theoretical limit on antenna size and performance, extensive investigations have been carried out by many researchers [*Chu*, 1948; *R Collin and Rothschild*, 1964; *Grimes and Grimes*, 1999; *McLean*, 1996]. Radiation Q (Q_{rad}) has been used a quantifier and was defined as [*Chu*, 1948]

$$Q_{rad} = e_r \frac{\left[W_{e,non-rad} + W_{m,non-rad}\right]_{Peak}}{\left\langle P_{rad} \right\rangle / \omega} = e_r \frac{2\left[\left\langle W_{e,non-rad} \right\rangle, \left\langle W_{m,non-rad} \right\rangle\right]_{Max}}{\left\langle P_{rad} \right\rangle / \omega} \approx \frac{1}{fractional\ BW_{3dB}} \quad (4.2)$$

where e_r is a radiation efficiency, ω is an angular frequency, P_{rad} is a radiated power, $W_e^{non-rad}$ is non-radiating electric energy, $W_m^{non-rad}$ is non-radiating magnetic energy, and $\langle \cdot \rangle$ means time average. This equation is a slight modification of that of Chu with generally accepted extensions. The radiation Q is proportional to the radiation efficiency and the amount of non-radiating energy, but as is well known to be inversely proportional to the fractional 3-dB bandwidth. Thus, radiation Q can be used to explain trade-offs between antenna size, bandwidth, and efficiency; note that only two of these can simultaneously be achieved close to the theoretical radiation Q limit. Due to this great advantage, the classic fundamental-limit theory focused on identifying the exact minimum amount of non-radiating energy or equivalently minimum radiation Q for an ideal dipole antenna case. Chu further extended his results to include higher order modes and also the trade-off with directivity that results from these modes.

In this section, the traditional approaches to obtain minimum radiation Q are reviewed. Implicit assumptions and errors in the approaches are explained based on what we have learned from the radiation and energy storage mechanism in Chapter 3.

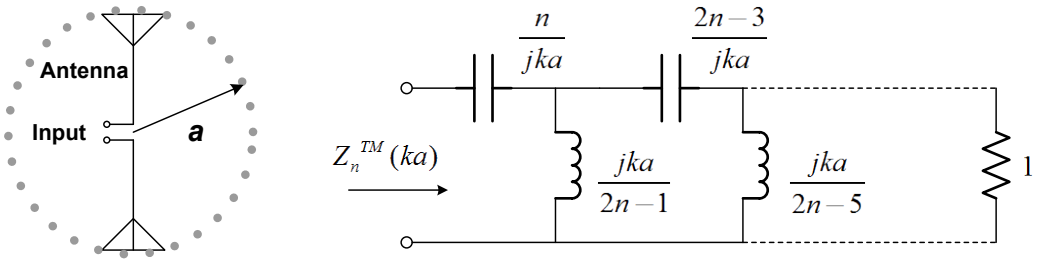
4.1.1 Circuit Approach

A year after Wheeler's paper, Chu published a comprehensive paper about the fundamental-limit theory on antenna size and performance [*Chu*, 1948]. He defined a sphere just enclosing an antenna structure (*antenna sphere*, see Figure 4.2a) and assumed that the antenna structure excites spherical modes. Because of the orthogonal properties of the spherical mode waves, the total energy outside the sphere is equal to the sum of the corresponding energies associated with each spherical mode. Similarly, the complex power leaving the surface of the sphere is identical to the sum of the complex powers associated with each spherical mode.

In order to identify the amount of non-radiating energy, Chu considered the radial wave impedance of the excited spherical modes. For spherical TM_{0n} modes, the normalized radial impedance is given by [*Chu*, 1948]

$$z_r^{TM_{0,n}}(ka) = \frac{1}{\eta_0}\frac{E_\theta^{TM_{0,n}}}{H_\phi^{TM_{0,n}}} = \frac{j\left(\hat{H}_n^{(2)}(ka)\right)'}{\hat{H}_n^{(2)}(ka)} \tag{4.3}$$

where η_0 is the free-space wave impedance, a is the radius of antenna sphere just enclosing the antenna and $\hat{H}_n^{(2)}(ka)$ is the spherical Hankel function of second kind (for $e^{j\omega t}$ time behavior). The Chu's equivalent circuit model of the normalized radial wave impedance for the spherical TM_{0n} modes is shown in Figure 4.2b.



(a) (b)

Figure 4.2 Chu's antenna model [*Chu*, 1948] – (a) Schematic diagram of a vertically polarized antenna and (b) Equivalent circuit of the normalized radial wave impedance for spherical TM_{0n} mode.

As an example, consider the spherical TM_{01} mode. The normalized radiation wave impedance is

$$z_r^{TM_{01}} = \frac{1}{j\omega(a/c_0)} + \frac{1}{\frac{1}{j\omega(a/c_0)} + \frac{1}{1}} \tag{4.4}$$

where c_0 is the speed of light. The equivalent circuit of (4.4) is shown in Figure 4.3a. We can observe that the input impedance of the equivalent circuit is capacitive for any ka. For $ka << 1$ (the antenna is very small or frequency is very low), the shunt inductor becomes electrically short. Thus, the real part of the input impedance becomes very small while the imaginary part becomes very capacitive. As is well known, this is the typical behavior for the input impedance of an electrically-small electric dipole antenna. For $ka >> 1$ (the antenna is very large or frequency is very high), the series capacitor becomes electrically short and the shunt inductor becomes electrically open. Thus, the input impedance of the equivalent circuit becomes the normalized free-space wave impedance (= 1), which can be seen as the input impedance of a traveling-wave antenna.

Since Chu intended to evaluate minimum radiation Q for multiple spherical-mode excitations, he considered that it is tedious to calculate the sum of the electric energies in the equivalent circuits. Thus, Chu considered an approximate simple series *RLC* resonant circuit. The approximate series *RLC* resonant circuit for $z_r^{TM_{01}}$ is shown in Figure 4.3b. The component values in the approximated circuit are found as

$$R_s = \frac{k^2a^2}{1+k^2a^2} \tag{4.5}$$

$$L_s = \frac{1}{\omega}\frac{ka}{\left(1+k^2a^2\right)^2} \tag{4.6}$$

$$C_s = \frac{1}{\omega}\frac{ka\left(1+k^2a^2\right)^2}{2k^2a^2+1} \tag{4.7}$$

where a portion of the shunt inductance is actually included as a capacitance effect in C_s. The power dissipated in the series resistor (P_r) and the electric energy stored in the series capacitor (W_e) can be written in terms of the current (I_s) in the circuit respectively, as

$$P_r = R_s \mid I_s \mid^2 = \frac{k^2a^2}{1+k^2a^2} \mid I_s \mid^2 \tag{4.8}$$

$$W_e = \frac{1}{2}\frac{\left|I_s\right|^2}{\omega^2 C_s} = \frac{1}{2\omega}\frac{\left(1+2k^2a^2\right)}{ka\left(1+k^2a^2\right)^2}\left|I_s\right|^2 \tag{4.9}$$

Since the spherical TM_{01} mode is the mode that an ideal dipole antenna excites, the electric energy is the dominant non-radiating energy. Using (4.2), the radiation Q is found as

$$Q_{Chu} = e_r \frac{2\left\langle W_{e,non-rad}\right\rangle}{\left\langle P_{rad}\right\rangle/\omega} = e_r \frac{1+2k^2a^2}{k^3a^3(1+k^2a^2)} \tag{4.10}$$

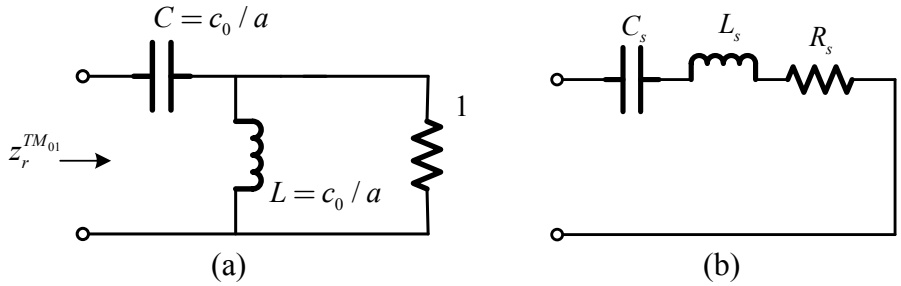


Figure 4.3 Chu's circuit representation of the normalized radial wave impedance for spherical TM_{01} mode – (a) Equivalent circuit and (b) Approximate equivalent circuit.

Later, McLean presented the radiation Q for the original equivalent circuit (Figure 4.3a), which is given as [*McLean*, 1996]

$$Q_{McLean} = e_r \left(\frac{1}{k^3 a^3} + \frac{1}{ka} \right) \tag{4.11}$$

Since any higher-order mode excitation increases the radiation Q, Chu and McLean's radiation Qs for the fundamental spherical mode (TM_{01}) were presented as a minimum radiation Q excitation for ideal electrical dipole antenna. In particular, (4.11) is often referred to as "Chu's exact result." As compared in Figure 4.4, both minimum radiation Qs of (4.10) and (4.11) approach $1/(k^3 a^3)$ as frequency or antenna size decreases. As ka increases, McLean's radiation Q becomes higher than that of Chu's approximate radiation Q.

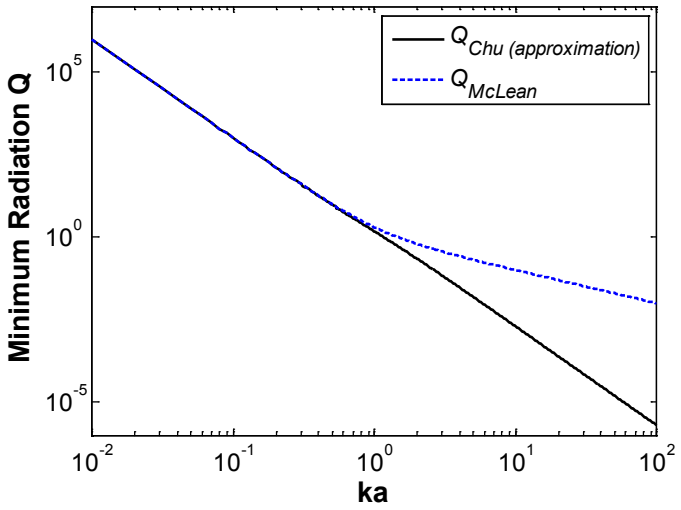


Figure 4.4 Minimum radiation Q assuming 100% radiation efficiency versus ka for Chu's approximation and McLean's results.

It may be worthwhile to comment on the angular wave impedance (Z_θ) that is not typically discussed in the literature. Chu and McLean's circuit approaches estimate the non-radiating energy directly, compared to the approach estimating non-radiating energy by subtracting radiating energy from the total energy [*R Collin and Rothschild*, 1964]. Because the angular wave impedance seems to be related to energy storage, the implication of the angular wave impedance needs to be addressed.

As discussed in Chapter 2, angular wave impedance could also be defined on the antenna sphere. For example, the normalized angular wave impedance for the spherical TM_{01} mode is found as

$$z_\theta^{TM_{0,1}} = -\frac{1}{\eta_0}\frac{E_r}{H_\phi} = -2\frac{\cot\theta}{j\omega(a/c_0)} \tag{4.12}$$

which is purely capacitive. Thus, there is no time average power in the angular direction. The corresponding capacitance is

$$C_\theta = 2\frac{1}{\omega}\frac{\cot\theta}{ka} \tag{4.13}$$

Notice that a net total angular capacitance on the antenna sphere is zero. Thus, the time average of the net total angular electric energy also would be zero. Thus, even though the net total angular electric energy is included in the calculation of the radiation Q, the results of (4.10) and (4.11) do not change.

The discussed circuit approach is simple and intuitive because the equivalent circuit analysis for the fundamental spherical mode giving the minimum radiation Q is almost trivial. However, we noticed that there is an excess phase delay in the energy storage and radiation process in Chapter 2. If the excess phase delay is considered, the resulting minimum radiation Q would be different from (4.11). The circuit viewpoint for minimum radiation Q, that includes this excess phase delay effect, is presented in Section 4.2.1.2, though the approximate from of Chu was one attempt to account for this excess delay.

4.1.2 The Frequency-Domain Approach

Collin and Rothschild calculated the minimum radiation-Q from the fields of spherical TM modes in frequency domain [*R Collin and Rothschild*, 1964]. In order to demonstrate the essence of their approach, consider the fields of the fundamental spherical TM_{01} mode with a source normalization factor ($Idl = 4\pi/k$) [*R F Harrington*, 2001]:

$$\vec{E}(r,\theta) = \hat{r}2\eta_0\frac{\cos\theta}{r}\left[-\frac{j}{kr}+\frac{1}{k^2r^2}\right]e^{-jkr} + \hat{\theta}\eta_0\frac{\sin\theta}{r}\left[1-\frac{j}{kr}-\frac{1}{k^2r^2}\right]e^{-jkr} \tag{4.14}$$

$$\vec{H}(r,\theta) = \hat{\phi}\frac{\sin\theta}{r}\left[1-\frac{j}{kr}\right]e^{-jkr} \tag{4.15}$$

The time average power through a spherical surface is found as

$$\left\langle P_{rad} \right\rangle = \Re\left\{ \frac{1}{2} \oint_S \vec{E} \times \vec{H}^* \cdot d\vec{s} \right\} = \eta_0 \frac{4\pi}{3} \tag{4.16}$$

The time-average energy velocity of radiating energy ($\left\langle v_{e,rad} \right\rangle$) can be defined as [*R E Collin*, 2001]

$$\left\langle v_{e,rad} \right\rangle = \frac{\left\langle P_{rad} \right\rangle}{\left\langle U_{rad} \right\rangle} \tag{4.17}$$

where $\left\langle U_{rad} \right\rangle$ is the time average of total radiating energy density per radial distance. Collin and Rothschild assumed that the time-average radiating energy velocity is constant and is equal to the speed of light. Thus, the time-average total radiating energy density per radial distance is given as

$$\left\langle U_{rad} \right\rangle = \frac{\left\langle P_{rad} \right\rangle}{c_0} = \mu_0 \frac{4\pi}{3} \tag{4.18}$$

In order to evaluate the radiation Q using (4.2), we need to indentify how much time-average electric energy density per radial distance ($\left\langle U_{e,rad} \right\rangle$) is included in $\left\langle U_{rad} \right\rangle$. Collin and Rothschild assumed that $\left\langle U_{e,rad} \right\rangle$ is identical to the time-average radiating magnetic energy per radial distance ($\left\langle U_{m,rad} \right\rangle$), but they did not provide any justification.

Assume that we can identify the radiating electric field ($\vec{E}_{rad}$) and magnetic field ($\vec{H}_{rad}$). Because only radial wave impedance is related to time-average real power, consider radial wave impedance (Z_r) for the radiating fields:

$$Z_r = \frac{E_{\theta,rad}}{H_{\phi,rad}} = \eta_0 \tag{4.19}$$

Then, it can easily be shown that the radiating electric energy density ($w_{e,rad}$) is identical to the radiating magnetic energy density ($w_{m,rad}$), i.e.

$$w_{e,rad} = \frac{\varepsilon_0}{2} \left| E_{\theta,rad} \right|^2 = \frac{\varepsilon_0}{2} \left| \sqrt{\frac{\mu_0}{\varepsilon_0}} H_{\phi,rad} \right|^2 = \frac{\mu_0}{2} \left| H_{\phi,rad} \right|^2 = w_{m,rad} \tag{4.20}$$

Therefore, for the spherical TM_{01} mode excitation case, we have using (4.18):

$$\left\langle U_{e,rad} \right\rangle = \left\langle U_{m,rad} \right\rangle = \frac{\mu_0}{2}\frac{4\pi}{3} \tag{4.21}$$

The time-average electric and magnetic energy densities per radial distance are found as

$$\left\langle U_e \right\rangle = \frac{1}{2}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\frac{\varepsilon_0}{2}\vec{E}\cdot\vec{E}^* r^2 \sin\theta d\theta d\phi = \frac{\mu_0}{2}\frac{4\pi}{3}\left(1+\frac{1}{k^2r^2}+\frac{3}{k^4r^4}\right) \tag{4.22}$$

$$\left\langle U_m \right\rangle = \frac{1}{2}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\frac{\mu_0}{2}\vec{H}\cdot\vec{H}^* r^2 \sin\theta d\theta d\phi = \frac{\mu_0}{2}\frac{4\pi}{3}\left(1+\frac{1}{k^2r^2}\right) \tag{4.23}$$

Thus, the time-average non-radiating energy densities per radian distance can be evaluated as the following:

$$\left\langle U_{e,non-rad} \right\rangle = \left\langle U_e \right\rangle - \left\langle U_{e,rad} \right\rangle = \frac{\mu_0}{2}\frac{4\pi}{3}\left(\frac{1}{k^2r^2}+\frac{3}{k^4r^4}\right) \tag{4.24}$$

$$\left\langle U_{m,non-rad} \right\rangle = \left\langle U_m \right\rangle - \left\langle U_{m,rad} \right\rangle = \frac{\mu_0}{2}\frac{4\pi}{3}\left(\frac{1}{k^2r^2}\right) \tag{4.25}$$

Obviously, $\left\langle U_{e,non-rad} \right\rangle$ is larger than $\left\langle U_{m,non-rad} \right\rangle$. The time-average non-radiating electric energy is found as

$$\left\langle W_{e,non-rad} \right\rangle = \int_a^{\infty}\left\langle U_{e,non-rad} \right\rangle dr = \eta_0\frac{2\pi}{3\omega}\left(\frac{1}{ka}+\frac{1}{k^3a^3}\right) \tag{4.26}$$

The corresponding radiation Q is

$$Q_{Collin} = e_r\frac{2\left\langle W_{e,non-rad} \right\rangle}{\left\langle P_{rad} \right\rangle / \omega} = e_r\left(\frac{1}{ka}+\frac{1}{k^3a^3}\right) \tag{4.27}$$

which is the same as Chu's "exact result."

The issue in Collin and Rothschild's work is that they assumed that the time-average energy velocity is constant and equal to the speed of light. As we observed in Chapter 2, the time-average energy velocity is not constant. Levis [*Levis*, 1957] and Rhodes [*Rhodes*, 1976] also noted that it is a hypothesis to use the speed of light for estimating the time-average radiating energy. In addition, we noticed from (4.25) that not all of the magnetic energy is used for radiation. In terms of minimum non-radiating energy viewpoint, it is hard to understand why all the magnetic energy is not used because the complex Poynting theorem [*R F Harrington*, 2001] implies that observable stored energy corresponds to the difference between electric and

magnetic energies inside a closed volume in the frequency domain. Thus, Collin and Rothschild's minimum radiation-Q result is questionable.

Later, McLean [*McLean*, 1996] reexamined the radiation-Q of the spherical TM_{01} mode. McLean considered the far-field terms, i.e. $1/r$ field terms, as the only terms contributing to the radiating energy. The time-average radiating electric and magnetic energy densities per radial distance from the $1/r$ field terms are identical to (4.21), giving the same minimum radiation Q that Collin and Rothschild found, i.e.

$$Q_{McLean} = e_r \left(\frac{1}{k^3 a^3} + \frac{1}{ka} \right) \tag{4.28}$$

which is also identical to the result of McLean's circuit analysis in (4.11).

Often, it is explained in the literature that McLean's result confirms that Collin and Rothschild's results. In a different sense, McLean's approach of taking the $1/r$ field terms for the radiating fields proves that Collin and Rothschild's approach using the constant speed of light for radiating energy velocity is based on a far-field approximation extrapolated to the near field. Thus, the minimum radiation Q result of (4.27) is an approximate result that is based on this far-field approximation.

4.1.3 The Time-Domain Approach

Grimes [*Grimes and Grimes*, 1999] recognized that the group velocity of the power flow in frequency-domain complex power notation exceeds the speed of light. He implied that there is a necessary phase delay and the phase delay cannot easily be observed with the phase notation in frequency domain. Based on the observation, Grimes claimed that the previous minimum radiation Q results of (4.10) and (4.27) are in error because the previous minimum radiation Q results did not reflect the phase delay. Thus, Grimes tried to obtain the minimum radiation-Q in the time domain.

In order to demonstrate the essential concept of Grimes' approach, consider Grimes' field expressions for the spherical TM_{01} mode:

$$\begin{aligned} \vec{E}(r,\theta) = \hat{r}2 &\left[-\frac{1}{k^2 r^2}\cos(\omega t - kr) - \frac{1}{k^3 r^3}\sin(\omega t - kr) \right] \cos\theta \\ -\hat{\theta} &\left[\frac{1}{k^2 r^2}\cos(\omega t - kr) - \left(\frac{1}{kr} - \frac{1}{k^3 r^3} \right) \sin(\omega t - kr) \right] \sin\theta \end{aligned} \tag{4.29}$$

$$\vec{H}(r,\theta) = -\hat{\phi}\frac{1}{\eta_0}\left[\frac{1}{k^2r^2}\cos(\omega t - kr) - \frac{1}{kr}\sin(\omega t - kr)\right]\sin\theta \tag{4.30}$$

where Grimes assumed a source normalization of $Idl = 4\pi/\eta_0/k^2$. The corresponding Poynting vector is found as

$$\begin{aligned}\vec{S}(r,\theta) &= \vec{E}(r,\theta)\times\vec{H}(r,\theta)\\ &= \hat{r}\frac{1}{2\eta_0}\left[\frac{1}{k^2r^2}\left[1-\mathrm{co}(\omega t - kr)\right] - \left(\frac{2}{k^3r^3} - \frac{1}{k^5r^5}\right)\sin 2(\omega t - kr) + \frac{2}{k^4r^4}\cos 2(\omega t - kr)\right]\sin^2\theta\\ &\quad + \hat{\theta}\frac{1}{2\eta_0}\left[\left(\frac{2}{k^3r^3} - \frac{2}{k^5r^5}\right)\sin 2(\omega t - kr) - \frac{4}{k^4r^4}\cos 2(\omega t - kr)\right]\sin\theta\cos\theta\end{aligned} \tag{4.31}$$

This equation differs from Grimes derivation [*Grimes and Grimes*, 1999] which was in an algebraic error. Thus, (4.31) is a corrected version of Grimes' result.

Using the differential form of the Poynting theorem [*R F Harrington*, 2001] in free space, Grimes tired to relate the Poynting vector representing power density and the non-radiating energy density ($w_{non-rad}$), i.e.

$$\nabla\cdot\vec{S} = -\frac{\partial}{\partial t}w_{non-rad} \tag{4.32}$$

Grimes integrated (4.32) over time to obtain

$$\begin{aligned}w_{non-rad} = \frac{\varepsilon_0}{4}\Bigg\{K + \Bigg[\left(\frac{4}{k^4r^4} - \frac{4}{k^6r^6}\right)\cos 2(\omega t - kr) + \frac{8}{k^5r^5}\sin 2(\omega t - kr)\Bigg]\cos^2\theta\\ - \frac{1}{k^6r^6}\mathrm{c}\cos 2(\omega t - kr)\sin^2\theta\Bigg\}\end{aligned} \tag{4.33}$$

where *K* is a time-independent integration constant. Grimes used the following condition in order determine *K*:

$$w_{non-rad} \geq 0 \tag{4.34}$$

which implies that the non-radiating energy density is equal to or greater than zero at any given time. Thus, the minimum non-radiating energy can be obtained after solving $w_{non-rad} = 0$ for *K*. However, there exist multiple solutions for *K* that satisfies the (4.34) condition. Thus, it is not straightforward to obtain *K* for minimum non-radiating energy density.

Grimes claimed that the time-independent constant K giving the minimum non-radiating energy is found as

$$K = \left(\frac{4}{k^4 r^4} + \frac{4}{k^6 r^6} \right) \cos^2 \theta + \frac{1}{k^6 r^6} \sin^2 \theta \tag{4.35}$$

Note that the non-radiating energy using (4.35) is not a minimum. We use (4.35) to obtain Grimes' minimum radiation Q result in following steps, not corrected in this work. The corresponding minimum non-radiating energy is found as

$$\begin{aligned} W_{non-rad} &= \int_{r=a}^{\infty} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \left[w_{non-rad} \right]_{\min} r^2 \sin\theta d\theta d\phi dr \\ &= \frac{2\pi\varepsilon_0}{3k^3} \left[-\frac{3}{3k^3 a^3} \left[1 - \cos 2(\omega t - kr) \right] - \frac{2}{kr} \left[1 + \cos 2(\omega t - kr) \right] - \frac{4}{2k^2 r^2} \sin 2(\omega t - kr) \right]_a^{\infty} \end{aligned} \tag{4.36}$$

When Grimes attempted to evaluate the minimum non-radiating energy using (4.35) [*Grimes and Grimes*, 1999], Grimes made an algebraic error. Thus, (4.36) is a corrected version of Grimes' result. The peak of (4.36) is found as

$$\left[W_{non-rad} \right]_{Peak} = \frac{4\pi\varepsilon_0}{3k^3} \left[\frac{1}{2k^3 a^3} \left(1 + \sqrt{4k^4 a^4 + 1} \right) + \frac{1}{ka} \right] \tag{4.37}$$

which is identical to what Grimes found. The time average power is obtained as

$$\left\langle P_{rad} \right\rangle = \left\langle \int_{r=a}^{\infty} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} S_r r^2 \sin\theta d\theta d\phi dr \right\rangle = \frac{4\pi}{3k^2 \eta_0} \tag{4.38}$$

The resulting minimum radiation Q is given as

$$Q_{Grimes} = e_r \frac{\left[W_{e,non-rad} + W_{m,non-rad} \right]_{Peak}}{\left\langle P_{rad} \right\rangle / \omega} = \frac{1}{2k^3 a^3} (1 + \sqrt{1 + 4k^4 a^4}) + \frac{1}{ka} \tag{4.39}$$

The Grimes' minimum radiation Q is compared to Chu's approximate and McLean's results in Figure 4.5. The Grimes' minimum radiation Q values also approach $1/(k^3 a^3)$ as ka decreases as do the Chu and McLean results. But the Grimes' minimum radiation Q is higher than Chu's and McLean's results for large ka.

Grimes' observation of group velocity exceeding the speed of light implies that there is dispersion in wave propagation [*Brillouin*, 1960]. Often group velocity is used for narrowband signal propagation instead of energy velocity because, sometimes, it is not straightforward to evaluate energy velocity. Mathematically, group velocity can vary from negative infinite to

positive infinite, but the velocity of wave propagation is finite and its speed is limited to the speed of light, as is well known. Thus, energy velocity should be used to represent the speed of wave propagation whenever group velocity exceeds the speed of light. Grimes did not extend his work to explore the complex power flow with energy velocity, but his observation on group velocity exceeding the speed of light was an important finding to correct the classic fundamental-limit theory.

Grimes' approach using the energy condition, i.e. energy should be equal to or greater than zero, was a new approach in order to obtain minimum non-radiation energy. However, he utilized the differential form of Poynting theorem that is expressed with power density and non-radiating energy density. As a result, the obtained non-radiating energy through the differential form of Poynting theorem seems to be a local minimum, rather than a global.

Perhaps, Grimes' work has not been referenced as much as it could be because this result does not agree with the exact Chu's result. Grimes' algebraic errors in [*Grimes and Grimes*, 1999] also reduces the credibility of Grimes' work.

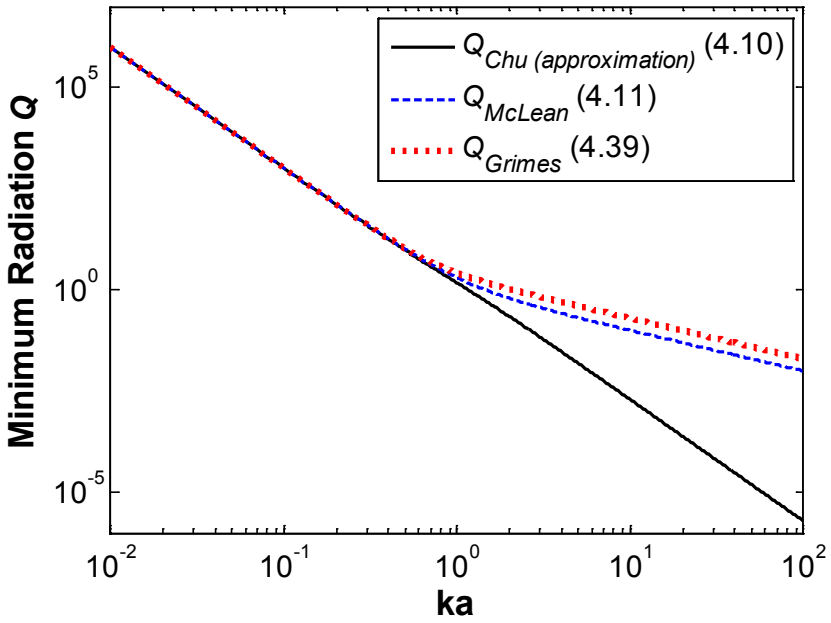


Figure 4.5 Minimum radiation Q assuming 100% radiation efficiency versus ka for Grimes' result, compared to Chu's approximated and McLean's results.

4.2 Extended Fundamental-Limit Theory

Recently, Davis *et al.* [*Davis et al.*, 2011] identified errors in previous fundamental limit theory and presented a new limit of radiation-Q using the fields of the spherical TM_{01} mode that

represent an ideal electric dipole. Their approach was based on the observation of excess delay in radiation process and the corresponding change in energy velocity. Initially, they examined the total power that was obtained from integrating Poynting vector over a closed surface (s) at r in time domain:

$$\begin{aligned} P(r,t) &= \oint_S \vec{E}\times\vec{H}\cdot d\vec{s} \\ &= \langle P_{rad}\rangle\left[1+\left(1-\frac{2}{k^2r^2}\right)\cos 2(\omega t-kr)+\left(\frac{2}{kr}-\frac{1}{k^3r^3}\right)\sin 2(\omega t-kr)\right] \end{aligned} \tag{4.40}$$

where $\langle P_{rad}\rangle$ is the time average power ($=4\pi\eta_0/3$) with the source normalization of $Idl = 4\pi/k$. Because the total power expression of (4.40) does not show a clear separation between real and reactive power terms, Davis *et al.* rearranged (4.40) as

$$P(r,t) = \langle P_{rad}\rangle\left[1+\cos 2(\omega t-kr-\zeta)+\frac{1}{k^3r^3}\sin 2(\omega t-kr-\zeta)\right] \tag{4.41}$$

where $\zeta = \tan^{-1}(1/kr)$ is the excess phase. The first and second terms of (4.41) are related to the radiated power [$\cos^2(\cdot)$], and the third term is related to reactive power [$\sin(\cdot)\cos(\cdot)$]. These radiated and reactive powers are 90° out of phase, which is consistent with circuit definitions of real and reactive power.

In the following sub-section, Davis *et al.*'s approach is reviewed and extended to obtain the minimum radiation Q of arbitrary-polarized antennas. In addition, we discuss the origin of the minimum non-radiating energy. Evaluation and measurement approaches for antenna radiation Q are briefly addressed.

4.2.1 Fundamental Spherical Modes (Spherical TM_{01} and TE_{01} Modes)

4.2.1.1 The Energy Velocity Viewpoint

The time-harmonic fields including the excess phase ($\zeta = \tan^{-1}(1/kr)$) for the spherical TM_{01} mode are given as [*Davis et al.*, 2011]

$$\begin{aligned} \vec{E}(r,t) &= \hat{r}2\eta_0\frac{\cos\theta}{r}\frac{\sqrt{k^2r^2+1}}{k^2r^2}\sin(\omega t-kr-\zeta) \\ &\quad+\hat{\theta}\eta_0\frac{\sin\theta}{r}\frac{kr}{\sqrt{k^2r^2+1}}\left[\cos(\omega t-kr-\zeta)+\frac{1}{k^3r^3}\sin(\omega t-kr-\zeta)\right] \end{aligned} \tag{4.42}$$

$$\vec{H}(r,t)=\hat{\phi}\frac{\sin\theta}{r}\frac{\sqrt{k^2r^2+1}}{kr}\cos\left(\omega t-kr-\zeta\right) \tag{4.43}$$

However, as we noted from the observation of Hertz [*Hertz*, 1893] in Chapter 2, the phase behavior for electric and magnetic fields are different in near-field region. This implies that the total power and field expressions could be rearranged in a different form with a different excess delay. In this sense, expressions of (4.41), (4.42), and (4.43) are the rearranged form referenced to the phase of magnetic field. The alternative field and power forms referenced to the *phase of angular electric field* are found as

$$\begin{aligned}\vec{E}(r,t)=\hat{r}2\eta_0\frac{kr}{\sqrt{k^4r^4-k^2r^2+1}}\left[\cos\left(\alpha-\zeta-\gamma\right)+\frac{1}{k^3r^3}\sin\left(\alpha-\zeta-\gamma\right)\right]\\ +\hat{\theta}\eta_0\frac{\sin\theta}{r}\frac{\sqrt{k^4r^4-k^2r^2+1}}{k^2r^2}\cos(\alpha-\zeta-\gamma)\end{aligned} \tag{4.44}$$

$$\vec{H}(r,t)=\hat{\phi}\frac{\sin\theta}{r}\frac{k^2r^2}{\sqrt{k^4r^4-k^2r^2+1}}\left[\cos(\alpha-\zeta-\gamma)-\frac{1}{k^3r^3}\sin(\alpha-\zeta-\gamma)\right] \tag{4.45}$$

$$P\left(r,t\right)=P_{av}\left[1+\cos2\left(\omega t-kr-\zeta-\gamma\right)-\frac{1}{k^3r^3}\sin2\left(\omega t-kr-\zeta-\gamma\right)\right] \tag{4.46}$$

where $\gamma=\tan^{-1}(k^3r^3)$ is an additional excess phase.

In order to identify which phase reference has the correct physical meaning, we first consider the phase velocity of a radiating energy packet in the radial direction that is referenced to the magnetic field phase in (4.43) first, which is found as [*Davis et al.*, 2011]

$$v_{p,H_\phi}=\left(1+\frac{1}{k^2r^2}\right)c_0 \tag{4.47}$$

where c_0 is the speed of light. Thus, we can observe that $v_{p,H}$ can exceed c_0, which mean that the energy packet is dispersive. On the other hand, the group velocity is found as [*Davis et al.*, 2011]

$$v_{g,H_\phi}=\left(1+\frac{1-k^2r^2}{k^4r^4+3k^2r^2}\right)c_0 \tag{4.48}$$

which suggests that the group velocity $\vec{v}_{g,H}$ obtained from the phase can exceed c_0 or be lower than c_0. Thus, $\vec{v}_{g,H}$ does not seem to be a physical quantity and also cannot represent the energy

velocity. The energy velocity ($v_{e,H}$) can be obtained from the following relationship: [*Brillouin*, 1960; *R E Collin*, 2001; *Ramo et al.*, 1994]

$$v_{e,H_\phi} = \frac{c_0^2}{v_{p,H_\phi}} = \left(1 - \frac{1}{k^2r^2 + 1}\right)c_0 \tag{4.49}$$

As shown in Figure 4.6a, the energy velocity v_{e,H_ϕ} becomes zero for $kr = 0$ while it approaches c_0 as kr increases, meaning that the energy packet is accelerated until it reaches to the speed of light.

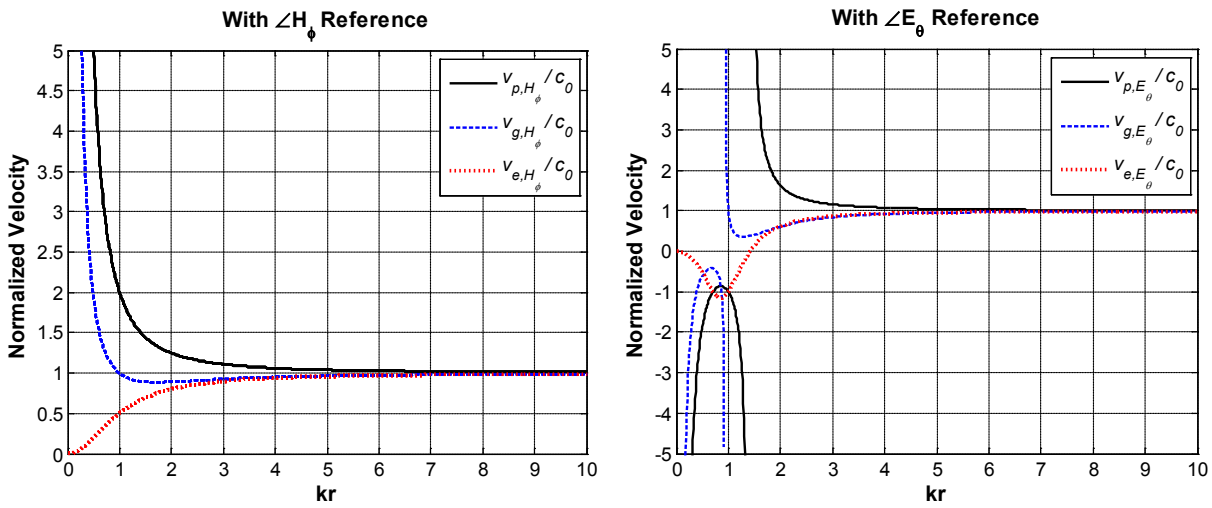
In a similar way, the phase, group, and energy velocities referenced to the angular electric field are found as

$$v_{p,E_\theta} = \left[1 + \frac{k^2r^2 + 1}{k^2r^2\left(k^2r^2 - 2\right)}\right]c_0 \tag{4.50}$$

$$v_{g,E_\theta} = \left[1 - \frac{k^6r^6 + 4k^4r^4 - 4k^2r^2 - 1}{k^2r^2\left(k^6r^6 - k^4r^4 + 7k^2r^2 - 6\right)}\right]c_0 \tag{4.51}$$

$$v_{e,E_\theta} = \left[1 - \frac{k^2r^2 + 1}{k^4r^4 - k^2r^2 + 1}\right]c_0 \tag{4.52}$$

These velocities are plotted in Figure 4.6b. The phase and group velocities have a singularity at $ka = \sqrt{2}$ and $ka \approx 0.933$, respectively. For $ka < 1$, all these velocities referenced to the angular electric fields are negative. Also, the magnitude of the energy velocity v_{e,E_θ} exceeds the speed of light for $ka < 1$. Thus, the velocities referenced to the angular electric fields do not represent physical phenomenon.



(a) (b)

Figure 4.6 Normalized phase, group, and energy velocities for spherical TM_{01} mode – (a) Referenced to H_ϕ phase and (b) Referenced to E_θ phase.

Using (4.17), the time-average radiating energy density per radial distance is found as

$$\langle U_{rad} \rangle = \frac{\langle P_{rad} \rangle}{\langle v_{e,H_\phi} \rangle} = \mu_0 \frac{4\pi}{3} \frac{k^2 r^2 + 1}{k^2 r^2} \tag{4.53}$$

Because the amount of time-average electric and magnetic energy densities per radial distance is the same, we have

$$\langle U_{e,rad} \rangle = \langle U_{m,rad} \rangle = \frac{\langle U_{rad} \rangle}{2} = \mu_0 \frac{1}{2} \frac{4\pi}{3} \left(1 + \frac{1}{k^2 r^2} \right) \tag{4.54}$$

The total electric and magnetic energy densities per radial distance are found as

$$\begin{aligned} U_e &= \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \frac{\varepsilon_0}{2} \vec{E} \cdot \vec{E} r^2 \sin\theta d\theta d\phi \\ &= \mu_0 \frac{4\pi}{3} \frac{1}{2k^4 r^4 \left(k^2 r^2 + 1 \right)} \left\{ \left[k^2 r^2 \left(k^2 r^2 + 1 \right) + 3 \right] \left(k^2 r^2 + 1 \right) \right. \\ &\quad \left. + \left(k^6 r^6 - 2k^4 r^4 - 4k^2 r^2 - 3 \right) \cos 2\left(\omega t - kr - \varsigma \right) + 2k^3 r^3 \sin 2\left(\omega t - kr - \varsigma \right) \right\} \end{aligned} \tag{4.55}$$

$$\begin{aligned} U_m &= \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \frac{\mu_0}{2} \vec{H} \cdot \vec{H}^* r^2 \sin\theta d\theta d\phi \\ &= \mu_0 \frac{4\pi}{3} \frac{k^2 r^2 + 1}{2k^2 r^2} \left[1 + \cos 2(\omega t - kr) \right] \end{aligned} \tag{4.56}$$

If we take time average on (4.55) and (4.56), the sinusoidal terms become zero. Thus, the resulting total time-average electric and magnetic energy densities per radial distance are

$$\langle U_e \rangle = \mu_0 \frac{1}{2} \frac{4\pi}{3} \left(1 + \frac{1}{k^2 r^2} + \frac{3}{k^4 r^4} \right) \tag{4.57}$$

$$\langle U_m \rangle = \mu_0 \frac{1}{2} \frac{4\pi}{3} \left(1 + \frac{1}{k^2 r^2} \right) \tag{4.58}$$

which are the same as (4.22) and (4.23) in Collin and Rothschild's approach. Thus, the time average non-radiating electric and magnetic energy densities per radian distance can be evaluated as follows:

$$\langle U_{e,non-rad} \rangle = \langle U_e \rangle - \langle U_{e,rad} \rangle = \mu_0 \frac{2\pi}{k^4 r^4} \tag{4.59}$$

$$\langle U_{m,non-rad} \rangle = \langle U_m \rangle - \langle U_{m,rad} \rangle = 0 \tag{4.60}$$

Compared to Collin and Rothschild's approach, the time average non-radiating magnetic energy density per radial distance is zero. Thus, the entire magnetic energy is used for radiation. Considering the complex Poynting theorem, this would give a minimum non-radiating energy.

The time-average non-radiating electric energy is found by integrating (4.59) from $r = a$ to ∞:

$$\langle W_{e,non-rad} \rangle = \int_a^\infty \langle U_{e,non-rad} \rangle \, dr = \eta_0 \frac{2\pi}{3\omega} \left(\frac{1}{k^3 a^3} \right) \tag{4.61}$$

Then, the resulting minimum radiation Q for spherical TM_{01} mode is given as

$$Q_{VTAG} = e_r \frac{2\langle W_{e,non-rad} \rangle}{\langle P_{rad} \rangle / \omega} = e_r \frac{1}{k^3 a^3} \tag{4.62}$$

which is what Davis *et al.* found in the recent work [*Davis et al.*, 2011]. As shown in Figure 4.7, Q_{VTAG} is lower than previous results for $ka > 1$ while Q_{VTAG} also equals $1/(k^3 a^3)$ as ka decreases. Interestingly, Davis *et al.* showed that Chu's approximate minimum radiation Q can be obtained when the group velocity (4.48) is used to estimate radiating energy density.

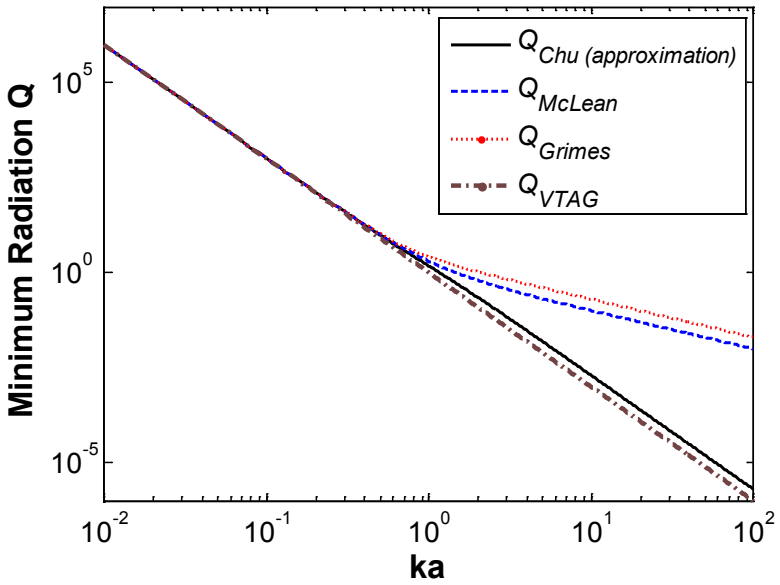


Figure 4.7 Minimum radiation Q assuming 100% radiation efficiency versus ka for Davis *et al.*'s result, compared to Chu's approximated, McLean's, and Grimes' results [*Davis et al.*, 2011]. Copyright © 2011 IET.

In terms of the spherical TE_{01} mode representing the ideal magnetic dipole (loop) antenna, it is a dual to the spherical TM_{01} mode. Thus, the entire time-average electric energy of the spherical TE_{01} mode is used for radiation. The time-average radiated power is the same as the spherical TM_{01} mode case. The time-average magnetic non-radiating energy is dominant for the spherical TE_{01} mode case. As a result, the amount of the time-average magnetic non-radiating energy is the same as the time-average electric non-radiating energy of the spherical TM_{01} mode case. The minimum radiation Q for spherical TE_{01} mode becomes the same as the minimum radiation Q for spherical TM_{01} mode:

$$Q_{VTAG}^{TE_{01}} = Q_{VTAG}^{TM_{01}} = e_r \frac{1}{k^3 a^3} \tag{4.63}$$

Since the amounts of the radiating and non-radiating energies were obtained during the minimum radiation Q derivation process, we are able to identify what energy (angular or radial) contributed to the radiation and non-radiation energies. For the spherical TM_{01} mode, the time-average radial and angular electric energy densities per radial distance are found as

$$\left\langle U_{e,r} \right\rangle = \mu_0 \frac{1}{2} \frac{4\pi}{3} \left(\frac{2}{k^2 r^2} + \frac{2}{k^4 r^4} \right) \tag{4.64}$$

$$\left\langle U_{e,\theta}\right\rangle = \mu_0 \frac{1}{2}\frac{4\pi}{3}\left(1-\frac{1}{k^2r^2}+\frac{1}{k^4r^4}\right) \tag{4.65}$$

In order to compare $\left\langle U_{e,r}\right\rangle$ and $\left\langle U_{e,non-rad}\right\rangle$, consider the difference:

$$\left\langle U_{e,r}\right\rangle - \left\langle U_{e,non-rad}\right\rangle = \mu_0 \frac{1}{2}\frac{4\pi}{3}\frac{2k^2r^2-1}{k^4r^4} \tag{4.66}$$

which suggests that $\left\langle U_{e,r}\right\rangle$ becomes smaller than $\left\langle U_{e,non-rad}\right\rangle$ for $kr<1/\sqrt{2}$. Thus, some of $\left\langle U_{e,non-rad}\right\rangle$ would be contributed from $\left\langle U_{e,\theta}\right\rangle$. Similarly, we have

$$\left\langle U_{e,\theta}\right\rangle - \left\langle U_{e,non-rad}\right\rangle = \mu_0 \frac{1}{2}\frac{4\pi}{3}\left(1-\frac{1}{k^2r^2}-\frac{2}{k^4r^4}\right) \tag{4.67}$$

which suggests that $\left\langle U_{e,\theta}\right\rangle$ becomes smaller than $\left\langle U_{e,non-rad}\right\rangle$ for $kr<\sqrt{2}$, and some of the $\left\langle U_{e,non-rad}\right\rangle$ would be contributed from $\left\langle U_{e,r}\right\rangle$. Plots for differences given by (4.66) and (4.67) are shown in Figure 4.8 versus *kr*. It is obvious that both the radial and angular electric energies contribute to the non-radiating energy.

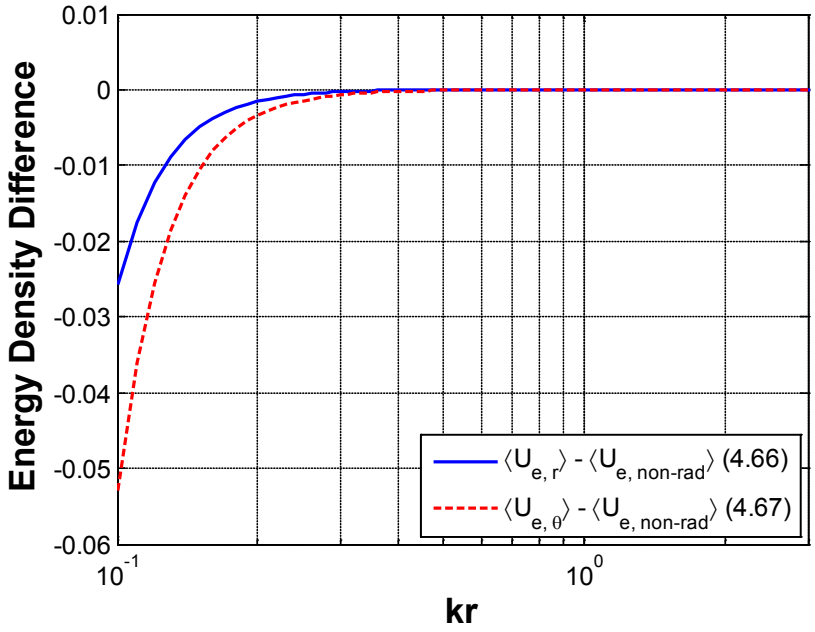


Figure 4.8 Difference between the time-average radial and angular electric energy densities per radial distance and the time-average electric non-radiating energy.

On the other hand, consider the difference between the time-average radiating radial electric energy and the time-average radiating energy

$$\left\langle U_{e,r} \right\rangle - \left\langle U_e^{rad} \right\rangle = \mu_0 \frac{1}{2} \frac{4\pi}{3} \left[\frac{2 + k^2 r^2 - k^4 r^4}{k^4 r^4} \right] \tag{4.68}$$

which suggests that $\left\langle U_{e,r} \right\rangle$ becomes smaller than $\left\langle U_e^{rad} \right\rangle$ for $kr > \sqrt{2}$. Thus, some of $\left\langle U_e^{rad} \right\rangle$ was contributed from $\left\langle U_{e,\theta} \right\rangle$. Comparing $\left\langle U_{e,\theta} \right\rangle$ and $\left\langle U_{e,rad} \right\rangle$, we have

$$\left\langle U_{e,\theta} \right\rangle - \left\langle U_{e,rad} \right\rangle = \mu_0 \frac{1}{2} \frac{4\pi}{3} \left(\frac{1 - 2k^2 r^2}{k^4 r^4} \right) \tag{4.69}$$

which suggests that $\left\langle U_{e,\theta} \right\rangle$ becomes smaller than $\left\langle U_{e,rad} \right\rangle$ for $kr > 1/\sqrt{2}$. Thus, some of $\left\langle U_{e,rad} \right\rangle$ should be contributed from $\left\langle U_{e,r} \right\rangle$. Plots for differences in (4.68) and (4.69) are shown in Figure 4.9 versus *kr*. Thus, it is clear that $\left\langle U_{e,rad} \right\rangle$ is contributed from both the radial and angular radiating energy densities per radial distance, not just from the angular radiating energy.

Combining all these observations through the energy density difference analysis in (4.66) through (4.69), conclude that both energy storage and radiation occur in both angular and radial directions. As we explored the energy radiation and storage mechanism in Chapter 2, the actual radiated energy was seen as an energy leakage during the absorbing and re-emitting process (circulation) between the antenna terminals and the virtual boundaries including radian sphere and causal surfaces. The trapped time during the circulation corresponds to the excess delay. Thus, the time-average radial energy velocity is not equal to the speed of light all the way from the antenna sphere to an infinite distance.

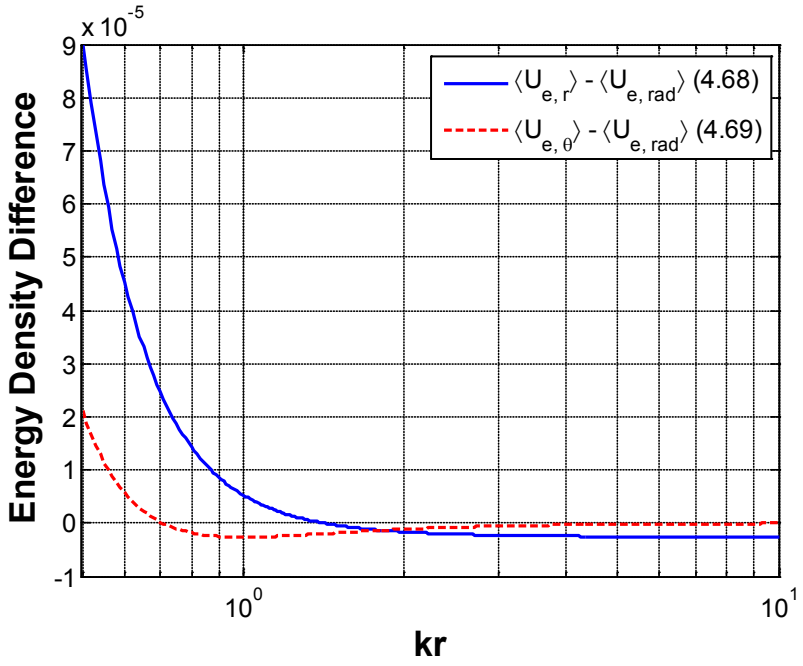


Figure 4.9 Difference between the time-average radial and angular electric energy densities per radial distance and the time-average electric radiating energy.

4.2.1.2 The Non-Negative Energy Condition Viewpoint

As pointed out earlier, Grimes tried to use the condition that energy is equal to or greater than zero to obtain a minimum non-radiating energy, which was a new approach. However, he used power and energy densities in the derivation process. It appears that Grimes found a local minimum of the non-radiating energy. In order to obtain a global minimum of non-radiating energy, we can utilize the same positive energy condition on the non-radiating energy, which is associated with reactive power.

The reactive power term for the spherical TM_{01} mode was recognized in the rearranged total power form of (4.41) [*Davis et al.*, 2011]. The relationship between the reactive power and the change of non-radiating energy about an antenna sphere can be written as

$$P_{reactive} = -\frac{\partial}{\partial t} W_{non-rad} \tag{4.70}$$

Integrating (4.70) from $r = a$ to ∞ yields

$$W_{non-rad} = \eta_0 \frac{4\pi}{3}\left[C - \frac{1}{2\omega k^3 a^3} \cos 2(\omega t - ka - \zeta_a) \right] \tag{4.71}$$

where C is a constant of integration. Because energy should not be negative for at given time, we have

$$C \geq \frac{1}{2\omega k^3 a^3} \tag{4.72}$$

Thus, the minimum non-radiating energy can be written as

$$W_{non-rad,\min} = \eta_0 \frac{4\pi}{3} \frac{1}{2\omega k^3 a^3} \left[1 - \cos 2(\omega t - ka - \zeta_a)\right] \tag{4.73}$$

Using (4.2), the minimum radiation Q is found as

$$Q_{rad,\min} = e_r \frac{\left[W_{non-rad,\min}\right]_{Peak}}{\left\langle P_{rad} \right\rangle / \omega} = e_r \frac{1}{k^3 a^3} \tag{4.74}$$

which is identical to Davis *et al.*'s result.

The non-negative energy condition can also be applied to Chu's equivalent circuit model. In order to demonstrate the concept, consider Chu's circuit model of normalized radial wave impedance for spherical the TM_{01} mode in Figure 4.10.

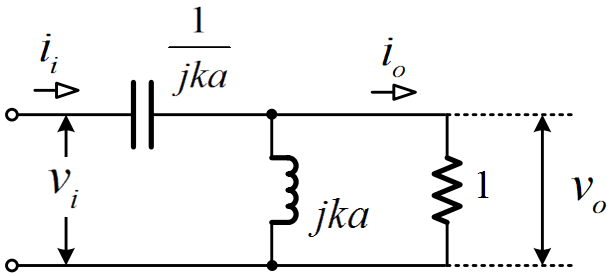


Figure 4.10 Chu's circuit representation of the normalized radial wave impedance for spherical TM_{01} mode in order to demonstrate the non-negative energy condition approach.

Assume that we have an input current $i_i = \cos(\omega t)$. The sum of currents on the inductor and resistor is the same as the input current:

$$i_i = v_o + \frac{c_0}{a} \int_t v_o(t) dt \tag{4.75}$$

where c_0 is the speed of light. If the excess phase (ζ) is considered, the output voltage can be written as

$$v_o = V_0 \cos(\omega t - \zeta) \tag{4.76}$$

where V_o is a peak voltage. Substituting (4.76) into (4.75) gives the following equation:

$$i_i = \cos(\omega t) = V_0 \left\{ \left[\cos(\zeta) - \frac{1}{ka} \sin(\zeta) \right] \cos(\omega t) + \left[\sin(\zeta) + \frac{1}{ka} \cos(\zeta) \right] \sin(\omega t) \right\} \quad (4.77)$$

Because $\cos(\omega t)$ and $\sin(\omega t)$ are orthogonal in time, the solution of (4.77) provides

$$\tan(\zeta) = -\frac{1}{ka} \quad (4.78)$$

$$V_o = -\frac{ka}{\sqrt{1+k^2a^2}} \quad (4.79)$$

Then, the input and output voltages are given as, respectively

$$v_i = \frac{k^2a^2}{1+k^2a^2} \left[\cos(\omega t) + \frac{1}{k^3a^3} \sin(\omega t) \right] \quad (4.80)$$

$$v_o = \frac{k^2a^2}{1+k^2a^2} \left[\cos(\omega t) - \frac{1}{ka} \sin(\omega t) \right] \quad (4.81)$$

The input power is

$$P_i = \langle P \rangle \left[1 + \cos 2(\omega t) + \frac{1}{k^3a^3} \sin 2(\omega t) \right] \quad (4.82)$$

where $\langle P \rangle$ is average power ($= 0.5k^2a^2 / (1+k^2a^2)$). Similar to the field approach of (4.41), the first and second terms of (4.82) are related to the radiated power, and the third term is related to reactive power ($P_{reactive}$).

The output power is found as

$$P_o = \langle P \rangle \left[1 + \cos 2(\omega t - \zeta) \right] \quad (4.83)$$

Thus, $\langle P \rangle$ in both the input and output power expressions indicate conservation of power. The excess phase ζ in (4.83) indicates the output power is shifted in time. In addition, there is no reactive power term is in (4.83).

Similar to the field approach, the reactive power ($P_{reactive}$) term (the third term) in (4.82) is related to the non-radiating energy ($W_{non-rad}$):

$$-\frac{\partial}{\partial t} W_{non-rad} = P_{reactive} \quad (4.84)$$

After integrating (4.84) over time from $r = a$ to ∞, we have

$$W_{non-rad} = \langle P \rangle \left[C - \frac{1}{2\omega k^3 a^3} \cos(2\omega t) \right] \tag{4.85}$$

where C is a constant of integration. For the non-negative energy condition, C needs to be equal to or greater than $1/\left(2\omega k^3 a^3\right)$. After plugging $C = 1/\left(2\omega k^3 a^3\right)$ in (4.85), we find the peak of the minimum non-radiating energy as $\langle P \rangle / \omega k^3 a^3$. Using (4.2), the resulting minimum radiation Q from the Chu's circuit analysis is found as

$$Q_{Chu-Circuit,VTAG} = e_r \frac{1}{k^3 a^3} \tag{4.86}$$

which is the same result as (4.62) of the field approach.

In this subsection, it has been demonstrated that the approach using the non-negative energy condition can be useful to obtain a minimum radiation Q for both field and circuit representations of spherical mode.

4.2.1.3 The Static Energy Viewpoint

The fundamental-limit theory indicates that not all energy outside an antenna can contribute to far-field radiation. Some non-radiating energy remains in the vicinity of an antenna interacting with the antenna and limits the antenna bandwidth or radiation efficiency. The question is where non-radiating energy originates. An intuitive answer would be that the static fields contribute to the non-radiating energy. Then, how much of the static energy is included in the amount of the identified minimum non-radiating energy. In order to answer to these questions, we investigate a static electric dipole problem.

Consider the static electric dipole shown in Figure 4.11. With $e^{j\omega t}$ time dependence, the conservation of charge requires that [*R F Harrington*, 2001]

$$I = -j\omega q \tag{4.87}$$

For an infinitesimal dipole or point source, we have

$$\lim_{\substack{I\to\infty \\ dl\to 0}} Idl = -j\omega \lim_{\substack{q\to\infty \\ dl\to 0}} qdl \tag{4.88}$$

Notice that both Idl and qdl are finite in the limit sense. The static potential for Figure 4.11b can be written as [*Cheng*, 1989]

$$\Phi(r,\theta,\phi)=\frac{1}{4\pi\varepsilon}\left(\frac{q}{\left|\vec{r}-\hat{z}dl/2\right|}-\frac{q}{\left|\vec{r}+\hat{z}dl/2\right|}\right) \tag{4.89}$$

If $dl \ll r$, we have

$$\left|\vec{r}-\hat{z}dl/2\right| \approx r-\frac{dl}{2}\cos\theta \tag{4.90}$$

$$\left|\vec{r}+\hat{z}dl/2\right| \approx r+\frac{dl}{2}\cos\theta \tag{4.91}$$

Thus, the potential is approximated as

$$\Phi(r,\theta,\phi)\approx\frac{q}{4\pi\varepsilon}\left(\frac{1}{r-\frac{dl}{2}\cos\theta}-\frac{1}{r+\frac{dl}{2}\cos\theta}\right)=\frac{qdl}{4\pi\varepsilon}\left[\frac{\cos\theta}{r^2-\left(\frac{dl}{2}\cos\theta\right)^2}\right] \tag{4.92}$$

For the case of a point-source case ($q\to\infty$ and $dl\to 0$) in (4.88), we have

$$\Phi(r,\theta,\phi)=\frac{qdl}{4\pi\varepsilon_0}\frac{\cos\theta}{r^2} \tag{4.93}$$

which is exact due in the limit process. The static potential of the spherical TM_{01} mode corresponds to (4.93). Notice that (4.93) is not spherically symmetric, i.e. not isotropic.

The corresponding static electric field is found from

$$\vec{E}=-\nabla\Phi=\frac{qdl}{4\pi\varepsilon_0}\frac{\left(\hat{r}2\cos\theta+\hat{\theta}\sin\theta\right)}{r^3} \tag{4.94}$$

Thus, the static electric field of the point source is inversely proportional to the cube of the distance, compared to the square of distance for the static potential case.

Using the same source normalization condition that we used in (4.14), we have

$$\vec{E}_{static}=-\eta_0 k\frac{\left[\hat{r}2\cos\theta+\hat{\theta}\sin\theta\right]}{k^3r^3} \tag{4.95}$$

The corresponding time-average electric energy density per radial distance is found as

$$\left\langle U_{e,static}\right\rangle=\frac{1}{2}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\frac{1}{2}\varepsilon_0\left|\vec{E}_{static}\right|r^2\sin\theta d\theta d\phi=\frac{2\pi\mu_0}{k^4r^4} \tag{4.96}$$

which is the same as (4.59) of the field approach using the energy velocity.

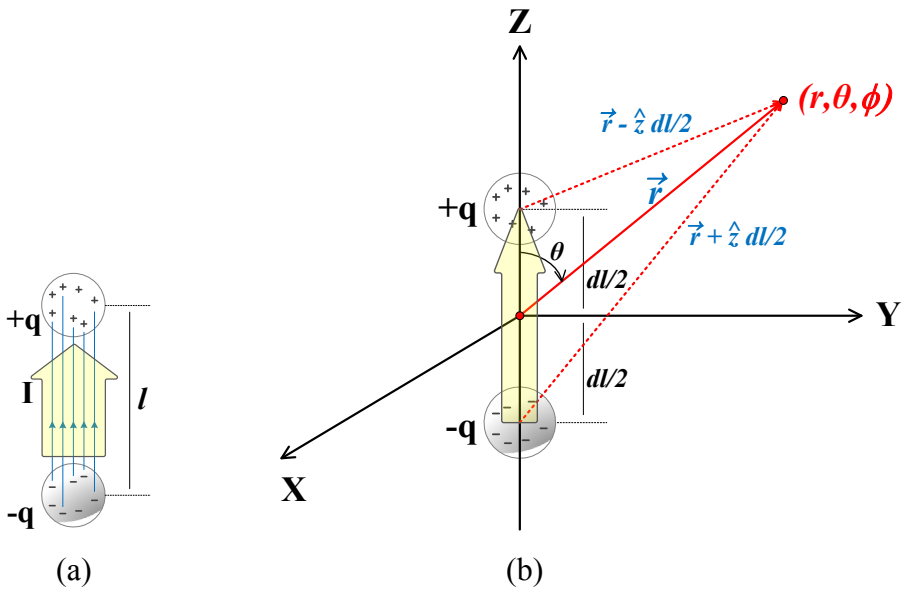


Figure 4.11 Static charge dipole problem – (a) Illustration of point source and (b) Illustration of parameters and coordinates for the static dipole.

Therefore, (4.96) suggest that that the minimum non-radiating energy is actually due to the static electric field exist between the charges for spherical TM_{01} mode case. In terms of static fields, this result implies that the amount of the static electric energy needs to be charged when the ideal dipole starts radiating. The corresponding time required to charge the static energy might be the delay that we observed in the time-domain approach.

In fact, because we know all magnetic energy is used for radiation in a single mode excitation of spherical TM modes, we could take a low-frequency approximation ($kr \to 0$) on the electric fields and evaluate the non-radiating energy from the low-frequency approximated fields. For the spherical TM_{01} mode, the $1/r^3$ terms could be conveniently chosen in order to evaluate the non-radiating energy.

Applying what we learned about the radiation process in Chapter 2, this result makes sense because the static fields are still bound to the charges. Thus, the amount of the bound static fields corresponds to the minimum non-radiating energy. As soon as we identify the origin of the minimum non-radiating energy, evaluation of minimum radiation Q becomes almost trivial for single mode excitation.

4.2.2 Spherical $TM_{-1,1}$ and $TM_{1,1}$ modes

The spherical $TM_{-1,1}$ and $TM_{1,1}$ modes are the lowest single mode to support circularly-polarized wave radiation while the spherical TM_{01} and TE_{01} modes are the lowest single mode to support a linearly-polarized wave radiation. In this subsection, we investigate the minimum radiation Q for the spherical $TM_{1,1}$ mode providing a left hand circular polarization (LHCP) with $e^{j\omega t}$ time dependence assumption.

Using the spherical mode expansion formula in Chapter 3, the time-harmonic fields for the spherical $TM_{1,1}$ mode with a source normalization factor ($Idl = 4\pi / k$) and an excess phase ($\zeta = \tan^{-1}\left[1/(kr)\right]$) are found as

$$\begin{aligned}\vec{E}(r,t) = {} & \hat{r}2k\eta_0\left(1-\cos^2\theta\right)^{1/2}\frac{\sqrt{k^2r^2+1}}{k^3r^3}\sin\left(\omega t - kr + \phi - \zeta\right) \\ & -\hat{\theta}\frac{k\eta_0}{\sqrt{k^2r^2+1}}\frac{\sin\theta\cos\theta}{\left(1-\cos^2\theta\right)^{1/2}}\left[\cos(\omega t - kr + \phi - \zeta) + \frac{1}{(kr)^3}\sin(\omega t - kr + \phi - \zeta)\right] \\ & -\hat{\phi}\frac{k\eta_0}{\sqrt{k^2r^2+1}}\frac{\left(1-\cos^2\theta\right)^{1/2}}{\sin\theta}\left[\cos(\omega t - kr + \phi - \zeta + \frac{\pi}{2}) + \frac{1}{k^3r^3}\sin(\omega t - kr + \phi - \zeta + \frac{\pi}{2})\right]\end{aligned} \tag{4.97}$$

$$\begin{aligned}\vec{H}(r,t) = {} & \hat{\theta}k\frac{\sqrt{k^2r^2+1}}{k^2r^2}\frac{\left(1-\cos^2\theta\right)^{1/2}}{\sin\theta}\cos\left(\omega t - kr + \phi - \zeta + \frac{\pi}{2}\right) \\ & -\hat{\phi}k\frac{\sqrt{k^2r^2+1}}{k^2r^2}\frac{\sin\theta\cos\theta}{\left(1-\cos^2\theta\right)^{1/2}}\cos\left(\omega t - kr + \phi - \zeta\right)\end{aligned} \tag{4.98}$$

The Poynting vector of the spherical $TM_{1,1}$ mode has the following form:

$$\vec{S} = \left[\hat{r}E_\theta H_\phi + \hat{\theta}E_r H_\phi\right] + \left[-\hat{r}E_\phi H_\theta + \hat{\phi}E_r H_\theta\right] \tag{4.99}$$

which may be seen as a combination of two orthogonal waves, i.e. one is E_θ (and H_ϕ) and the other is E_ϕ (and H_θ). Both E_ϕ and H_θ hav2 $\pi/2$ phase offset compared to the other set. The fields contributing to the radial power are orthogonal to each other, but the energy storage processes are correlated via the field component E_r. Thus, the spherical $TM_{1,1}$ mode is a different from the linear combination of spherical TM and TE modes to obtain a circular polarization. In other words, the spherical $TM_{1,1}$ is a separate, unique mode.

The radial Poynting vector is found as

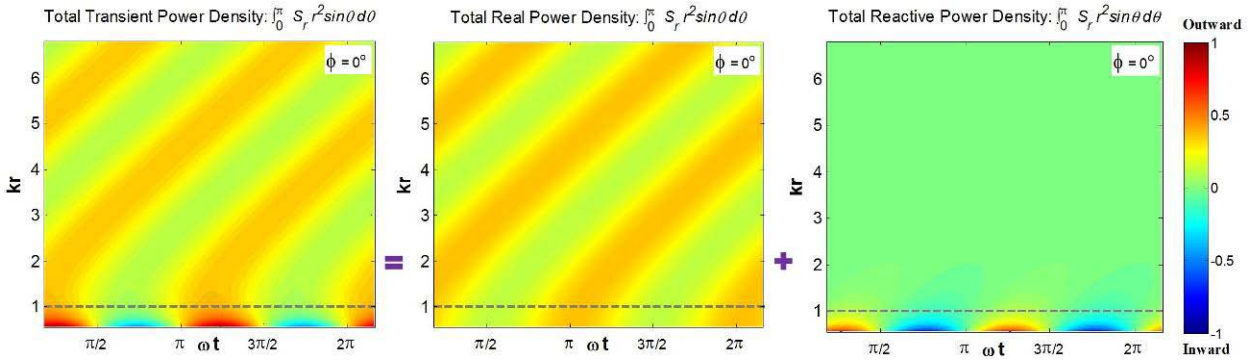
$$\begin{aligned} S_r &= E_\theta H_\phi - E_\phi H_\theta \\ &= \frac{1}{2}\frac{\eta_0}{r^2}\cos^2\theta\left[1+\cos 2\left(\omega t - kr + \phi - \zeta\right) + \frac{1}{(kr)^3}\sin 2(\omega t - kr + \phi - \zeta)\right] \\ &\quad + \frac{1}{2}\frac{\eta_0}{r^2}\left[1+\cos 2\left(\omega t - kr + \phi - \zeta + \frac{\pi}{2}\right) + \frac{1}{k^3 r^3}\sin 2(\omega t - kr + \phi - \zeta + \frac{\pi}{2})\right] \end{aligned} \tag{4.100}$$

In order to understand the spherical $TM_{1,1}$ mode, consider a power density function $p(\phi)$ found by integrating radial Poynting vector over θ only:

$$\begin{aligned} p(\phi) &= \int_{\theta=0}^{\pi}\left(E_\theta H_\phi - E_\phi H_\theta\right)\cdot \hat{r} r^2 \sin\theta d\theta \\ &= \frac{1}{3}\eta_0\left[1+\cos 2\left(\omega t - kr + \phi - \zeta\right) + \frac{1}{(kr)^3}\sin 2(\omega t - kr + \phi - \zeta)\right] \\ &\quad + \frac{3}{3}\eta_0\left[1+\cos 2\left(\omega t - kr + \phi - \zeta + \frac{\pi}{2}\right) + \frac{1}{(kr)^3}\sin 2(\omega t - kr + \phi - \zeta + \frac{\pi}{2})\right] \end{aligned} \tag{4.101}$$

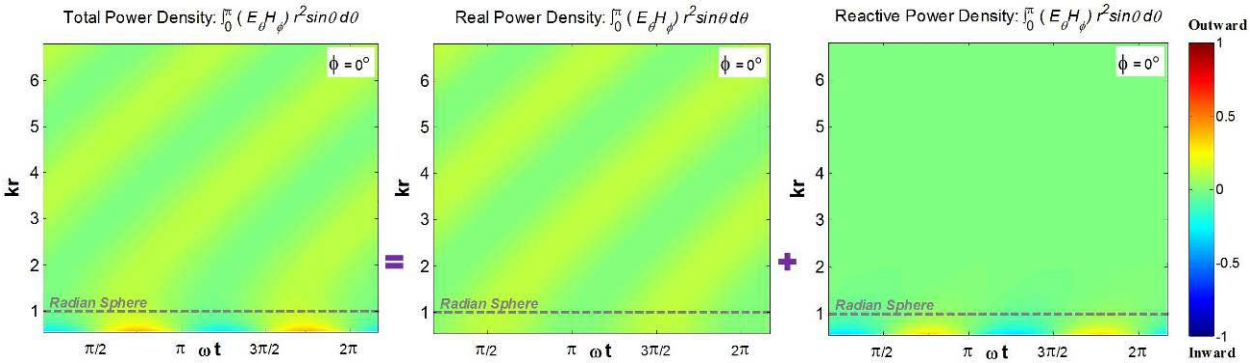
The third and forth terms with $\pi/2$ phase offset in (4.101) dominate the radiation property of the spherical TM_{11} mode, i.e. E_ϕ and H_θ govern the radiation property of the spherical TM_{11} mode due to the higher power radiation. The radial Poynting vector of the $E_\theta H_\phi$ product has θ dependence, i.e. $\cos^2\theta$, which gives a peak at $\theta = 0$, π and a null at $\theta = \pi/2$. However, The radial Poynting vector of the $E_\phi H_\theta$ product is not a function of θ. Thus, the radiated power from the third term in (4.101) should be isotropic.

Time-space plots of the power density function $p(\phi)$ with fixed ϕ are shown in Figure 4.12 through Figure 4.14. Although both $E_\theta H_\phi$ and $E_\phi H_\theta$ products radiate with $\pi/2$ phase offset, we can clearly observed that $E_\phi H_\theta$ product dominate the radiated power. An alternative representation may be a time-varying plot of (4.101) with a varying ϕ. The time-varying plots in Figure 4.15 clearly show that the spherical TM_{11} mode supports left-handed circular polarization.



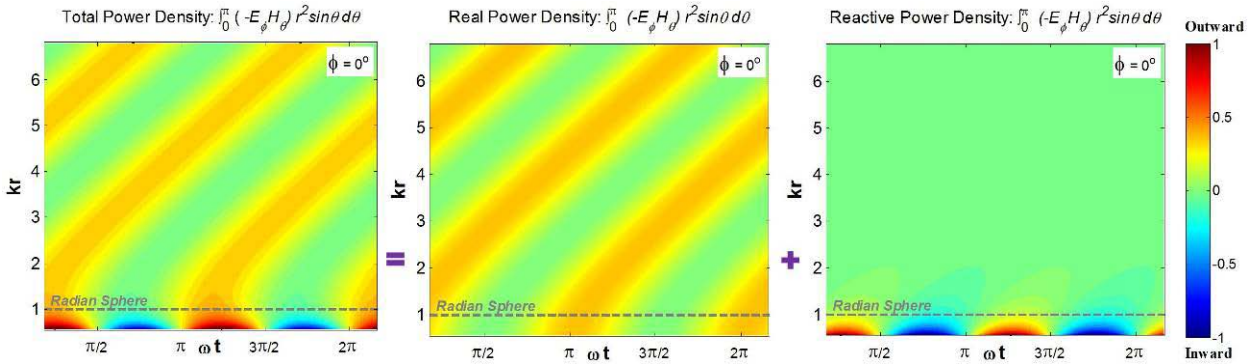
(a) (b) (c)

Figure 4.12 Time space plot of power density $p(\phi = 0^o)$ for the spherical TM_{11} mode. Amplitude was normalized by the peak of the total transient power density.



(a) (b) (c)

Figure 4.13 Time space plot of the first and the second terms of power density $p(\phi = 0^o)$ for the spherical TM_{11} mode. Amplitude was normalized by the peak of the total transient power density.



(a) (b) (c)

Figure 4.14 Time space plot of the third and fourth terms of power density $p(\phi = 0^o)$ for the spherical TM_{11} mode. Amplitude was normalized by the peak of the total transient power density.

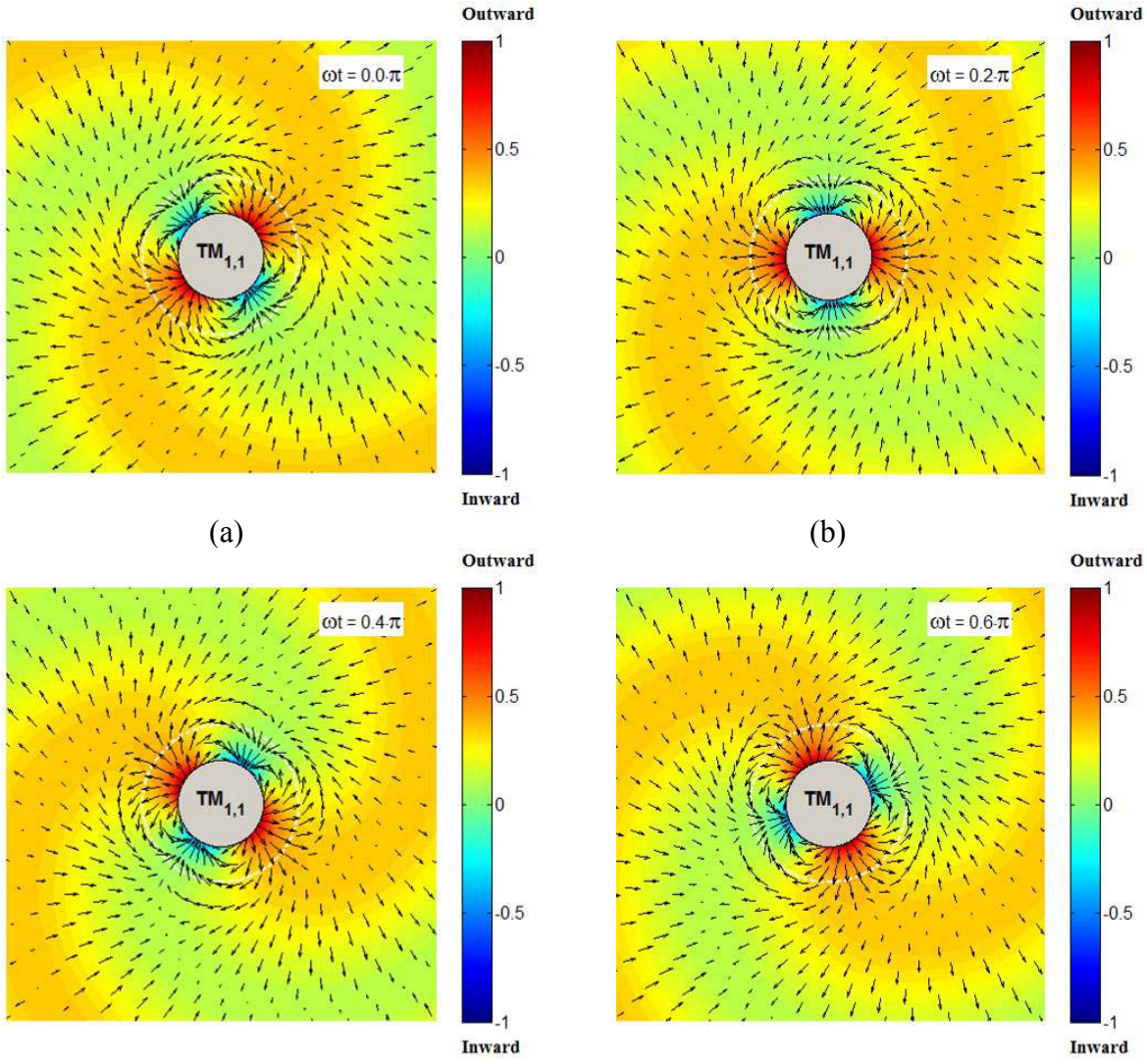
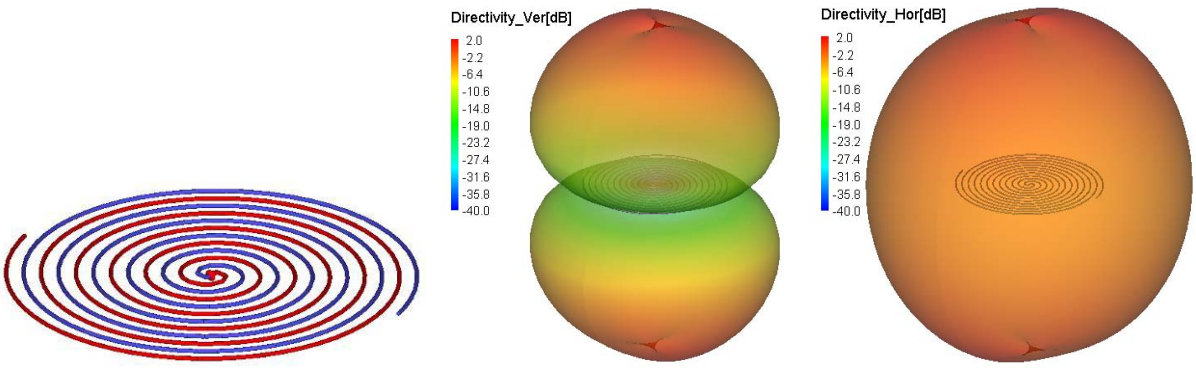


Figure 4.15 Time-varying plot of power density function $p(\phi)$ for the spherical TM_{11} mode (LHCP) in *x*-*y* plane – (a) $\omega t = 0$, (b) $\omega t = 0.2\pi$, (c) $\omega t = 0.4\pi$, and (d) $\omega t = 0.6\pi$. Note that the arrows represent a gradient. The circle represents the antenna sphere.

The closest structure supporting the spherical TM_{11} mode among well-known antennas is the Archimedean spiral. A LHCP Archimedean spiral was modeled using the FEKO code [*FEKO*, 2007]. The geometry and simulated partial directivity patterns (D_θ and D_ϕ) for each polarization are shown in Figure 4.16. Similar to the spherical TM_{11} mode, D_ϕ pattern is almost isotropic while D_θ pattern has a peak at $\theta = 0$ and a null at $\theta = \pi / 2$.



(a) (b) (c)

Figure 4.16 Simulated partial radiation patterns of Archimedean spiral at the lowest operating frequency- (a) Geometry supporting LHCP, (b) D_θ and (b) D_ϕ. A commercial method of moments code [*FEKO*, 2007] was used for the simulation.

The time average power is given as

$$\begin{aligned}\langle P\rangle &= \oint_S \vec{E}\times\vec{H}\cdot d\vec{s} \\ &= \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi} (\hat{r}E_\theta H_\phi - \hat{r}E_\phi H_\theta)\cdot \hat{r}r^2\sin\theta d\theta d\phi \\ &= \eta_0 \frac{2\pi}{3}(1+3)\end{aligned} \tag{4.102}$$

which is double the power for the spherical TM_{01} or TE_{01} cases in previous section. In addition, the power density is spit into $E_\theta H_\phi$ and $E_\phi H_\theta$ with 1:3 ratios.

In order to find the phase velocity (v_p) of the outgoing radiated power, we examine a constant phase point:

$$\omega t - kr + \phi - \zeta = \text{constant} \tag{4.103}$$

After taking a total derivative of the constant phase, the resultant phase velocity is found as

$$v_p = \left(1+\frac{1}{k^2r^2}\right)c_0 \tag{4.104}$$

which is the same as the phase velocity of the spherical TM_{01} mode of (4.47). The other phase form shown in the fields includes the constant $\pi/2$ phase offset, but the constant phase offset does not change the phase velocity. The corresponding energy velocity is obtained from

$$v_e = \frac{{c_0}^2}{v_p} = \left[1-\frac{1}{1+k^2r^2}\right]c_0 \tag{4.105}$$

Thus, the time-average radiating energy density per radial distance is found as

$$\langle U_{rad} \rangle = \frac{\langle P_{rad} \rangle}{v_e} = \mu_0 \frac{8\pi}{3}\left(1+\frac{1}{k^2 r^2}\right) \tag{4.106}$$

The time-average electric and magnetic energy densities per radial distance are given as, respectively

$$\langle U_e \rangle = \mu_0 \frac{4\pi}{3}\left(1+\frac{1}{k^2 r^2}+\frac{3}{k^4 r^4}\right) \tag{4.107}$$

$$\langle U_m \rangle = \mu_0 \frac{4\pi}{3}\left(1+\frac{1}{k^2 r^2}\right) \tag{4.108}$$

The time-average electric non-radiating energy density per radial distance is found as

$$\langle U_{e,non-rad} \rangle = \langle U_e \rangle - \frac{\langle U_{rad} \rangle}{2} = \mu_0 \frac{4\pi}{k^4 r^4} \tag{4.109}$$

which is double that in (4.59) for the spherical TM_{01} mode. The time-average non-radiating magnetic energy density per radial distance becomes zero. Since the radiated power of the spherical $TM_{1,1}$ mode are also doubled, the resulting radiation Q of the spherical $TM_{1,1}$ mode is the same as the radiation Q of the spherical TM_{01} mode (see Figure 4.17). It can be shown that all spherical modes with different degrees (m, ϕ behavior), yet with the same order (n, θ and r behavior), have the same minimum radiation Q. In fact, the $TM_{-1,1}$ ($TE_{-1,1}$) or $TM_{1,1}$ ($TE_{1,1}$) modes provide bidirectional right-hand circular polarization (RHCP) and left-hand circular polarization (LHCP),respectively, but with the same Q as linear antennas. In summary, we have

$$Q_{VTAG}^{TM_{1,1}} = Q_{VTAG}^{TM_{-1,1}} = Q_{VTAG}^{TE_{1,1}} = Q_{VTAG}^{TE_{-1,1}} = e_r \frac{1}{k^3 a^3} \tag{4.110}$$

Alternatively, it can be shown that we can obtain the same result as (4.110) with a low-frequency approximation.

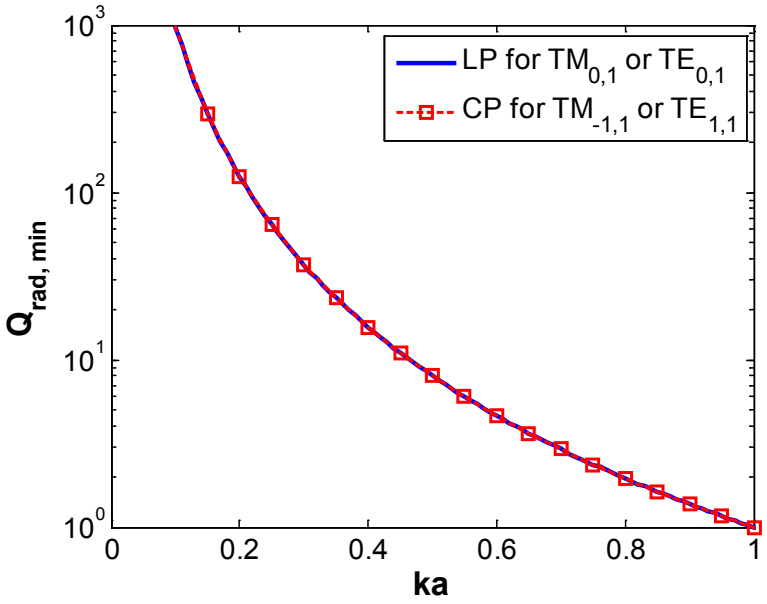


Figure 4.17 Minimum radiation Q comparison of the fundamental spherical $TM_{0,1}$ and $TE_{0,1}$ modes (linear polarization) and the spherical $TM_{-1,1}$ and $TM_{1,1}$ modes (circular polarization) [*Yang and Davis*, 2009]. Copyright © 2009 IEEE.

4.2.3 Arbitrarily Polarized Antenna

Each fundamental spherical mode has a linear polarization. However, the combination of spherical TM_{01} and TE_{01} modes can address the minimum radiation Q of an arbitrary-polarized antenna. The fields of each fundamental spherical mode for the combined mode can be written as

$$\begin{aligned}\vec{E}_{TM} &= \hat{r}2\eta_0 \frac{\cos\theta}{r}\frac{\sqrt{1+k^2r^2}}{k^2r^2}\sin(\omega t - kr - \zeta) \\ &+ \hat{\theta}\eta_0 \frac{\sin\theta}{r}\frac{kr}{\sqrt{1+k^2r^2}}\left\{\cos(\omega t - kr - \zeta) + \frac{1}{k^3r^3}\sin(\omega t - kr - \zeta)\right\}\end{aligned} \tag{4.111}$$

$$\vec{H}_{TM} = \hat{\phi}\frac{\sin\theta}{r}\frac{\sqrt{1+k^2r^2}}{kr}\cos(\omega t - kr - \zeta) \tag{4.112}$$

$$\vec{E}_{TE} = -A\hat{\phi}\frac{1}{\eta_0}\frac{\sin\theta}{r}\frac{\sqrt{1+k^2r^2}}{kr}\cos(\omega t - kr - \zeta + \gamma) \tag{4.113}$$

$$\begin{aligned}\vec{H}_{TE} &= A\hat{r}2\frac{\cos\theta}{r}\frac{\sqrt{1+k^2r^2}}{k^2r^2}\sin(\omega t - kr - \zeta + \gamma) \\ &A\hat{\theta}\frac{\sin\theta}{r}\frac{kr}{\sqrt{1+k^2r^2}}\left\{\cos(\omega t - kr - \zeta + \gamma) + \frac{1}{k^3r^3}\sin(\omega t - kr - \zeta + \gamma)\right\}\end{aligned} \tag{4.114}$$

where A is an amplitude scaling factor and γ is a phase difference between the modes. The total power of the combined spherical mode representing an arbitrarily-polarized antenna can be written as

$$P_{Arbitrary-Pol.}^{TM_{01}+TM_{01}}(r,t) = \langle P \rangle \left[1 + \cos 2(\omega t - kr - \zeta) + \frac{1}{(kr)^3} \sin 2(\omega t - kr - \zeta) \right] + A^2 \langle P \rangle \left[1 + \cos 2(\omega t - kr - \zeta + \gamma) + \frac{1}{(kr)^3} \sin 2(\omega t - kr - \zeta + \gamma) \right] \quad (4.115)$$

where A is an amplitude scaling factor, $\langle P \rangle$ is time average power for each spherical mode, and γ is a phase difference between the modes. The time-space plot of (4.115) for an elliptically-polarized case (A = 0.5 and $\gamma = \pi/2$) is shown in Figure 4.18. We can clearly observe in time that as one mode is radiating energy, the other mode will be storing energy.

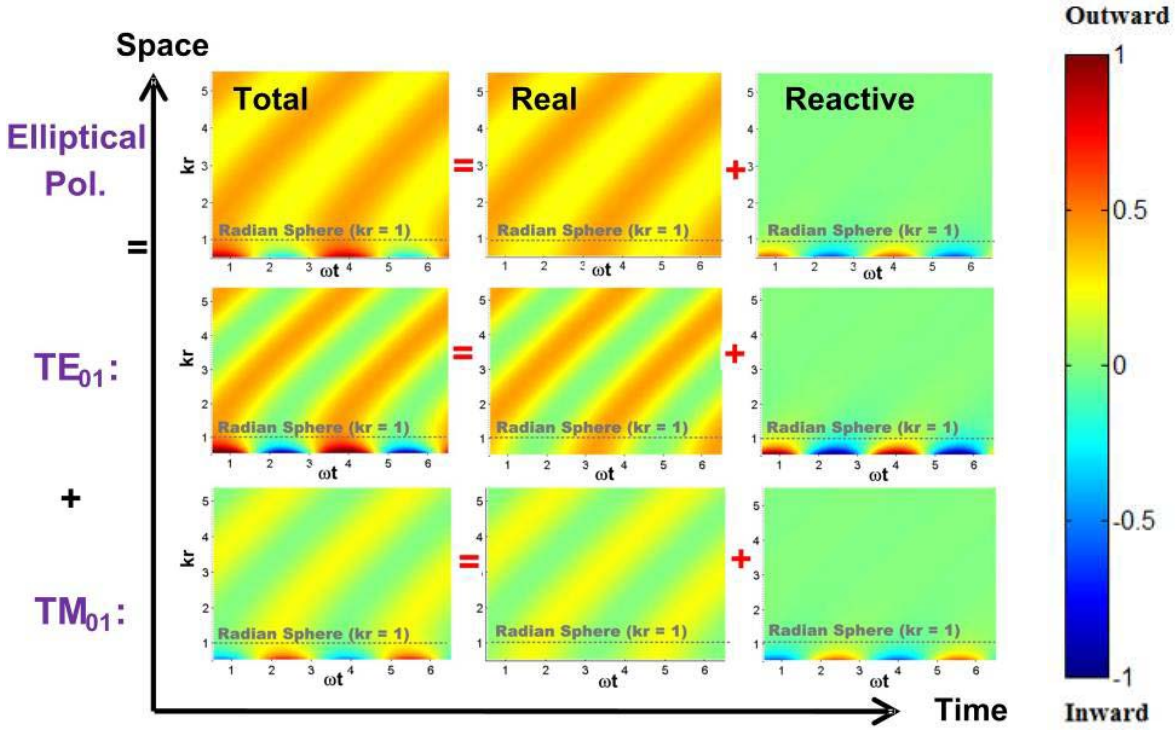


Figure 4.18 Time-space plot of total power for elliptically polarized antenna with both spherical TM_{01} and TE_{01} mode excitations in phase quadrature case (A = 0.5 and $\gamma = \pi/2$) [*Yang and Davis*, 2009].

Consider two independent antennas with independent sources, i.e. independent feeds. The time-average combined power can be written as

$$\left\langle P_{Arbitrary-Pol.}^{TM_{01}+TM_{01}} \right\rangle = \langle P \rangle \left(1 + A^2\right) \tag{4.116}$$

Then, the reactive power of the combined modes is given as

$$P_{reactive} = \frac{\langle P \rangle}{k^3 r^3} \left[\sin 2(\alpha - \zeta) + A^2 \sin 2(\alpha - \zeta + \gamma) \right] \tag{4.117}$$

As discussed in the previous section, the relationship between the reactive power and the change of total non-radiating energy about an antenna sphere can be obtained using (4.70), in the following total non-radiating energy expression:

$$W_{non-rad} = \langle P \rangle \left\{ C - \frac{1}{2\omega k^3 a^3} \left[\cos 2(\omega t - ka - \zeta_a) + A^2 \cos 2(\omega t - ka - \zeta_a + \gamma) \right] \right\} \tag{4.118}$$

where C is a constant of integration. In order not to have negative energy for (4.118), the range of C is limited to

$$C \geq \frac{1 + A^2}{2\omega k^3 a^3} \tag{4.119}$$

The corresponding time-average minimum non-radiating energy is found as

$$\left\langle W_{non-rad,\min} \right\rangle = \langle P \rangle \frac{1 + A^2}{2\omega k^3 a^3} \tag{4.120}$$

Using (4.2), the minimum radiation Q of arbitrarily-polarized (AP) antenna with an independent feed (IF) is found as

$$Q_{rad,AP\ (IF)} = e_r \frac{1}{k^3 a^3} \tag{4.121}$$

which is the same result as the minimum radiation Q of the fundamental spherical modes found in (4.63). The reason for the same radiation Q is that both the total non-radiating energy and power increase or decrease at the same time for the combined mode with the independent feed assumption.

On the other hand, if a single source (corporate feed) is assumed, the combined magnetic and electric energies can be balance each other. Thus, the resulting minimum radiation Q can be reduced or the same minimum radiation Q as for a linearly polarized antenna. Thus, in the elliptically-polarized case with the corporate feed assumption gives a minimum radiation-Q between the limits of linearly and circularly polarized antennas. For example, the minimum radiation-Q for the elliptically polarized case with $\gamma = \pi / 2$ for corporate feed (CF) is found as

$$Q_{rad,\min\ (CF)}^{Elliptically-Pol.} = \frac{AR^2}{1+AR^2}\frac{1}{k^3a^3} \tag{4.122}$$

where AR is axial ratio. The plot of (4.122) is shown in Figure 4.19 for sample cases. The minimum radiation Qs of antennas with various polarizations and feed configurations are summarized in Table 4-1.

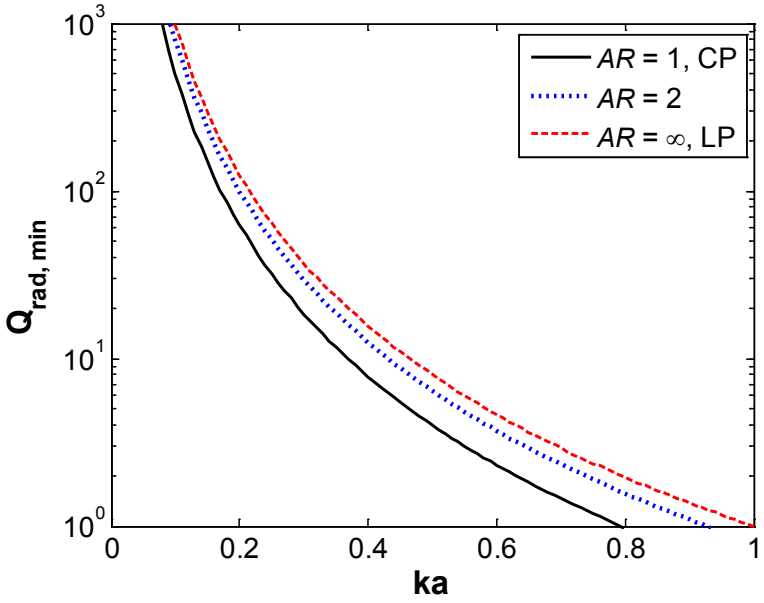
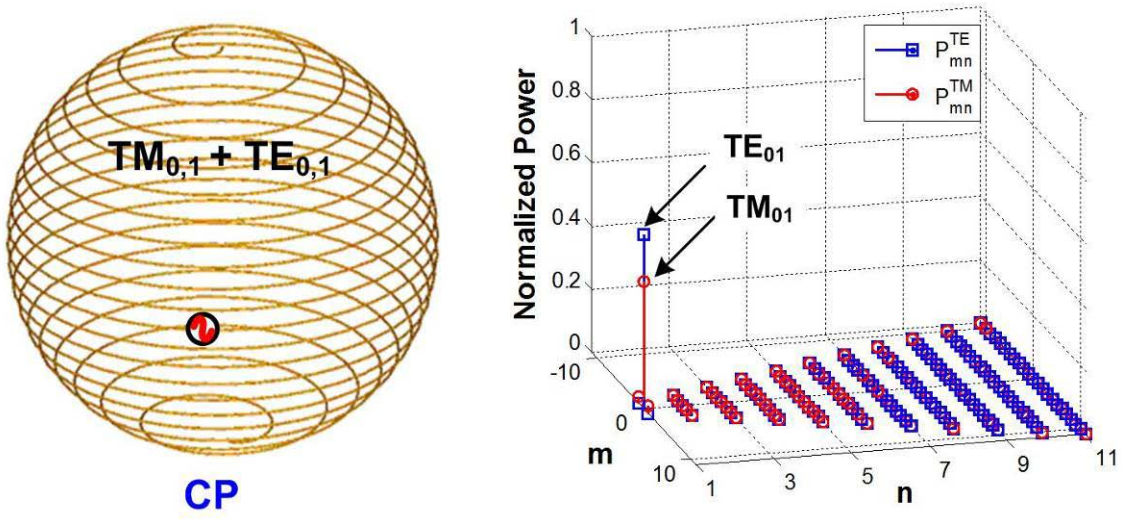


Figure 4.19 Minimum radiation Q for comparison between various polarizations, assuming corporate feed to combine the spherical TM_{01} and TE_{01} modes [*Yang and Davis*, 2009].

Table 4-1 Summary of minimum radiation Q for antennas with various polarizations and feed types

	Minimum Radiation Q	
	Corporate Feed	**Independent Feed**
Linear Polarization (A=1, $\gamma = 0$)	$e_r \frac{1}{k^3a^3}$	$e_r \frac{1}{k^3a^3}$
Circular Polarization (A=1, $\gamma = \pi/2$)	$e_r \frac{1}{2}\frac{1}{k^3a^3}$	$e_r \frac{1}{k^3a^3}$
Elliptical Polarization ($\gamma = \pi/2$)	$e_r \frac{AR^2}{1+AR^2}\frac{1}{k^3a^3}$	$e_r \frac{1}{k^3a^3}$

Because the case combining the fundamental modes with an isotropic circular polarization and an integrated corporate feed provides the lowest radiation Q, it is of interest to identify what antenna structure can potentially achieve the lowest radiation Q. As an example antenna structure, consider the spherical helix [*Cardoso and Safaai-Jazi*, 1993; *Safaai-Jazi and Cardoso*, 1996] in Figure 4.20a. The modal power analysis through the spherical wave expansion (SWE) in Chapter 3 clearly shows that both fundamental spherical modes are excited in the spherical helix, but with a slightly different radiation power ratio between the modes (see Figure 4.20b). Thus, the particular spherical helix design in Figure 4.20a will excite elliptically-polarized waves and the resulting radiation Q would is given by (4.122). The unbalanced power might be caused from the offset feed location.



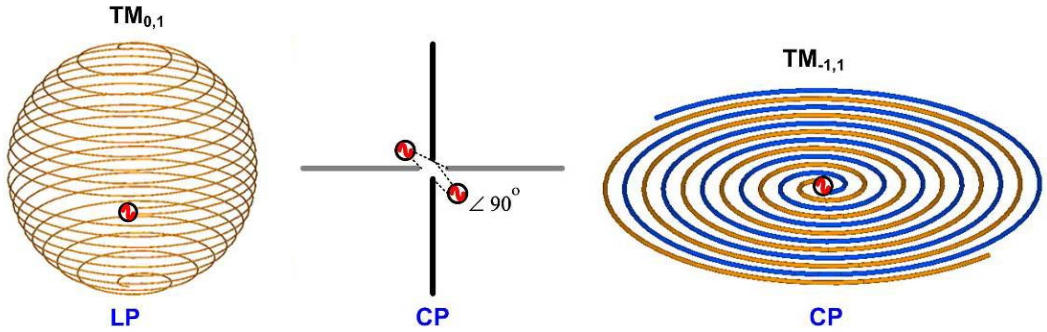
(a) (b)

Figure 4.20 Excited spherical modes of spherical helix [*Yang and Davis*, 2009] – (a) Geometry of spherical helix and (b) Modal power analysis. Copyright © 2009 IEEE.

On the other hand, an electrically-small spherical helix dipole (Figure 4.21a) may look similar to the spherical helix in Figure 4.20a. However, the minimum radiation Q limit for the spherical helix dipole would be the same as the radiation Q of an electrically-small hemispherical helix over an infinite ground due to the cancellation of the image current. Both dual-linear antennas with phase quadrature and the spiral in Figure 4.21b and Figure 4.21c will excite

circularly-polarized waves, but the resulting minimum radiation-Q will be the same as linearly-polarized cases.

Thus, one cannot use the radiation Q limit blindly because the theoretical limit varies with source configurations and antenna geometry.



(a) (b) (c)

Figure 4.21 Excited spherical modes of various antennas [*Yang and Davis*, 2009] – (a) Linearly-polarized electrically-small spherical dipole, (b) Circularly-polarized cross dipole with independent source, and (c) Archimedean spiral antenna with right-hand circular polarization.

4.2.4 Radiation-Q Evaluation and Measurement

Once we find the fundamental-limit curve of antenna radiation Q, the next question is how we can estimate and/or measure the radiation-Q of the antenna under evaluation to compare with the theoretical limit curve. It is well known that the 3-dB impedance bandwidth (BW_{3dB}) of a simple *RLC* resonant circuit (parallel or series) is the inverse of the Q of the resonant circuit given by

$$Q_{rad} \simeq \frac{1}{BW_{3dB}} \tag{4.123}$$

The 3-dB impedance bandwidth can also be converted from the impedance bandwidth with voltage standing wave ratio (*VSWR*) criteria by using [*Volakis*, 2007]

$$BW_{3-dB} = BW_{VSWR=S}\frac{2\sqrt{S}}{S-1} \tag{4.124}$$

where *VSWR* is represented by S. In terms of the *VSWR* bandwidth (BW_{VSWR}), (4.123) can be rewritten as

$$Q_{rad} \simeq \frac{VSWR-1}{BW_{VSWR}\sqrt{VSWR}} \tag{4.125}$$

Therefore, the radiation Q of a designed antenna can be estimated from measured VSWR bandwidth. If a given antenna can be modeled as the simple single-resonance circuit (i.e. high-Q), the estimated Q is close to the actual radiation-Q of the antenna because (4.123) and (4.124) were obtained with an electrically-small antenna assumption or directly from the simple *RLC* resonant circuit. Nevertheless, this method has been frequently used because of its simplicity.

Geyi [*Geyi*, 2003] proposed a Q estimation method based on his Foster's reactance theorem development for antennas, which is given by

$$Q_{rad} = \frac{\omega \frac{dX_A}{d\omega} \pm X_A}{2R_A^{rad}} \tag{4.126}$$

where X_A and R_A^{rad} are reactance of antenna input impedance and radiation resistance, respectively. Either + or – sign is chosen to give the higher Q. The essence of the conventional Foster's reactance is "The slope of the reactance or susceptance for a loss-free one-port network is always positive" [*R F Harrington*, 2001]. Because of radiation, an antenna is inherently a lossy one-port problem. As Best [*Best*, 2004b] pointed out, Geyi's Foster reactance development and Q formula is invalid at the anti-resonance of the antenna input impedance.

Yaghjian and Best [*Yaghjian and Best*, 2005] proposed an approximate antenna radiation-Q formula for tuned antennas, given by

$$Q_{rad} \approx \frac{\omega \left| \frac{dZ_A}{d\omega} \right|}{2R_A^{rad}} \Bigg|_{\omega \to \omega_0} \tag{4.127}$$

where Z_A is the antenna input impedance and ω_0 is the resonant frequency. Unlike Geyi's formula (4.126), (4.127) can be used even at the anti-resonant frequency. Except for the consideration of the source condition at the antenna terminal, the basic derivation processes of both (4.126) and (4.127) are not much different from the process used to obtain Chu's equivalent circuit [*Chu*, 1948]. In addition, both (4.126) and (4.127) have same problems as with (4.123) for electrically-large antennas. Thus, (4.126) and (4.127) are good approximations for electrically-small antennas. In this sense, the antenna radiation Q estimate using (4.126) or (4.127) is very similar to the estimate of (4.125) when an electrically-small antenna is considered.

Some approaches to estimate radiation-Q for an arbitrarily-shaped antenna have been proposed. Based on the time-domain Poynting theorem, Geyi [*Geyi*, 2003] estimated the radiation Q for small dipole, small loop, and small inverted-L antennas,. His approach requires the current distribution on the antenna to estimate the radiation Q. Liu [*Liu et al.*, 1999] demonstrated the approach to measure the radiation-Q of antenna in the time-domain. They measured peak energy returning to the source after turning off the source and assumed the returned energy as non-propagating energy.

All the proposed estimation methods for the radiation Q are either approximated with limited assumptions or are of questionable accuracy. Even though the radiation Q can be estimated accurately with (4.126) or (4.127), the estimated radiation Q is for a single frequency. Thus, both (4.126) or (4.127) are not very useful for frequency-independent or ultra-wideband antennas. In this sense, (4.125) is more practical.

The minimum radiation-Q curves, with the assumption of 100% radiation efficiency, are shown in Figure 4.22 along with the typical radiation Q of various antenna types. The antenna radiation Q was evaluated using (4.125). Clearly, the limit curve shows that radiation Q should increase if antenna size (ka) decreases. In order to maintain impedance bandwidth as the size is reduced, the radiation efficiency must be decreased accordingly. This is the trade-off that the classic fundamental-limit theory suggests, i.e. only two of the three parameters of antenna size, radiation efficiency (or gain), and bandwidth can be optimized close to the theoretical limit curve.

Notice that the radiation Q of typical antennas in Figure 4.22 is not even close to the theoretical limit curve. Especially, the legacy performance line connecting a dipole to a spiral antenna appears as a simple shift of the limit curve, i.e. resulting in a higher radiation Q or narrower impedance bandwidth. The shift occurs because those antennas do not use the volume of the antenna sphere efficiently or do not integrate wideband impedance-matching topology into the antenna, compared to the folded hemisphere and Goubau antennas that fill the volume.

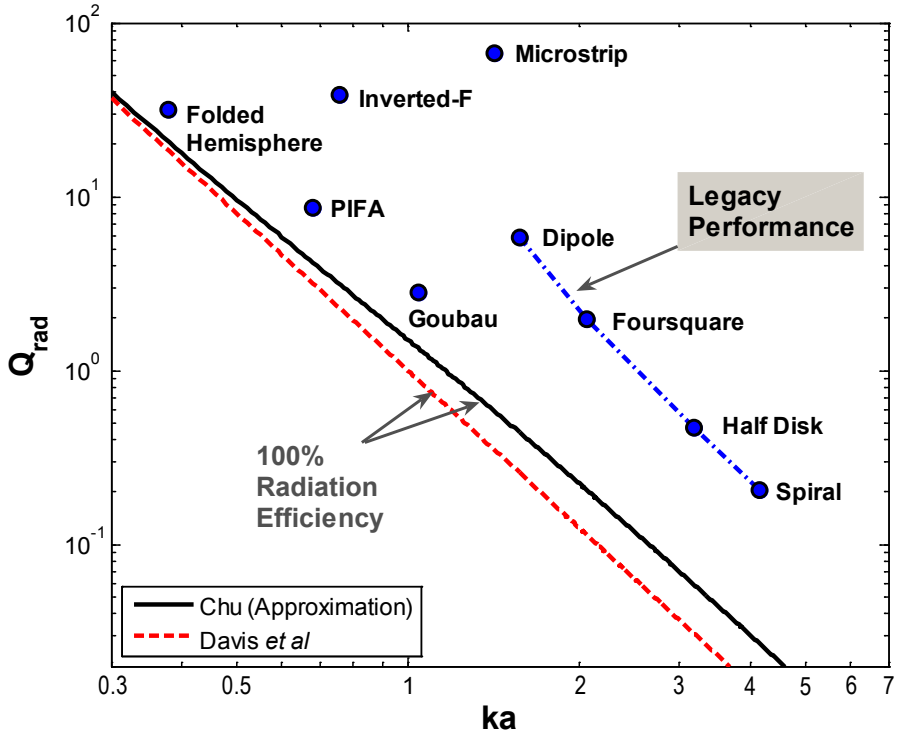


Figure 4.22 Comparison of theoretical radiation-Q limit curves with 100 % radiation efficiency and estimated radiation Q of various antennas in a log-log scale [*Yang and Davis*, 2006].

Thus, the size and bandwidth (or efficiency) specification for typical real-world antenna design can be estimated from Figure 4.22, i.e. the radiation Q corresponding to the targeted antenna size and bandwidth (or efficiency) is typically located between the limit curve and legacy performance curve.

4.3 Consideration for Frequency-Independent and Ultra-Wideband Antennas

Classical fundamental limit theories propose various minimum radiation Q expressions. For either field or circuit approaches, the differences in the minimum Q formulas are due to the way the amount of non-radiating stored energy is estimated. As we demonstrated in the previous sections, the resulting minimum Q formula should be the same for both field and circuit approaches if one is consistent in estimating the amount of non-radiating stored energy in both approaches. Thus, the choice of a field or a circuit representation of the fundamental-limit problem depends on one's preference. In this section, we re-examine Chu's equivalent circuit

model of spherical modes in order to obtain some fundamental-limit aspects of ultra-wideband antennas.

One frequently-asked question about Chu's equivalent circuit in Figure 4.2b is where is the antenna located? Actually, one implicit assumption of Chu's circuit is that the arbitrarily-shaped antenna in the antenna sphere has an all-pass-filter characteristic. The antenna is not shown in Chu's equivalent circuit. The phase delay of energy propagation to the far-field region is also not included because this delay is irrelevant to the evaluation of radiation Q and can be modeled as a transmission line on the output of the model. In addition, Chu's model basically represents a one-port problem.

If we consider how much power is delivered to space (radiation), Chu's ladder circuit becomes to a two-port problem as shown in Figure 4.23. The all-pass-filter antenna and the propagation phase delay is included to provide a better radiation viewpoint including the antenna, even though both extended and original circuits are equivalent in terms of the impedance viewpoint. As explained earlier, Chu's circuit model has characteristics of a high pass filter, as modeled by the ladder network.

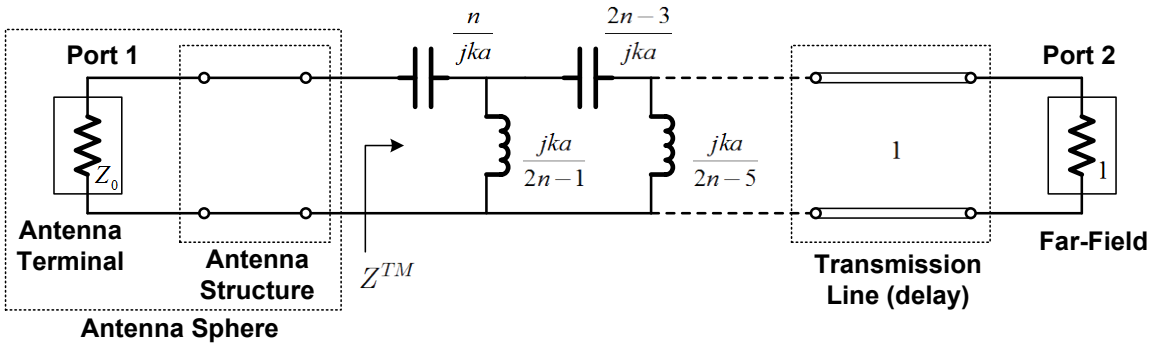


Figure 4.23 The mode circuit model – (a) Definition of antenna sphere, (b) Chu's original circuit model and (c) Extended circuit model with two-port scattering parameter representation. The ladder network continues until just before a change in component sign. A dual ladder network applies to spherical TE modes.

In the extended circuit model (Figure 4.23), port 1 is the antenna terminal with a system reference impedance z_0. Port 2 represents the far-field region where radiated power is delivered. Therefore, $|s_{21}|^2$ of the extended model represents the transfer of power from the antenna

terminal to the radiated far field of the mode being considered. For example, s_{21} for the spherical TM_{01} mode of the two-port equivalent of Chu's model (neglecting the transmission line) can be written as

$$s_{21} = \frac{2s^2}{2s^2 + 2s + 1} = 1 - 0.5\left(\frac{1}{s - (-0.5 + j0.5)} + \frac{1}{s - (-0.5 - j0.5)}\right) \quad (4.128)$$

where $s = jka$. The reference impedance of Port 1 is assumed to be same with that of Port 2 ($z_0 = 1$). This form shows that s_{21} for the fundamental spherical mode (TM_{01}) consists of an entire function and two complex poles (see Figure 4.27a). This represents what we refer to as an *ideal antenna* which meets the theoretical performance and size limit. The ideal antenna based on the spherical TM_{01} mode has the well-known donut-shaped radiation pattern with an omni-directional pattern in the azimuth plane. The entire frequency-domain and time-domain responses of the ideal antenna can be completely described with only two poles and an entire function. Thus, the characteristics of the ideal antenna can be easily included in simulations to evaluate communication system performance without introducing a significant computational load.

As we observed in Chapter 3, many real antennas have more complicated pole-residue structures than the ideal antenna. Thus, the antenna performance and the location of the poles changes because of various factors. These factors are depicted in Figure 4.24. First, the source conditions, such as voltage, current, and matched source, can move the location poles or change the number of them. In addition, radiation performance also varies, depending on the source condition. For example, assume that only the spherical TM_{01} mode is excited for an electrically small antenna and consider source loading effects for voltage, matched, and current source, as shown in Figure 4.25.

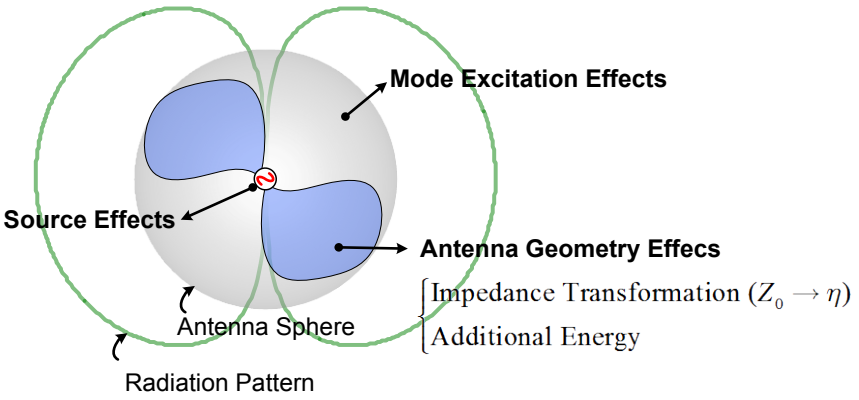


Figure 4.24 Illustration of realistic antenna in term of design viewpoint.

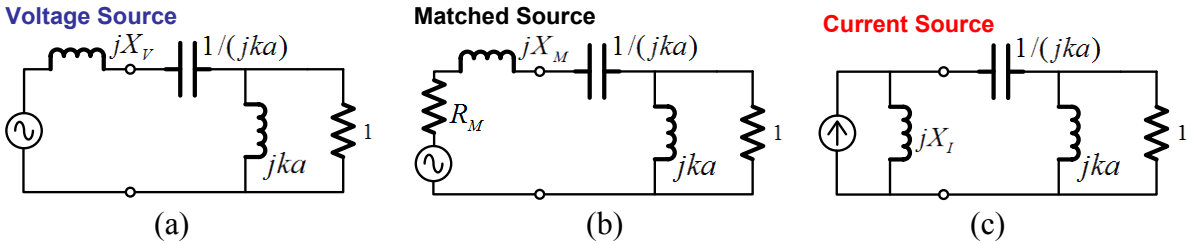


Figure 4.25 Equivalent circuits for spherical TM_{01} mode with various source conditions – (a) Voltage source, (b) Matched source, and (c) Current source. Note that X_V, X_M, R_M, and X_I are determined at resonance frequency.

As shown in Figure 4.26, different source conditions result in different radiation performances. At resonance, the radiation performance for all source conditions is the same, but with the small antenna ($k_0a = 1$), the radiation performance varies a lot around the resonant frequency. Therefore, when an electrically small antenna is required, one may consider different sources in addition to the conventional matched source to improve radiation performance. Secondly, since an antenna provides an impedance transformation from the system reference impedance (Z_0, typically 50 Ω) to the intrinsic wave impedance (η_0, 377Ω), this impedance transformation effect needs to be considered. An antenna may also have additional energy storage structures, corresponding to equivalent capacitance and inductance for an impedance matching. This additional energy storage structure may move the location of poles. Thirdly, a specific antenna structure may result in exciting a specific spherical mode strongly, as we observed in the Chapter 3. These effects may be coupled and eventually change the antenna performance and location of poles.

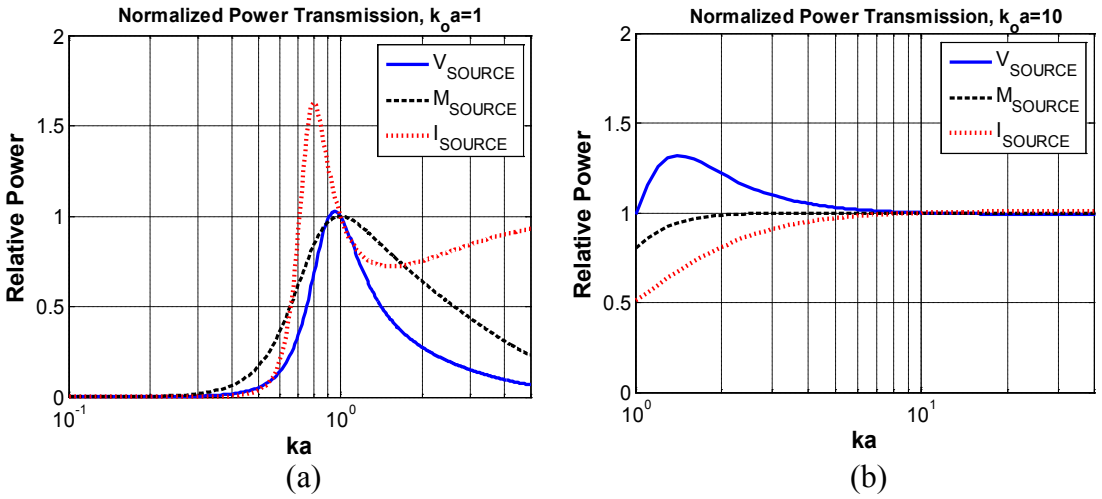
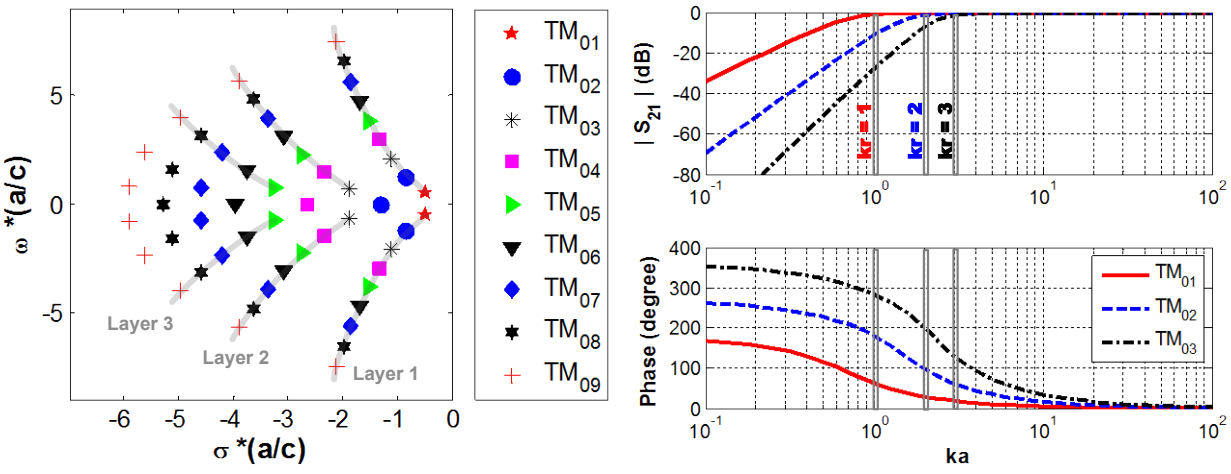


Figure 4.26 Comparison of source loading effects for various source conditions of the TM_{01} model – (a) Tuned at k_0a =1 and (b) Tuned at k_0a = 10 from the circuits in Figure 4.25. Each radiated power plot is normalized with the radiated power at resonant frequency.

The Bode plots of the pole-residue structures for some selected spherical modes in Figure 4.27b show that each mode has the characteristic of a high-pass filter. It is clearly seen that each mode has a cut-off frequency that increases with the mode order n. The product of the directivity (given in terms of the θ and ϕ performance of the mode) and $|s_{21}|^2$ for each spherical mode corresponds to the realized gain of the mode. Thus, the $|s_{21}|^2$ plot of the Figure 4.27b dictates the variation of gain with frequency. The s_{21} representation corresponds to the matched source condition. An antenna exciting a specific spherical mode will have relatively constant gain above the cut-off frequency. This suggests that the ideal antenna has ultra-wideband characteristics.

The pole structure of s_{21} for various spherical TM modes with a matched source is shown in Figure 4.27a, with an obvious layering of the poles by mode. A similar layering phenomenon of the poles was shown by Teche [*Tesche*, 1973], but Figure 4.27a clearly shows what spherical modes are actually involved in each specific layer. These poles offer additional insight into the performance of antenna radiation. The spherical TM_{01} mode is the simplest mode with s_{21} determined from the distance between the poles and the frequency of interest, plus an entire function. As the mode number (n) increases, the number of poles needed to represent the mode increases. For each mode, the poles with the smallest real part of the complex frequency typically dominate the late-time response of an antenna. The late-time response of the antenna is

one of the important characteristics in pulsed ultra-wideband systems because the late-time response affects the data throughput.



(a) (b)

Figure 4.27 Pole-residue and Bede plot analysis for the extended Chu's circuit model – (a) Pole locations in complex frequency plane for various spherical TM modes (c is the speed of light, a is radius of antenna sphere, ω is angular frequency, and σ is damping coefficient) and (b) Bode plot for various spherical TM modes. Radius of radian sphere corresponding to each mode is also shown in vertical lines.

The impulse responses of Chu's equivalent circuits are shown in Figure 4.28 for various spherical TM modes. The responses correspond to the late-time performance of the ultra-wideband, ideal antennas in the time domain. As the mode number increases, the pulse width narrows. An increase in the size of the antenna sphere (a) results in a wider pulse width, corresponding to a lower cut-off frequency. Because higher-order spherical modes have higher radiation-Q values, the resulting late-time responses have some ringing. However, this ringing decays fast.

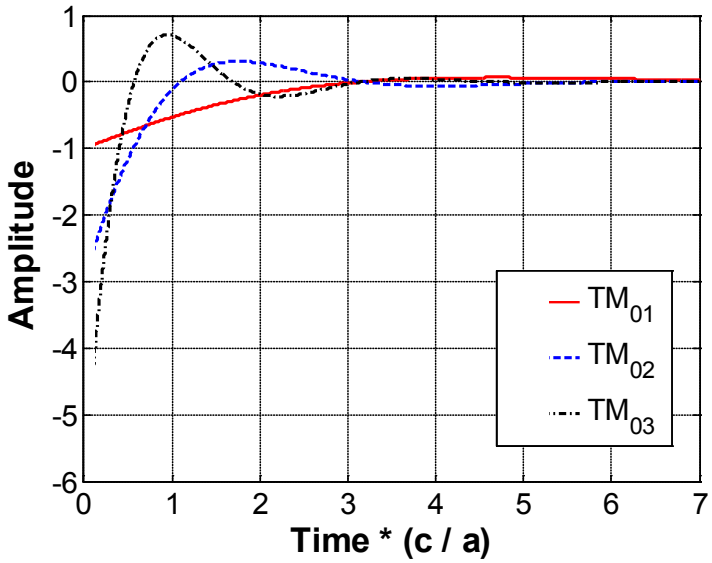
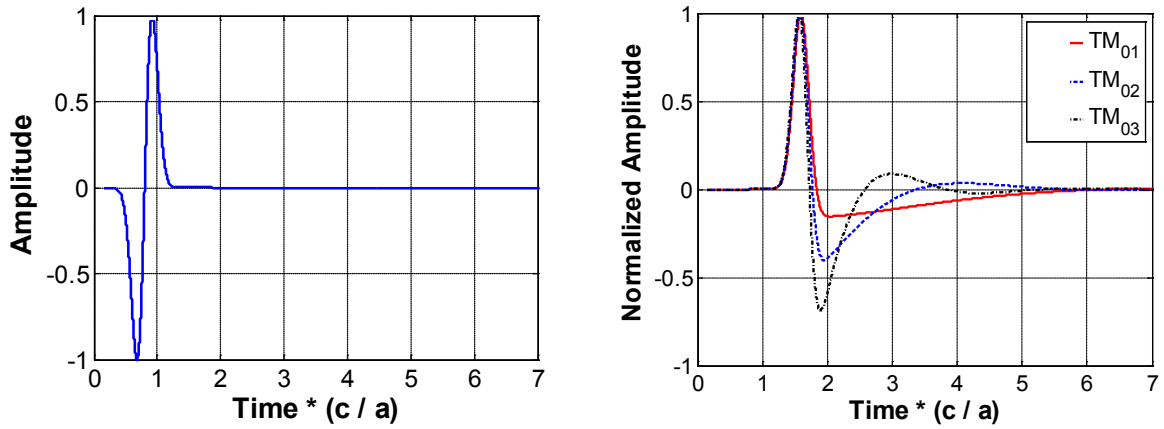


Figure 4.28 Impulse response of s_{21} for the extended circuit model.

Radiated pulses of several spherical TM waves in the far field are compared in Figure 4.29b. These pulses are the response of a Gaussian pulse input. However, the derivative of the Gaussian pulse (doublet waveform, see Figure 4.29a) was convolved with the impulse response of each spherical mode in order to obtain the radiated pulse, which dictates the radiated electric-field in the far field. The doublet was used since the circuit model does not explicitly include the derivative effect of antenna radiation process (see Chapter 6 for details). The radiated pulse of Figure 4.29b suggests that we need an antenna exciting higher-order spherical modes (i.e. high gain antennas) or an antenna having a ringing (i.e. high Q antennas) in order to obtain a doublet shape of the radiated pulse. Figure 4.29b also shows that an antenna exciting a fundamental mode over a large spectral range will have the radiated pulse shape of a Gaussian pulse for a Gaussian pulse input. These results are consistent with the measurement results of Licul [*Licul*, 2004]. Licul observed that the radiated pulse of a Vivaldi antenna (high gain antenna) had a doublet waveform while the radiated pulse of a bicone (exciting fundamental mode dominantly) had a Gaussian pulse shape for a Gaussian input pulse.



(a) (b)

Figure 4.29 Radiated pulses of spherical TM waves – (a) Input pulse and (b) Radiated pulse in the far field (*c* is the speed of light, *a* is radius of antenna sphere).

A finite-size bicone antenna (see Figure 4.30a) is a good example to demonstrate how the desired antenna functions can be implemented in a real antenna structure. The bicone antenna is a linearly-tapered structure, providing a wideband impedance transformation. To evaluate the performance of the bicone antenna, the antenna of Figure 4.30a was characterized with a commercial moment method code [*FEKO*, 2007]. As shown in Figure 4.30b, the upper bound of the impedance bandwidth of the bicone antenna is unbounded and is typically limited only by the accuracy details of the feed region. Thus, the bandwidth is extremely large and the conventional radiation-Q formula, based on impedance bandwidth, becomes inaccurate and loses meaning for the bicone antenna. The radiation Q evaluation formula of (4.127) [*Yaghjian and Best*, 2005] can be used, but it only provides performance at a single frequency. In addition, (4.127) is accurate if the antenna is electrically small. Thus, the previous radiation Q evaluation formulas are not applicable to determine overall size and performance limits for ultra-wideband or frequency-independent antennas.

The gain versus frequency plot in Figure 4.30c shows relatively constant gain over a wide frequency range (pattern is found to also be relatively constant: see the bicone antenna example in Chapter 3), implying a single spherical mode is strongly excited and dominates the antenna performance. Figure 4.30d shows the radiated E-field pulse at 30 m from the source for a 55 pico second Gaussian input pulse. This radiated E-field pulse dictates the antenna transfer function and has a shape of monocycle pulse with slight ringing.

Overall, the bicone antenna is a good ultra-wideband antenna. However, we still need to compare the size of the designed bicone antenna to the theoretical limit. The size of an ultra-wideband antenna is limited by the lower bound of operating bandwidth as characterized by the excited fundamental mode. Thus, we relate the size of an ultra-wideband antenna with the fundamental-mode 3-dB cut-off frequency of $|s_{21}|^2$ shown in Figure 4.27b. However, Chu's circuit model does not consider the actual antenna structure and the connection to the spherical modes.

The introduced antenna transfer function model of Chapter 3, based on the singularity expansion method (SEM) and the spherical mode expansion (SWE), can provide the relationship between an antenna input port and spherical wave mode excitation with scattering parameters. The excited mode of an arbitrary antenna can be found from measured or simulated near-field data relative to the input port by using spherical mode expansions.

As observed in Chapter 3, the bicone antenna excites the spherical TM_{01} mode dominantly over the entire frequency range and the lower bound of operational bandwidth is determined by the spherical TM_{01} mode. This observation suggests that the size of ultra-wideband antennas is limited by the excited fundamental spherical mode. Thus, the radiation performance of the bicone antenna as a function of size is best represented with the lower frequency cut-off (3-dB gain reduction) versus size, since the antenna is essentially unlimited on the upper frequency. When the 3-dB cut-off frequency of the bicone antenna is compared with the theoretical 3-dB cut-off frequency for the spherical TM_{01} mode in Figure 4.31, it is apparent that it might be possible to design an antenna with a size below that of the bicone without sacrificing radiation efficiency. However, this reduction offers a major challenge for an ultra-wideband antenna that avoids the trade-off in size, impedance bandwidth, and gain.

Thus, the size and performance limitations of frequency independent or ultra-wideband antennas can be characterized with the theoretical limits of pulse width and 3-dB cut-off frequency.

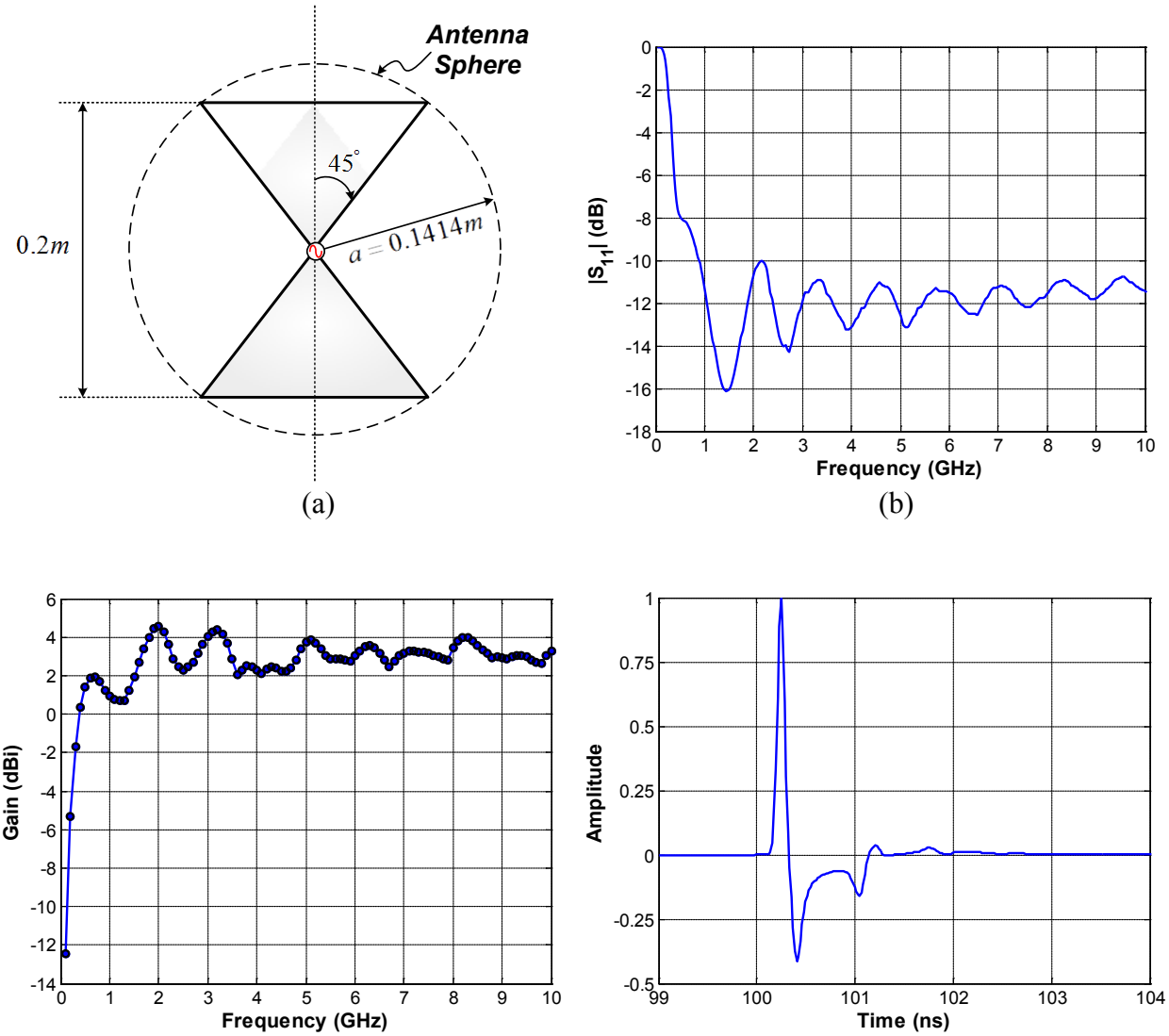


Figure 4.30 Example bicone characteristics: (a) Cross section view of finite-size bicone, (b) Return loss versus frequency ($Z_0 = 100\Omega$), (c) Gain versus frequency, (d) Radiated E-field at 30 m with Gaussian pulse excitation. Data were obtained from simulations using a commercial method of moments code [*FEKO*, 2007].

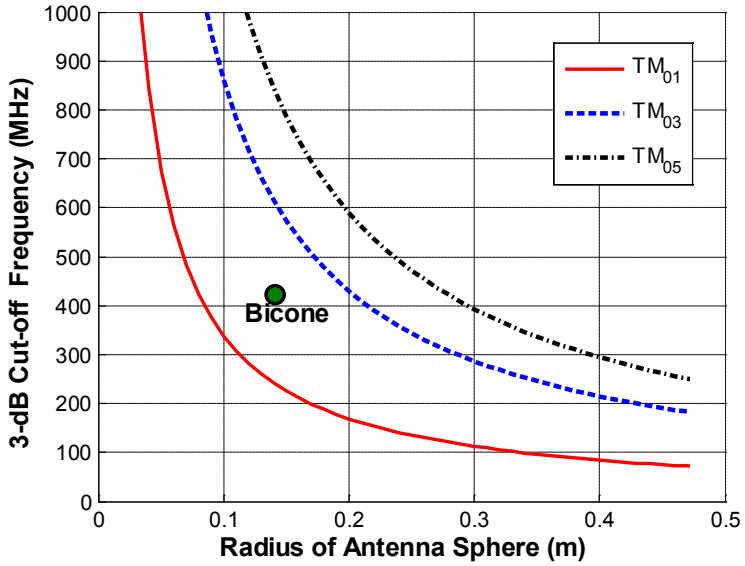


Figure 4.31 Fundamental limit 3 dB cut-off frequency versus antenna size for various spherical TM_{0n} modes and an example bicone.

4.4 Chapter Summary

In essence, the fundamental-limit theory for antennas shows that size, efficiency, and bandwidth are trade-offs in the design process. Based on the radiation Q versus ka curve, one can evaluate the feasibility of realizing specifications and determine how close the designed antenna is to the limit. Thus, the accuracy of the theoretical limit has been emphasized and corresponding research also has been carried out. In order to understand the errors and assumptions in historical limit theory research, classic fundamental limit theory was reviewed first. We showed that the frequency-domain approach of Chu [*Chu*, 1948], Collin and Rothschild [*R Collin and Rothschild*, 1964], and McLean [*McLean*, 1996] is essentially based on a far-field approximation for the radiated fields or the assumption that the radiating energy travels with the speed of light all the way from the antenna sphere to the far-field. As we observed from the investigation on antenna radiation and the energy storage mechanism in Chapter 2, both assumptions in the classic fundamental-limit theory are in error. Grimes [*Grimes and Grimes*, 1999] recognized that there may be an excess delay in the antenna radiation process and they tried to obtain a radiation Q limit with a new non-negative energy condition. However, their minimum radiation Q seemed to be a local minimum because their approach was based on the differential form of the Poynting theory with the power density and stored energy densities.

Based on the recent work of Davis *et al.* [*Davis et al.*, 2011], it was demonstrated how classic limit theory can be corrected using the time-average radial energy velocity and the non-negative energy condition. The corrected minimum radiation Q was consistent in both field and Chu's equivalent circuit approaches. The analysis based on the static energy of a charge dipole indicated that the minimum non-radiating energy of the spherical TM_{01} is identical to the amount of the static electric energy between the positive and negative charges, suggesting that the minimum non-radiating energy comes from the electric fields bound to charges and all decoupled energy contributes to far-field radiation. Our investigation of the minimum radiation Q was extended to consider the spherical TM_{11} (circular polarization) mode. The investigation showed that the minimum radiation Q of the spherical TM_{11} is the same as that in the spherical TM_{01} and TE_{01} modes, which suggests that all spherical modes with different degrees (m, ϕ behavior), yet with the same order (n, θ and r behavior), have the same minimum radiation Q. Our investigation of minimum radiation Q was further extended to consider an antenna with arbitrary polarization. The minimum radiation Qs were evaluated for various polarizations and feed configurations.

Classical fundamental-limit theory on antenna size was interpreted from an ultra-wideband antenna perspective. The frequency response of the extended Chu's equivalent model for spherical TM modes suggests the concept of an ideal antenna, with the entire frequency response described by only two complex poles and an entire function. It was shown that the ideal antenna naturally has ultra-wideband characteristics. The size limitations of ultra-wideband antennas in terms of pulse width and 3-dB cut-off frequency were developed and its usefulness was demonstrated with a bicone antenna example. The 3-dB cut-off frequency criterion can also be used to determine which antenna is more suitable in a frequency range of interest, either resonant, ultra-wideband, or frequency independent antennas.

The previous research activities over the last half century focused on the accuracy of the fundamental-limit theory itself rather than discussing why conventional antennas are not close to the theoretical limit and how we can make the performance of the antenna close to the limit. In addition, the trade-off relation of the limit theory between impedance bandwidth (inversely proportional to radiation-Q) and size of electrically small antennas has often been emphasized. However the basic radiation Q formula is directly related to the non-propagating near-field

energy, i.e. a low Q antenna has lower non-radiation energy than that of high Q antenna. Thus, a low Q antenna may show less interaction with objects in the vicinity of the antenna. This can be a clue as to how to mitigate some near-field interaction problems. These aspects have been overlooked for a long time.

Thus, based on what we learned from the fundamental-limit theory and the antenna radiation physics, we try to address these two questions in the next chapters:

1. How can we solve near-field interaction problems?
2. How we can make the performance of the antenna close to the limit?

Chapter 5

A Frequency-Domain Application: The Cellular Phone Near-Field Interaction Problem

In previous chapters, antenna radiation physics with energy storage mechanism, antenna transfer functions, and fundamental-limit theory on antenna size and performance were investigated. Observations during the investigation suggested that a low-Q antenna approach may mitigate near-field interaction problems due to the reduced stored energy in a reactive near-field region.

The low-Q antenna solution approach is demonstrated in this chapter through cellular phone near-field interaction problems as an example. Other similar near-field interaction problems are also addressed.

5.1 Overview of Cellular Phone Near-Field Interaction Problems

Cellular phones have become an indispensible item even for children. However, electromagnetic interference problems between cellular phones and medical devices have been reported in recent years [*Witters*, 2000]. The medical devices include electric powered wheelchairs, hearing aids, and cardiac pacemakers, as illustrated in Figure 5.1. Some hospitals have restricted the use of wireless phones inside the hospital buildings due to the concern of the interactions. However, it will be practically impossible to restrict the use of cellular phones in many public places.

The health effects of extended exposure to RF energy from handheld wireless transceivers have been debated for years. What is not debatable is that RF energy absorbed by the human operator is wasted energy. Thus, even from purely a performance standpoint, the goal is to reduce RF exposure. Thus, the goal is to reduce both cellular phone interaction with medical devices and human body exposures. Indeed, it is very challenging to mitigate the near-field interactions without scarifying the far-field communication performance.

In the following subsections, details of the interaction problems for a hearing aid and human head are explained in order to understand the problem. Related Federal Communications Commission (FCC) regulations are also briefly reviewed.

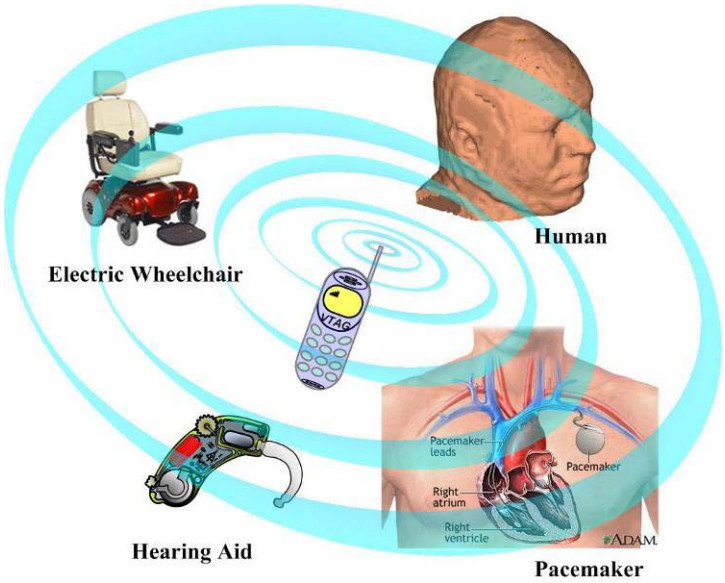


Figure 5.1 Various near-field interaction issues around cellular phone.

5.1.1 Interaction with Hearing-Aid

Demographers estimate that there are approximately 28 million people in the U.S. who suffer from hearing impairment, and 5.6 million of them wear hearing aids [*NIDCD*, 2006]. Unfortunately, these people are troubled by a buzzing noise in the hearing aid caused by interference from digital cellular phones. Though not all combinations of digital cellular phones and hearing aids have interference, in some cases the resulting buzzing noise is overwhelming and effectively prevents the user from hearing the cell phone audio. Interference can also be caused by a person using a digital cellular phone in close proximity to a person wearing a

hearing aid. The hearing-aid industry, the cell-phone industry, and consumers have been collaborating to solve this incompatibility problem. As a consequence of the efforts, the American National Standards Institute (ANSI) formed a task group (C63.19) in 1996 [*Berger*, 2001] to address the issue. The hearing aid compatibility (HAC) standard, developed in 2001 and updated in 2006 [*ANSI*, 2006] and 2011 [*ANSI*, 2011], provides acceptable performance levels for the measurement and evaluation of wireless phone, near-field strength. The standard also addresses immunity levels for hearing aids. The new rules, approved on April 9, 2011 by FCC (effective on August 16, 2012), allow the use of either ANSI C63.19-2007 or ANSI C63.19-2011 to test new wireless phone models during 12 month transition period from the effective date. After the transition period, cellular phone industry has to comply with the new ANSI standard.

An FCC rule and order in 2003 requires that phone manufacturers and carriers offer at least two phones that meet the HAC standard. Specifically, phones that can be used with hearing aid telecoils were required by 2006 and at least half of all mobile telephone handset models must not exceed the specified RF interference level by February 18, 2008. Thus, the industry has been seeking a solution to meet the FCC requirement [*Tom Victorian*, 2004]. Since then, industry's efforts continue to comply with updated FCC regulations. Early solutions, such as a neck loop or hands-free option, provide enough distance between the cellular phone and the hearing aid to mitigate the interference [*Berger*, 2001]. An emission blocker which claims to create a radiation null toward the hearing aid is also available on the market. These early solutions were supplied in the form of accessories and were often considered to be inconvenient. Furthermore, these solutions do not satisfy the FCC requirements that address the RF emission from the wireless phone. Many efforts shown in the literature have focused on measurement, modelling, and evaluation of the interference and RF emission [*Caputa et al.*, 2000; *Okoniewski and Stuchly*, 1998; *Skopec*, 1998], not on the fundamental issue of controlling field strength levels.

5.1.1.1 Structure of Hearing Aid

There are three basic types of hearing aids: behind-the-ear (BTE), in-the-ear (ITE), and in-the-canal (ITC), named in accordance with placement. The structure of the conventional BTE hearing aid is shown in Figure 5.2.

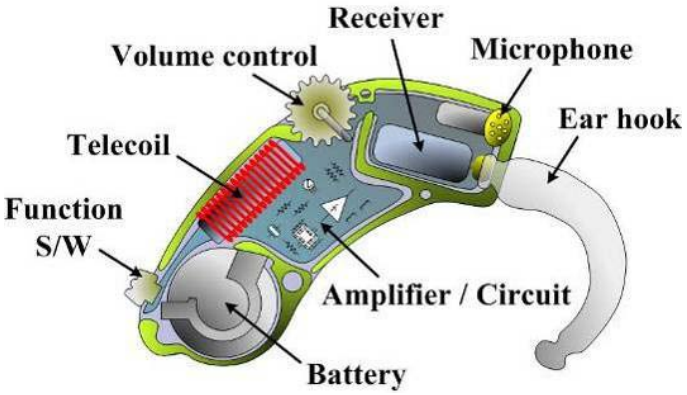


Figure 5.2 Structure of behind-the-ear (BTE) hearing aid [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

Either an acoustic mode or a telecoil mode can be selected. In the acoustic mode, a microphone converts acoustic energy to electrical energy, and in the telecoil mode, the telecoil detects magnetic-field leakage from the phone speaker and converts the detected magnetic energy to electrical energy. The converted electric energy is amplified and converted again to acoustic energy by the receiver, finally being delivered to the ear through the ear hook [*Dillon*, 2001].

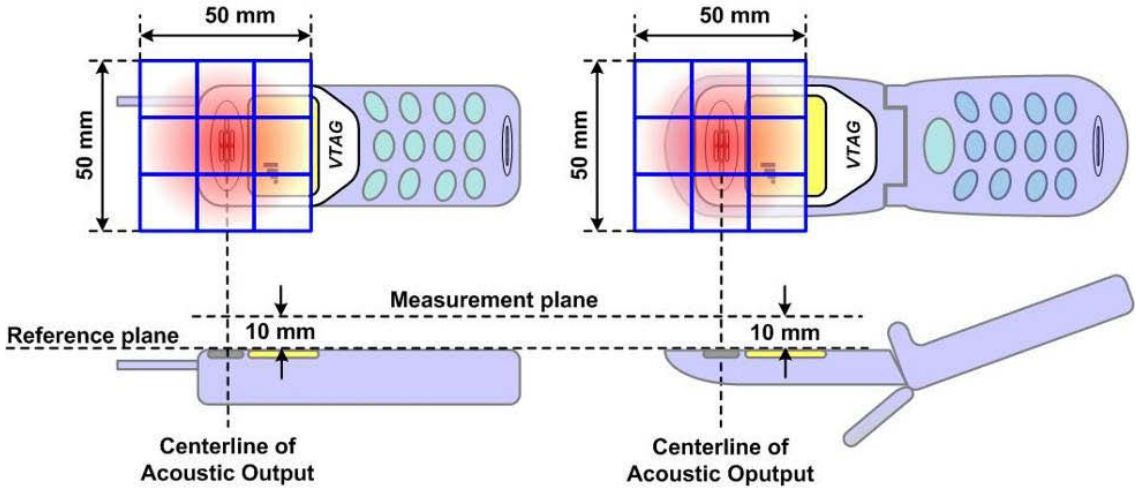
The advantage of the telecoil mode is that no audio background noise is received. Therefore, hearing-aid users receive a better audio signal in the telecoil mode than in the acoustic mode. However, the telecoil can act as an unintentional antenna and couple energy from cellular phones in close proximity to the hearing aid. The coupled energy through this unintentional antenna is demodulated, also unintentionally, in the amplifier circuits of the hearing aid, causing a buzzing noise to hearing-aid users. The typical frequencies of the interfering acoustic signal are 50 Hz, 217 Hz and their harmonics. In the acoustic mode, the microphone becomes the unintentional receiving antenna [*Kuk and Nielsen*, 1997; *Skopec*, 1998; *TA Victorian*, 1998].

The cellular phone antenna is the major source of strong electromagnetic energy and causes interference in both the acoustic and telecoil modes. In addition, any functional operation associated with a varying current in a cellular phone, such as battery surge currents, power-supply components, keyboard scanning, and display refresh, can induce magnetic fields and create interference in the telecoil mode [*Skopec*, 1998; *TA Victorian*, 1998]. In the following discussion, we focus on RF energy emission from the cellular-phone antenna.

5.1.1.2 Hearing Aid Compatibility (HAC) Standard

HAC standard ANSI C63.19-2006 [*ANSI*, 2006] covers wireless devices from 800 MHz to 3 GHz. The telephone emission below 960 MHz was lowered by 10 dB from the 2001 HAC standard. Therefore, it is likely equipment will meet the HAC standard below 960 MHz, but a significant challenge remains above 960 MHz.

Figure 5.3 illustrates the 50 mm square measurement area defined in the standard. The meshed area located in the measurement plane (10 mm above the reference plane) is evenly divided into 9 cells and is centered on the wireless-device speaker. Field strength measurements are made by scanning a probe over this meshed area. Based on the measured peak near-field strength, the cellular phone is placed into one of the categories listed in Table 5-1. The categories are 5 dB apart and GSM cellular phones have an additional -5 dB weighting factor. In the far-field region, the electric and magnetic fields are related by the intrinsic wave impedance. However, in the near-field region, the electric and magnetic fields are not strongly coupled and cannot be related in terms of the intrinsic wave impedance. In addition, there may be other magnetic interference sources, as explained in the previous section. Therefore, the HAC standards categorize the electric and magnetic fields separately.



(a) Candybar type (b) Cramshell type

Figure 5.3 Reference and measurement planes for field strength measurement complying to ANSI standards C63.19-2006 [*ANSI*, 2006; *Yang et al.*, 2008]. Copyright © 2008 IEEE.

Table 5-1 Telephone emission limits and categories [*ANSI*, 2006]

Category	Telephone RF Parameters (above 960 MHz)	
Near-field	**E-field emission (dBV/m)**	**H-field emission (dBA/m)**
Category **M1**	46 to 51 + 0.5 × AWF	-4.4 to -0.6 + 0.5 × AWF
Category **M2**	41 to 46 + 0.5 × AWF	-9.4 to -4.4 + 0.5 × AWF
Category **M3**	36 to 41 + 0.5 × AWF	-14.4 to -9.4 + 0.5 × AWF
Category **M4**	< 36 + 0.5 × AWF	< -14.4 + 0.5 × AWF
AWF(Articulated weighted factor): UMTS(WCDMA) = 0, CDMA = 0, TDMA = 0, and GSM = -5.		
Note: The levels below 960 MHz are 10 dB below the values indicated.		

The standard for hearing aid immunity to RF interference is measured using a balanced dipole antenna and has four similar categories (also denoted by M1 – M4). The FCC requires HAC compliance for both hearing aids and cellular phones independently, providing a variety of choices for users to select a combination of hearing aids and cellular phones. The HAC standard treats the hearing aid and cellular phone together as a system, classifying the system performance by summing the category values of each as in Table 5-2. For example, a phone with M2 emission level and a hearing aid with M3 immunity fall in the normal use category (2+3=5). A category sum below 4 suggests the system is not likely to be useable.

Table 5-2 System performance classification [*ANSI*, 2006].

System Classification	Category Sum (Hearing aid category + Telephone category)
Useable	4
Normal Use	5
Excellent Performance	6

The recently-adopted HAC standard ANSI C63.19-20011 [*ANSI*, 2011] covers wireless devices from 698 MHz to 6 GHz to take into account for newer wireless technologies. H-field emission testing is no longer required. The distance between wireless device and the field probe changed from 10 mm to 15 mm. The articulated weighted factor (AWF) compensating for a

different communication modulation is not going to be used. Instead, the E-field emission range of each category has been decreased by 6 dB, compared to the previous case of zero AWF weight. For example, the E-field emission range of Category M1 in the ANSI C63.19-20011 is 40 to 45 dBV/m for frequency above 960 MHz. The 10 dB relaxation for frequencies below 960 MHz still remains same, i.e. 50 to 55 dBV/m. Overall, it appears that the new HAC compliance requirements are relaxed, compared to the previous ANSI C63.19-2006.

5.1.2 Interaction with Human Body

Since cellular phones are typically used in very close proximity of human operators, there have been concerns about possible biological effects of radio-frequency (RF) radiation from a cellular phone. RF exposure standards [*Ahlbom et al.*, 1998; *ANSI/IEEE*, 2006; *CENELEC*, 1995; *IEC*, 2005] have been developed by various organizations and countries. The cellular phone safety guidelines are largely based on the biological effects caused from heating of tissue by electromagnetic energy, i.e. thermal effect for low-power non-ionizing radiation cases (< 2 Watts in practice). Non-thermal effects due to the long-term exposure from low-level RF energy have not been included in these safety guidelines [*Cleveland et al.*, 1999]. But concerns remain [*de Salles*, 1999; *Goldberg*, 1996] about biohazard, from long-term exposure.

The cell phone interactions with human operators arise due to strong fields, similar to medical device cases. However, unlike the hearing aid case of the previous section, the thermal effects of cellular phones on the human head are mainly affected by electric near fields. This is because the human body consists of mainly electric materials, not magnetic materials. Thus, a better antenna type to reduce the electric near fields would be a magnetic-field driven antenna such as a loop antenna. However, the operational bandwidth of magnetic dipole antennas is typically narrower than that of electric dipole antennas. This is because the radiation resistance of loop antennas decreases with $(frequency)^4$ while the radiation resistance of dipole antennas decreases with $(frequency)^2$. If the antenna size is decreased, bandwidth becomes must be further narrowed in order to maintain the same radiation efficiency. As a result, alternative approaches have been introduced in the literature, including a material loading [*Kitra et al.*, 2007], increasing the distance between the antenna and the human body, or by field blocking [*Kusuma et al.*, 2011] to redirect the peak electric near field away from human head. However, these alternative solutions require additional material (cost issue) or a blocking structure (size issue).

In addition, they do not address the fundamental problem, i.e. strong electric near fields generated from cellular antenna.

5.1.2.1 The Structure of Cellular Phone

The current cellular phone market is heavily directed toward smart phones. Compared to old feature phones, smart phones provide a better user experience by adopting sensors, finger touch input, an intuitive user interface, high-speed data connectivity, and a powerful processor with a variety of applications. In order to maximize user experience, smart phones typically have a large flat display with a touch input interface and high-speed data connectivity (web surfing) feature. Thus, many smart phones have a thin brick shape, as shown in Figure 5.4.

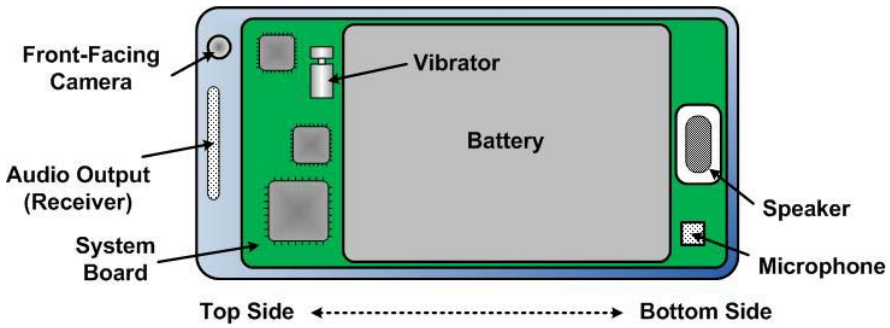


Figure 5.4 Example of smart phone structure with typical internal components.

In the early generation of cellular phones like feature phones, a long linear antenna was located outside of the housing at the top of the phone. Later, a retractable antenna was used, but soon was replaced by internal antennas in order to avoid visual impact and physical antenna damage. Because there are many components inside cellular phones, the effective internal antenna volume was reduced. Extra loss from the components and phone external materials also had to be considered in the design of the internal antenna. Basically, a more efficient antenna is needed with small volume that meets antenna performance specifications, including total radiated power (TRP) and total isotropic sensitivity (TIS). At the same time, antenna engineers have to consider radio frequency (RF) exposure level compliance. Due to the limited antenna footprint and location choices inside the cellular phone, cellular antenna design requires state-of-the-art skills to meet both performance and regulatory exposure level specifications. So, some cellular phone designs have the antenna on the bottom of the phone because the near-field

interaction can be reduced due to larger spacing between the antenna and the human body. Since the hearing aid compatibility regulations requiring both electric and magnetic fields began, most cellular phone manufactures have located the antenna at the bottom of the phone. However, antennas located at the bottom of cellular phones tend to experience hand effects, i.e. frequency detuning and significant loss, based on how users are holding the phone. Thus, there are trade-offs between antenna placement, antenna performance, and user experience.

Many contemporary smart phones have a penta-band main antenna at the bottom and a receive diversity antenna at the top. In addition, antennas for global positioning system (GPS), dual-band wireless local area network (WLAN), Bluetooth, near field communication (NFC), and wireless charging antennas need to be allocated together with the main and diversity antennas inside the smart phone. These antennas also need to be collocated with other components including speaker, vibrator, microphone, receiver, and camera. Thus, smart phone antenna design is very challenging.

The emerging "true" 4^{th}-generation cellular network requires multiple-input multiple-output (MIMO) technology, which may require multiple antennas for both transmission and reception. The multiple antennas need to be separated in distance (spatial diversity) and /or orthogonal in radiation pattern (pattern diversity) to maximize a diversity gain. Alternatively, MIMO can be supported through multiple bands (frequency diversity), but this configuration requires multiple RF front-end chips and/or corresponding increased loss and RF layout size. For spatial diversity case, one of MIMO antennas is most likely located at top side of the smart phone, creating near-field interaction issues with human head and hear-aid.

The limited volume and increased number of antennas in smart phones cause antenna design challenges in meeting RF exposure compliance and antenna performance specifications.

5.1.2.2 Specific Absorption Rate (SAR) Standard

As is well known, more than half of human body is basically water [*Guyton and Hall*, 2006]. When a time-varying electric field is applied to the body, water molecules vibrate. The collision between molecules causes heat. Thus, the applied electric energy is converted into the heat. Other electrical materials in human body show a similar loss mechanism to the water case, which is a microscopic or a quantum mechanics phenomenon.

From a macroscopic viewpoint, the loss mechanism is modeled by using loss factor. For electrical materials, the relative permittivity (ε_r) can be defined as [*R F Harrington*, 2001]

$$\varepsilon_r = \varepsilon_r' - j\varepsilon_r'' \tag{5.1}$$

where ε_r' (dielectric constant) is related to electric energy storage and ε_r'' is related to energy dissipation or loss in the media. In general, the relative permittivity varies with frequency. Alternatively, electric conductivity (σ_e) can be used to account for the energy dissipation and is written as

$$\sigma_e = 2\pi f \varepsilon_0 \varepsilon_r'' \tag{5.2}$$

where f is frequency in interest and ε_0 is free-space permittivity ($\approx 0.854^* 10^{-12}$ F/m). Then, the absorbed power (P_a) per unit volume is proportional to the electric conductivity and the peak magnitude of the electric field ($\left|\vec{E}\right|$) [*Stutzman and Thiele*, 2012], i.e.

$$P_a = \frac{1}{2}\sigma_e \left|\vec{E}\right|^2 \tag{5.3}$$

In terms of the absorbed power per unit weight viewpoint, absorption power ratio (SAR) can be defined as

$$SAR = \frac{P_a}{\rho} = \frac{1}{2}\frac{\sigma_e \left|\vec{E}\right|^2}{\rho} \tag{5.4}$$

where ρ is material density in $\mathrm{Kg/m^3}$. SAR is expressed in units of W/Kg or equivalently mW/g.

Averaging SAR over a specific volume of body has been found to be useful in practice. Thus, peak spatial-average SAR has been used as a figure of merit to qualify the thermal effects from RF radiation [*ANSI/IEEE*, 2003]. For example, a specific anthropomorphic mannequin (SAM) head phantom has been popular for SAR performance compliance testing in wireless mobile industry. The shell of SAM phantom head is made of low permittivity ($\varepsilon_r' < 5$) and low loss ($\tan\delta_\varepsilon = \varepsilon_r'' / \varepsilon_r' < 0.05$) materials [*Seabury*, 2005]. A homogeneous tissue equivalent liquid is filled inside the SAM head shell. The dielectric properties of the liquid in some selected frequencies are listed in Table 5-3.

Table 5-3 Dielectric properties of liquid for SAM head phantom [*ANSI/IEEE*, 2003]

Frequency (MHz)	Relative Permittivity (ε_r')	Electric conductivity (σ_e)
300	45.3	0.87
450	43.5	0.87
835	41.5	0.90
900	41.5	0.97
1450	40.5	1.20
1800 - 2000	40.0	1.40
2450	39.2	1.80
3000	38.5	2.40

Typically, SAR measurement is carried out using a multi-axis robot to locate a small electric-field probe inside the SAM phantom. So, a small area on the top of the SAM phantom is open for the automated scanning of the probe. Alternatively, open half of the phantom (cut in a vertical direction) can be used because the peak SAR location is typically in the half phantom closest to the wireless device. After scanning, the peak spatial-average SAR is determined through a post processing.

The head SAR measurement procedure defined in [*ANSI/IEEE*, 2003] requires testing in both "cheek" and "tilt" modes, which are illustrated in Figure 5.5 and Figure 5.6, respectively. An additional requirement is that the wireless mobile device in the SAR test has to be set to a condition giving maximum radiated power. For example, multiple radios are used simultaneously for some smart phones to support simultaneous voice and data services. In this case, SAR testing has to be carried out with all the radios are on. Thus, the worst peak spatial-average SAR value should meet the required SAR compliance limit specified in Table 5-4 for various countries.

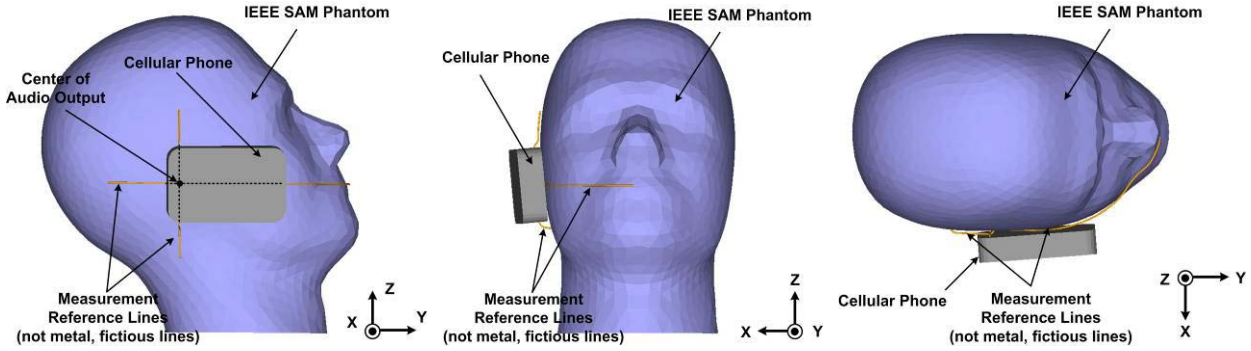


Figure 5.5 Cellular phone position in cheek mode – (a) Side view, (b) Front view, and (c) Top view.

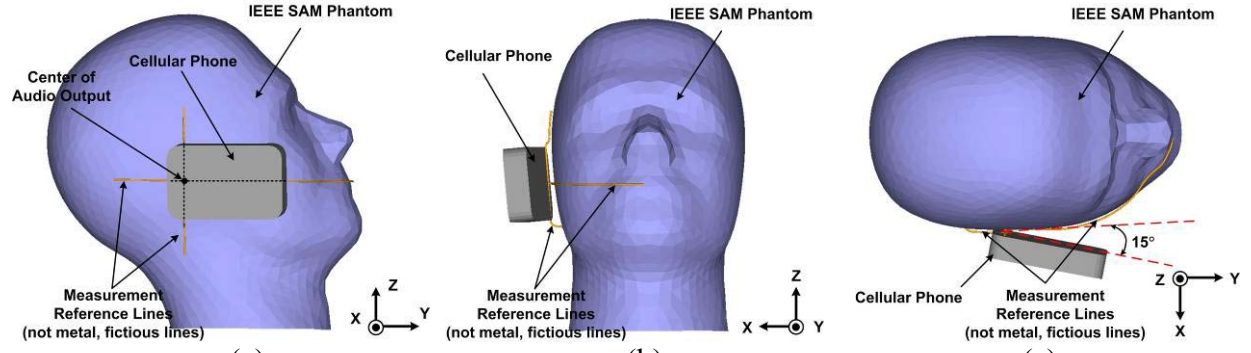


Figure 5.6 Cellular phone position in tilt mode with 15 degree angle – (a) Side view, (b) Front view, and (c) Top view.

Table 5-4 Summary of peak spatial-average specific absorption rate (SAR) limits

	International (ICNIRP)	**United States (ANSI/IEEE)**	**European Union (CENELEC)**
Whole body SAR	0.08 W/kg	0.08 W/kg	0.08 W/kg
Local SAR	2 W/Kg	1.6 W/Kg	2 W/kg
Average Mass	100 g (continuous tissue)	1 g (cube)	10 g (cube)

5.2 A Low-*Q* Antenna Solution Approach

In a classical view, the source of antenna radiation is the acceleration/deceleration of charges moving along a metallic antenna structure, corresponding to equivalent time-varying currents on the structure. The generated electric and magnetic fields from the electric and magnetic currents ($\vec{J}$ and $\vec{M}$) are given by [*R F Harrington*, 2001; *Stutzman and Thiele*, 2012]

$$\begin{bmatrix}\vec{E}\\ \vec{H}\end{bmatrix}=\begin{bmatrix}-\frac{1}{\varepsilon}\nabla\times\vec{F}+\frac{1}{j\omega\mu\varepsilon}[k^2\vec{A}+\nabla(\nabla\bullet\vec{A})]\\ \frac{1}{\mu}\nabla\times\vec{A}+\frac{1}{j\omega\mu\varepsilon}[k^2\vec{F}+\nabla(\nabla\bullet\vec{F})]\end{bmatrix} \tag{5.5}$$

where $\vec{A}$ and $\vec{F}$ are magnetic and electric vector potentials respectively and $k=2\pi/\lambda$. The vector potentials can be written in terms of the Green's function, G, as

$$\begin{bmatrix}\vec{A}\\ \vec{F}\end{bmatrix}=\begin{bmatrix}\mu\\ \varepsilon\end{bmatrix}\int_V G(R)\begin{bmatrix}\vec{J}\\ \vec{M}\end{bmatrix}dv' \tag{5.6}$$

King pointed out that the vector potential dictates the spatial variation of the current distribution [*R E Collin and Zucker*, 1969].

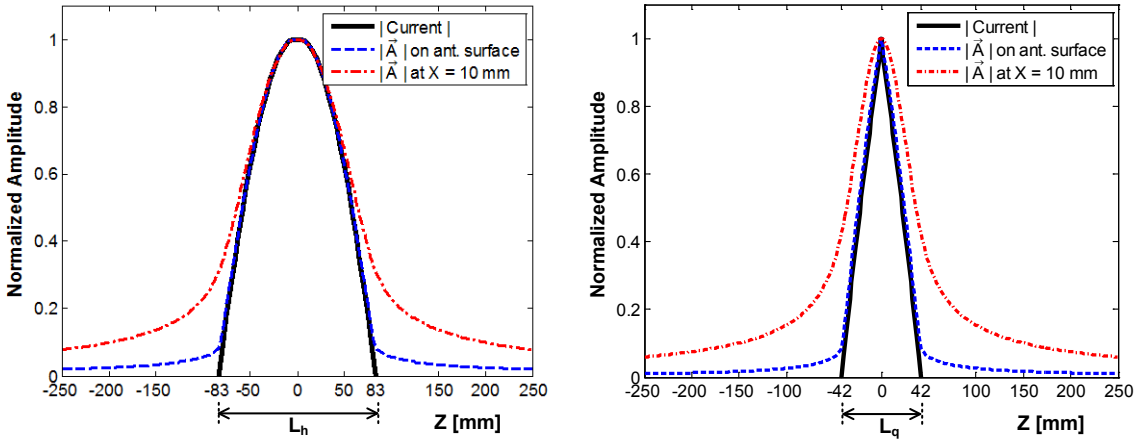
To illustrate this effect, the normalized $|\vec{A}|$ and $|\vec{J}|$ distributions of *z*-axis directed half and quarter-wavelength dipoles are compared in Figure 5.7. The variations of $|\vec{A}|$ and $|\vec{J}|$ are similar along the surface of antenna cylinder, with less variation of $|\vec{A}|$ near the ends of the dipole. If the normalized $|\vec{A}|$ is evaluated at a larger radial distance from the antenna, the distribution becomes smoother, but the normalized $|\vec{A}|$ still follows the basic behavior of the normalized $|\vec{J}|$.

Therefore, a current slope discontinuity can cause a significant increase of the near-field strength because the double derivative $\nabla(\nabla\bullet\vec{A})$ and $\nabla(\nabla\bullet\vec{F})$ terms in (5.5) dominate the near fields.

The conventional half-wavelength dipole in Figure 5.7a has a slope discontinuity in the current at the each end of the wire. Thus, we can expect a strong electric near-field, $|\vec{E}|$, around the ends of the wire. For a quarter-wavelength dipole, there is an additional slope discontinuity in current at the feed point (center), resulting from the approximately triangular current distribution (Figure 5.7b). As antenna length decreases, interference issues increase due to the increasing current slope discontinuity. This current slope discontinuity partially explains the following observation by [*Kuk and Nielsen*, 1997]: "Cellular phones with built-in and short antennas typically cause more interference than those with a long antenna."

To mitigate the current slope discontinuity, a fat dipole with a tapered feed can be used to spread the current over a wide surface. It is also well known that the impedance bandwidth of

dipoles increases as the dipole cross section increases. Thus, the fat dipole with a tapered feed would be an ultra-wideband antenna.



(a) (b)

Figure 5.7 Comparison between the normalized magnitude of the current distribution and the magnetic vector potential ($\vec{A}$) on the surface and at 10 mm distance along the Z-axis (X = 10 mm) for (a) 900 MHz half cylindrical (L_h / 2 = 83.33 mm) and (b) 900 MHz quarter wavelength cylindrical dipole (L_q / 2 = 41.67 mm). Radius of antenna cylinder is 0.01 mm and each curve is normalized [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

On the other hand, the definition of radiation Q in (4.2) shows that the radiation Q is proportional to the amount of non-radiating near-field energy. This relationship indicates that an antenna with smaller non-radiating near-field energy has a lower radiation Q. Since the radiation Q is approximately inversely proportional to the impedance bandwidth, an antenna with wide instantaneous impedance bandwidth, such as an ultra-wideband antenna, inherently has low non-radiating energy in the near field.

As explained in Chapter 2, the non-radiating near fields circulating between the antenna feed and the radian sphere boundary is the cause of the near-field interaction problem. Therefore, a low-Q antenna, such as an ultra-wideband antenna, will provide low near-field energy around the antenna and minimize interactions with surrounding objects.

In the following sub sections, the basic concept of low-Q antenna solution approach is demonstrated with simple thin and fat dipole examples.

5.2.1 Near-Field Strength versus Distance

In order to demonstrate this low-Q antenna approach, a half wavelength dipole (high-Q) and a planar fat circular dipole (low-Q), shown in Figure 5.8, were investigated using a commercial computational code based on the method of moment [*FEKO*, 2006].

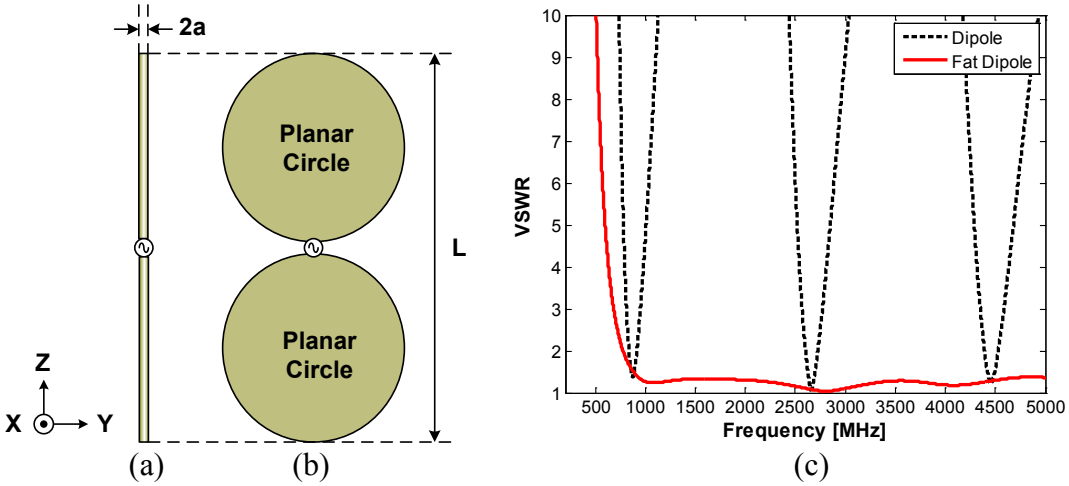
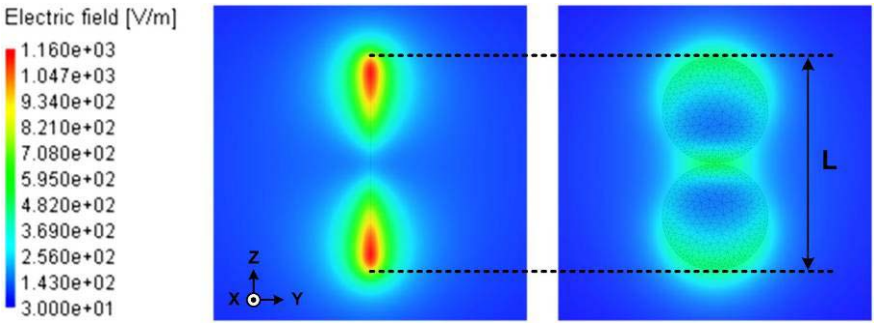


Figure 5.8 Antenna geometry and VSWR performance [*Yang et al.*, 2008] - (a) Dipole antenna (narrow band, high Q), (b) Fat dipole antenna (ultra-wideband, low Q), and (c) VSWR comparison. Note that L = half wavelength at 900 MHz (166.67 mm) and a = 0.01 mm.

The planar, fat circular dipole is an ultra-wideband variation on the conventional dipole antenna [*Yang et al.*, 2004]. The power delivered to both antennas was 32 dBm. As shown in Figure 5.8c, the conventional half-wavelength dipole is narrowband (high-Q) with periodic resonances at odd harmonics of 900 MHz. The planar, fat circular dipole shows an ultra-wideband, high-pass characteristic. Both antennas have standard dipole radiation patterns, with the radiation pattern of the fat dipole slightly more directive in the plane of the antenna due to the slight focusing effect of the taper. The azimuthal variation is less than 0.5 dB at 900 MHz, providing a reasonably omni-directional pattern.

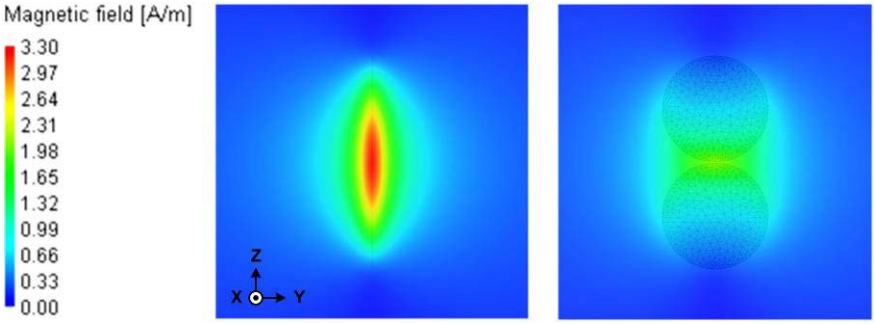
The E-field distribution at the perpendicular distance X = 10 mm is plotted in Figure 5.9. As predicted from the theory, the E-field is concentrated near the wire ends of the conventional dipole (Figure 5.9a) and is distributed over the circular surface of the circular fat dipole (Figure 5.9b). Similarly, the H-field strength of the fat dipole, shown in Figure 5.9, is significantly lower than the conventional dipole, due in large part to the wider distribution of the current. As summarized in Table 5-5, the peak $|\vec{E}|$ and $|\vec{H}|$ near-field 10 mm away from the surface of the

ultra-wideband, circular, fat dipole (low-Q) is reduced from the simple dipole (high-Q) by approximately 7 dB and 4 dB respectively. Nevertheless, the far-field performance is almost the same. Because there is 5 dB between HAC emission limits categories (see Table 5-1), the 7 dB reduction offers a significant improvement.



(a) (b)

Figure 5.9 Comparison of the calculated E-field amplitude distribution at 900 MHz (10 mm away from the antenna surface) [*Yang et al.*, 2008] – (a) Dipole and (b) Fat dipole. Copyright © 2008 IEEE.



(a) (b)

Figure 5.10 Comparison of the calculated H-field amplitude distribution at 900 MHz (10 mm away from antenna surface) [*Yang et al.*, 2008] – (a) Dipole and (b) Fat dipole. Copyright © 2008 IEEE.

Table 5-5 Comparison of $|\vec{E}|$ and $|\vec{H}|$ amplitudes for 900 MHz conventional and fat dipoles at 10 mm distance from the antenna surface [*Yang et al.*, 2008]

	Dipole (Narrowband)	**Fat Dipole (Ultra-wideband)**	**|Difference|**
Peak |E|	1153.96 (61.24 dBV/m)	502.20 (54.02 dBV/m)	7.22 dB
Peak |H|	3.27 (10.29 dBA/m)	2.03 (6.15 dBA/m)	4.14 dB

In order to investigate the relationship between antenna size and the effectiveness of the low-Q antenna approach, the peak $|\vec{E}|$ and $|\vec{H}|$ values versus distance from the antenna were compared for conventional and fat dipoles of both half- and quarter-wavelengths. The quarter-wavelength antenna was tuned to deliver 32 dBm of power as in the previous simulation. As seen in Figure 5.11, both half-wavelength low- and high-Q antennas have nearly the same field strength beyond the antenna sphere (radian sphere < antenna sphere). Similarly, both quarter-wavelength antennas have almost the same field strength beyond the radian sphere (radian sphere > antenna sphere). These observations suggest that a low-Q antenna lowers the near fields within the larger of the radian or antenna sphere, while retaining the same radiation performance beyond the larger sphere.

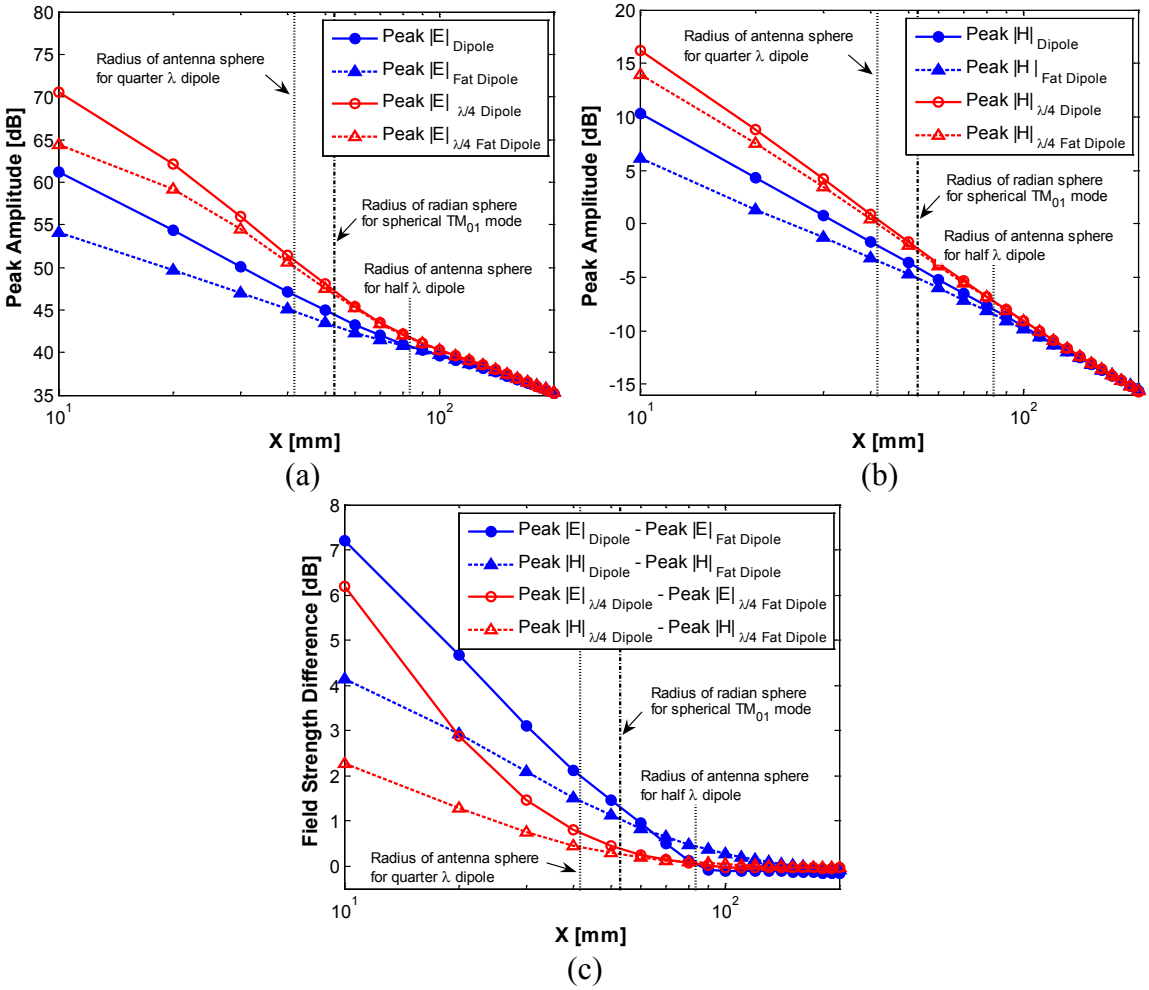


Figure 5.11 Comparison of fields found by simulations for half- and quarter-wavelength dipoles [*Yang et al.*, 2008] – (a) Peak $|\vec{E}|$, (b) Peak $|\vec{H}|$, and (c) Difference of peak field strength versus a distance (X) from antenna. Copyright © 2008 IEEE.

5.2.2 Near Field Strength and SAR versus Distance in the Presence of a Simple Human Head Model

To demonstrate that a low-Q antenna will reduce electric and magnetic fields in the vicinity of the human head, two antennas with high and low radiation Qs were considered. The thin dipole (high-Q) and the planar fat-dipole with a tapered feed (low-Q) are shown in Figure 5.12a and Figure 5.12b, respectively. Both antennas have a half-wavelength height at 1880 MHz with 50 degree angle bend (see Figure 5.12c). The human head model is a small dielectric sphere, which represents a child's head (radius of the head = 45 mm). The dielectric characteristics of the sphere were set to $\varepsilon_r' = 40$, $\sigma_e = 1.387$ S/m, $\rho = 1000$ Kg/m^3 to provide an approximate macro model for the head. The detailed simulation configuration for each antenna is illustrated in Figure 5.13.

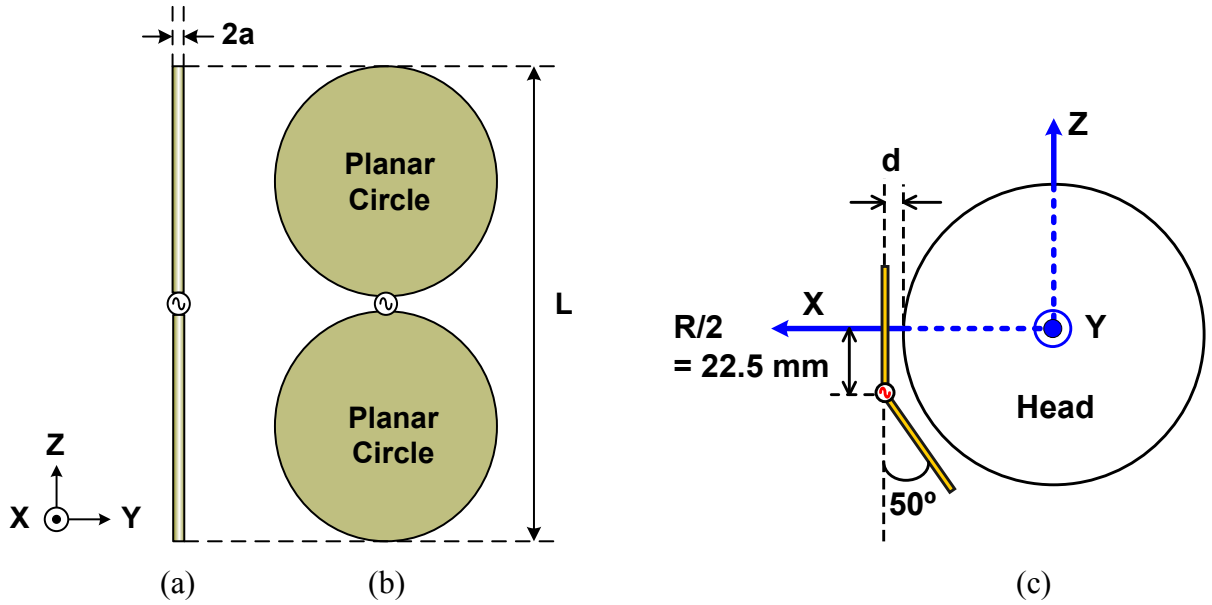
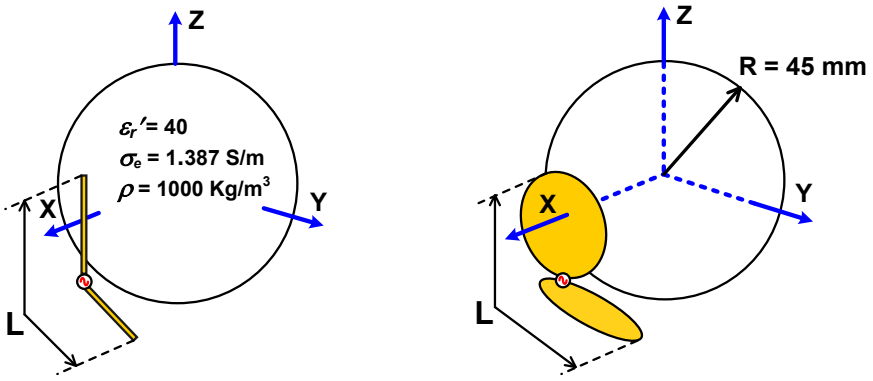


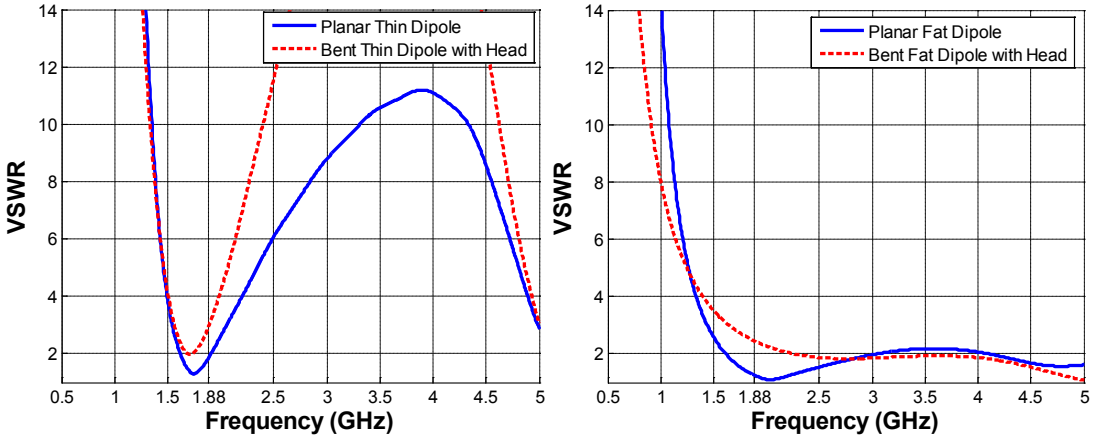
Figure 5.12 Antenna geometry and location from simple child head model – (a) Planar thin dipole antenna (narrow band, high Q), (b) Planar fat dipole antenna (ultra-wideband, low Q), and (c) Configuration of a bent antenna and simplified head model. Note that L = half wavelength at 1880 MHz (79.79 mm) and a = 0.1 mm.



(a) Bent thin dipole (low-Q) (b) Bent fat dipole (high-Q)

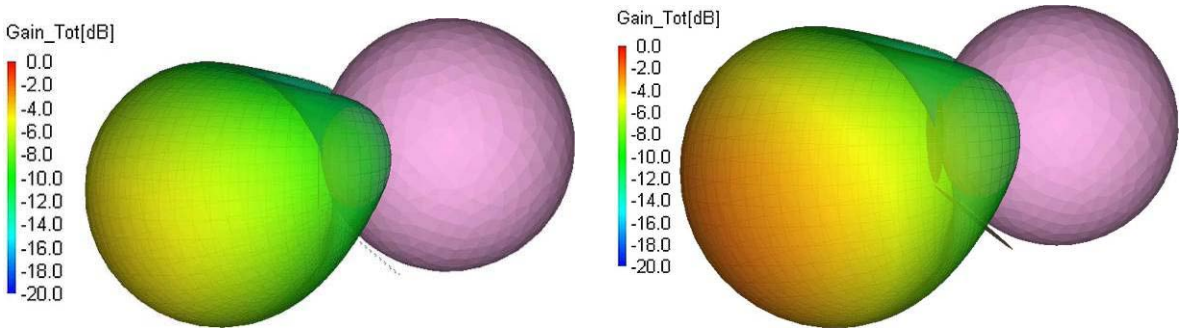
Figure 5.13 Simulation setup for bent thin and fat dipole antennas in the presence of a simple child head model.

The computed VSWR values for the cases with and without the head are compared in Figure 5.14 for each antenna. The compared VSWR results show that the detuning due to the child head model was minimal for both the bent thin dipole (low-Q) and the bent fat dipole (high-Q) antenna (separation between the head and the antenna was d = 5 mm). This lack of detuning might be due to the high loss of the head absorbing the majority of energy generated from the antennas in the direction of the head. This absorption causes the gain to be low. As shown in Figure 5.15, the gain of bent thin and fat dipole antenna was -4.2 dB and -1.67 dB, respectively. The gain of the bent fat dipole (low-Q) had a higher gain than the bent thin dipole (high-Q) by 2.53 dB, suggesting that the near-field intensity of the low-Q antenna is lower than that of high-Q antenna.



(a)Thin dipoles (high-Q) (b) Fat dipole (low-Q)

Figure 5.14 Computed VSWR Comparison between thin and fat dipoles for the cases with and without the simple child head model.



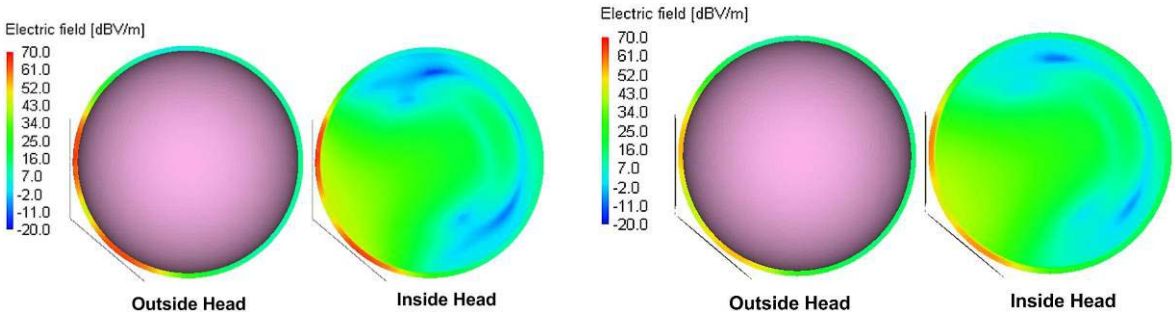
(a) Bent thin dipole
(G_{MAX} = -4.20 dB, e_r = 10.48 %)

(b) Bent fat dipole
(G_{MAX} = -1.67 dB, e_r = 16.96 %)

Figure 5.15 Computed gain pattern and radiation efficiency comparison between bent thin and fat dipole in the presence of simple child head model.

Computed electric and magnetic field distributions for each antenna are compared in Figure 5.16 and Figure 5.17, respectively. The delivered power to each antenna was 21 dBm. The E-field maximum occurs at the end of the wire and the H-field maximum occurs closest to the antenna feed (center). For either the |E| or the |H| field, the field strength had a peak at the surface of the head and rapidly decayed inside the head.

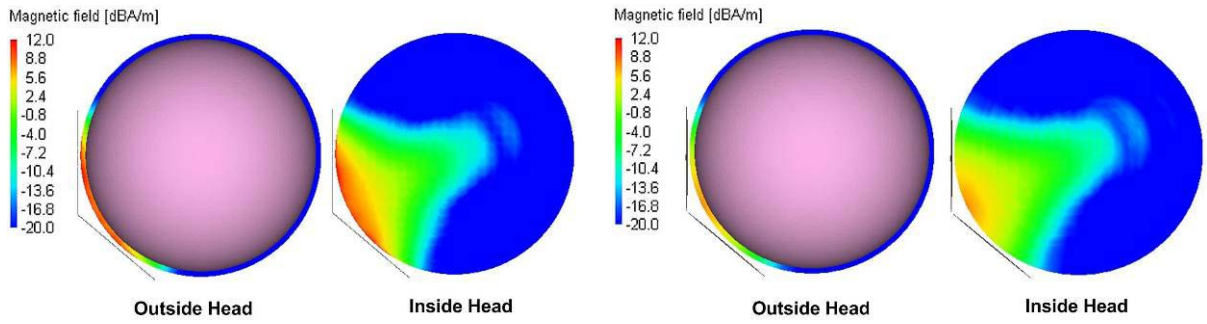
Separation distance (d) versus peak field amplitude is shown in Figure 5.18. At d = 5 mm, the peak E and H-field strength of the bent fat dipole (low-Q) was lower than the bent thin dipole (high-Q) by approximately 14 dB and 9.5 dB respectively. Above d > 60 mm, the field strength of the bent fat dipole was slightly higher due to the higher far-field gain of the antenna.



(a) Bent thin dipole

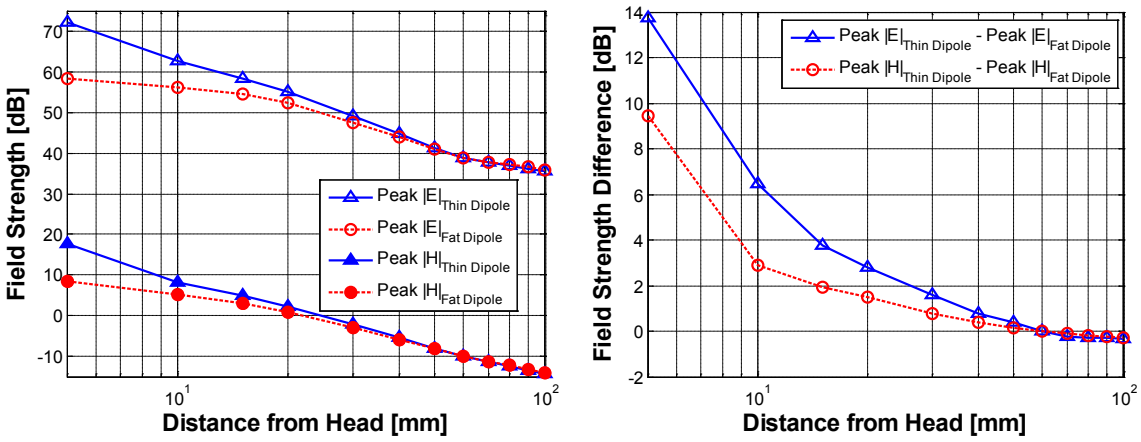
(b) Bent fat dipole

Figure 5.16 Comparison of computed E-field distribution outside and inside simple child head model at d = 5 mm.



(a) Bent thin dipole (high-Q) (b) Bent fat dipole (low-Q)

Figure 5.17 Comparison of computed H-field distribution outside and inside simple child head model at d = 5mm.



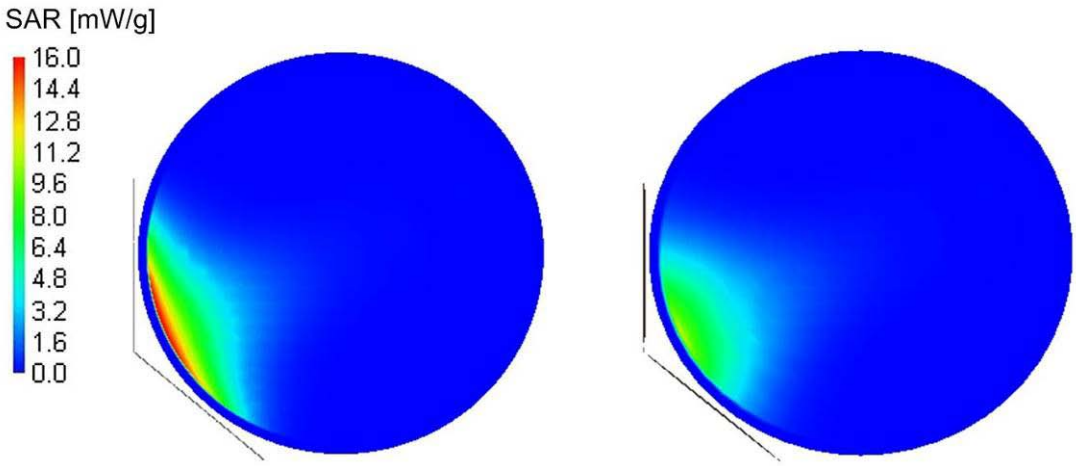
(a) (b)

Figure 5.18 Separation distance from head versus field amplitude for dipole (high-Q) and fat dipole (high-Q) antennas in the presence of simple child model – (a) Comparison of peak near-field strength around simple head model versus distance and (b) Comparison of peak near-field strength difference versus distance.

Keeping the same separation distance between the bent antennas and the simple child head-model (d = 5 mm), the distributions of the specific absorption rate (SAR) were computed and compared in Figure 5.19 (delivered power to each antenna = 21 dBm). We can observe that peak SAR of the fat dipole (low-Q antenna) is lower than the thin dipole (high-Q antenna) case. As shown in Figure 5.20, the bent thin dipole (high-Q) has a 32.4% higher peak SAR at the surface of the head than the case of the fat dipole. For the spatial-average SAR (1g) comparison case, the bent fat dipole antenna (high-Q) was lower than the bent thin dipole by approximately 20%.

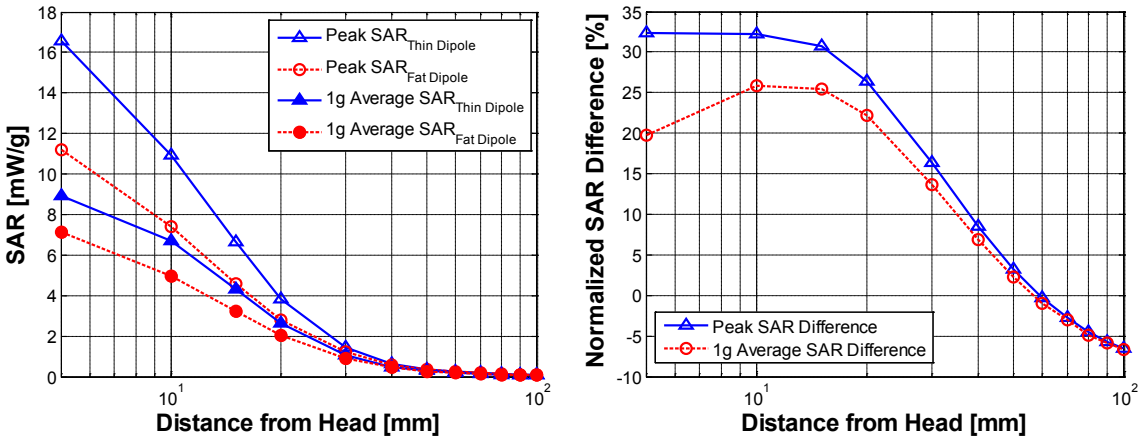
It is interesting to note that both peak and spatial-average SAR (1g) of the bent fat dipole antenna (low-Q) is higher than the cases of thin dipole antenna (high-Q) above 60 mm distance

from the antenna, even though the overall SAR values are very low. This might be because the energy from the bent thin dipole antenna was rapidly absorbed in the head issue at the close distance to the antenna. So, the left energy at the other side of the head for the thin dipole case would be smaller than the fat dipole case.



(a) Bent thin dipole (high-Q) (b) Bent fat dipole (low-Q)

Figure 5.19 Comparison of computed SAR distribution inside simple child head model at d = 5. mm.



(a) (b)

Figure 5.20 Peak and spatial-average SAR comparison between dipole (high-Q) and fat dipole (low-Q) – (a) Comparison of peak and spatial-average SAR (1g) versus distance from the child head model and (b) Comparison of peak and spatial-average SAR (1g) difference. The SAR difference ($SAR_{Thin\ Dipole} - SAR_{Fat\ Dipole}$) is normalized by $SAR_{Thin\ Dipole}$ at each sampled distance.

As summarized in Table 5-4, the current regulated maximum peak spatial-average SAR value is 2 W/Kg or equivalently 2 mW/g. Compared with this value, the computed peak spatial-

average SAR value for the small child head are extremely higher than the regulated value. Lin [*Lin*, 2002] pointed out that, SAR inside heads exposed to RF energy is a function of the size of the head. On the other hand, some early results in literature [*Hombach et al.*, 1996; *Kuster and Balzano*, 1992; *Schönborn et al.*, 1998] claimed that there are no significant differences in the absorption of electromagnetic radiation between adult and child heads. The conclusion of these papers was often based on the computed or measured results below 900 MHz. This could be a critical issue for the mobile communications industry and wireless mobile users because the current methodology of the SAR testing [*ANSI/IEEE*, 2003] only recommends using adult phantoms heads.

In general, the electromagnetic coupling efficiency between the wireless mobile device and human head is affected by many factors including device form factors and types (candybar and cramshell), distance from head, device holding methods, and head size. Thus, some results for certain test conditions may not provide a generalized conclusion because there are many correlated factors and almost random factors that we cannot exclude or include in the study. However, it is often recommended that young children should not use cellular phone in excessive way, regardless the controversial results.

5.2.3 Spherical Mode Decomposition Viewpoint

In previous sections, results of investigation of simple antennas and head models showed that the low-Q antenna has a reduced near-field strength and SAR value without sacrificing far-field radiation performance, compared to the high-Q antenna case. This result may also be explained with the spherical mode decomposition method (see Chapter 3 for details).

When multiple spherical modes are excited (see Figure 5.21), the radian sphere region excited lowest-order spherical mode ($r \le 1/k$) is the most concentrated field region with non-radiating energy, which is often referred as a reactive near-field region. Beyond the radian sphere of the excited highest-order spherical mode, we have the far-field region. Region between the reactive near-field region and the far-field region is a radiating near-field region, where both radiating and non-radiating energies coexist due to the multiple spherical mode excitations.

Typical cellular phone antennas strongly excite a fundamental spherical mode, i.e. spherical TM_{01} mode for a monopole type antenna, in order to cover a wide angular range (omni-directional pattern in azimuth direction). Unfortunately, conventional cell phone antennas often

also excite higher order spherical modes as well as the fundamental mode. Even though the higher order modes are not strongly excited, they can increase near-field strength significantly. The increased near-field intensity is manifested as an increased reactance in antenna input impedance. However, the higher-order modes decay fast in the near-field region and do not contribute to far-field radiation performance due to the weak excitation.

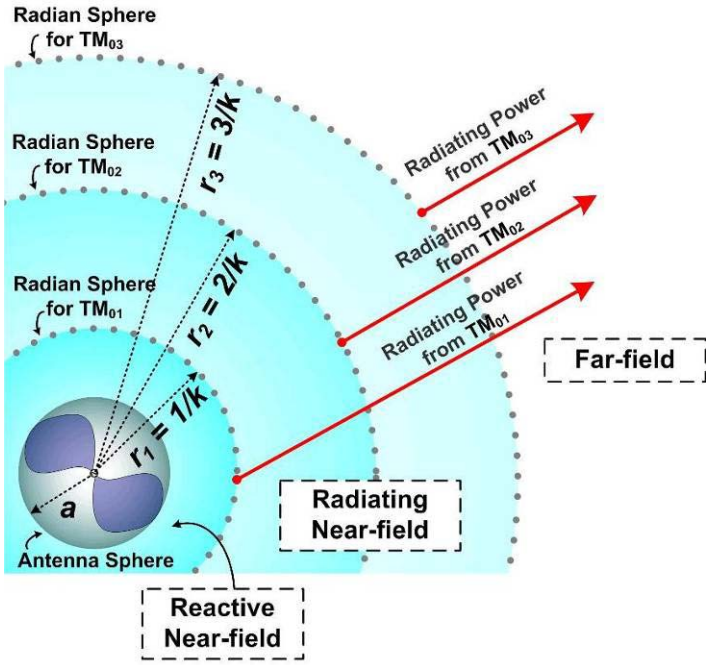


Figure 5.21 Illustration of radian spheres for various spherical TM_{0n} modes and antenna range.

As an example, we consider a simple 900MHz dipole antenna in Figure 5.8, which excites spherical TM_{0n} modes. With $\exp(j\omega t)$ time dependence and Debye gauge condition, we can write the magnetic vector potential in terms of spherical wave functions, which is given by

$$\vec{A} = \hat{r}\sum_{n=1}^{\infty}\sum_{m=-n}^{n} A_{mn}^{TM}\hat{H}_n^{(2)}(kr)P_n^{|m|}(\cos\theta)e^{jm\phi} \tag{5.7}$$

where $\hat{H}_n^{(2)}(kr)$ is the Schelkunoff spherical Hankel function of the second kind, $P_n^{|m|}(\cos\theta)$ is the associated Legendre function of the first kind, and A_{mn}^{TM} is the spherical coupling coefficients of the expansion for TM_{mn} mode. Then, the field for a specific mode can be easily from Maxwell's equations, as we discussed in Chapter 3. Therefore, the field strength is proportional

to the number of the excited spherical modes and the spherical coefficients, i.e. A_{mn}^{TM} for the TM_{mn} mode. The spherical TM mode coefficients can be obtained from two orthogonal fields, i.e.

$$A_{mn} = \frac{\frac{j\omega\varepsilon a}{2\pi k}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\left(\left[E_\theta \frac{d}{d\theta}P_n^{|m|} + \frac{-jm}{\sin\theta}E_\phi P_n^{|m|}\right]e^{-jm\phi}\right)\sin\theta d\theta d\phi}{\hat{H}_n^{(2)\prime}\frac{2}{2n+1}\frac{(n+m)!}{(n-m)!}n(n+1)} \tag{5.8}$$

where a is the radius of the spherical surface that the near fields are sampled. The radiated power associated with each spherical mode is found as

$$P_{rad,mn}^{TM} = \frac{2n(n+1)}{2n+1}\frac{(n+|m|)!}{(n-|m|)!}\left(\eta_0 |A_{mn}|^2\right) \tag{5.9}$$

Therefore, the radiated power of each mode is proportional to $|A_{mn}|^2$.

Decomposed near-field patterns of the 900MHz thin dipole antenna (Figure 5.8a) are shown in Figure 5.22. Very close to the dipole (r = 0.3 λ), it is observed that several higher-order spherical modes are excited. Notice that the near-field pattern of the spherical TM_{01} mode is close to an isotropic pattern shape. At r = 1.5 λ distance, most of higher order modes have decayed. The near-field pattern of the spherical TM_{01} mode is close to donut shaped, which is the well-known far-field pattern of the dipole antenna. In Table 5-6, the spherical coupling coefficients of the decomposed spherical modes are compared for the thin and fat dipoles. The coefficients of the fundamental mode dominating far-field performance are the same for both antennas. However, the fat dipole (low-Q antenna) has lower coefficients for the higher-order modes, leading to significantly stronger near fields than for the thin dipole (high-Q antenna). The difference of the spherical coefficients between the thin and fat dipole may not seem significant. However, the near-field strength is extremely high due to the property of the spherical Hankel function. Therefore, a small difference between the spherical coefficients can create significant increase of near-field strength.

In summary, analysis using the spherical mode decomposition explained why the low-Q antenna has lower near-field strength than that of the high-Q antenna, even though their spherical coupling coefficient of the fundamental mode, dominating the far-field performance, is same.

This result is consistent with the results from other two near-field theories discussed in previous sections.

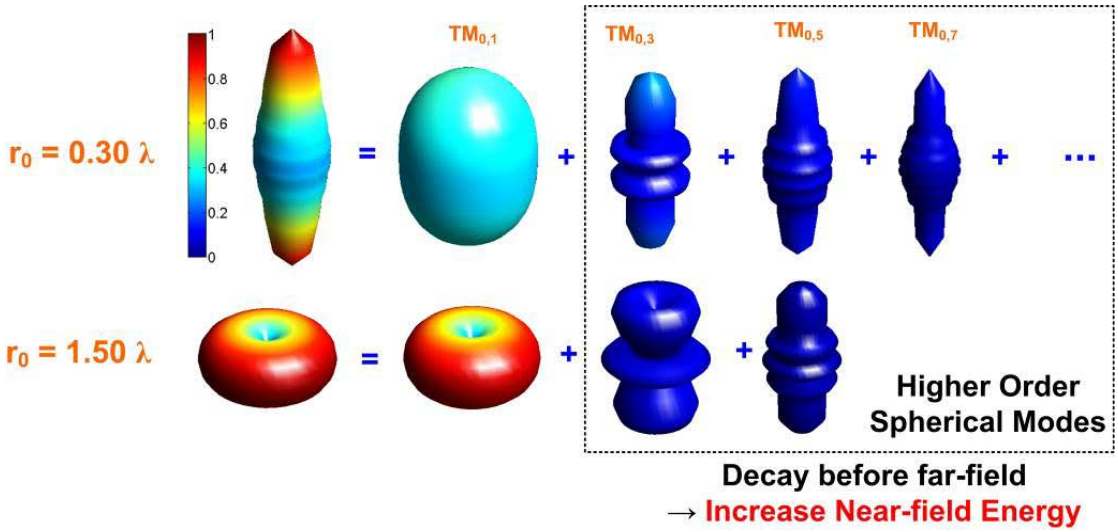


Figure 5.22 Decomposed near-field patterns of 900 MHz half-wavelength dipole (Figure 5.8a) at 0.3 and 1.5 wavelength from the center of the dipole.

Table 5-6 Comparison of the amplitudes of spherical couple coefficient A_{0n}^{TM} for thin dipole (Figure 5.8a) and fat dipole antenna (Figure 5.8b).

n	**Thin Dipole (High-*Q*)**	**Fat Dipole (Low-*Q*)**
	$\lvert A_{0n}^{TM} \rvert$	$\lvert A_{0n}^{TM} \rvert$
1	3.163e-02	3.163e-02
3	1.034e-03	0.892e-03
5	1.757e-05	1.071e-05
7	1.714e-07	0.826e-07

5.3 Near-Field Distributions of Mock Cellular Phone Antennas

5.3.1 Test Antennas

To evaluate realistic systems, the two test antennas were mounted on identical, mock cellular phones for testing. The first test antenna was a dual-band planar inverted-F antenna (PIFA) with the overall dimensions of $40\times10\times10$ mm ($L\times W\times H$). A meandered top plate similar to one

used in [*Ollikainen et al.*, 2000] was used to tune the antenna (see Figure 5.23). The second test antenna, shown in Figure 5.24, was a spoon-shaped ultra-wideband (UWB) antenna with dimensions of $45.52 \times 20 \times 21.1$ mm ($L \times W \times H$), designed by mapping a planar UWB antenna onto a curved surface to form a spoon shape for size reduction. Both test antennas were mounted on the metallic, mock cellular phones as shown in Figure 5.25 with the dimensions of $40 \times 20 \times 95$ mm ($L \times W \times H$) and rounded corners.

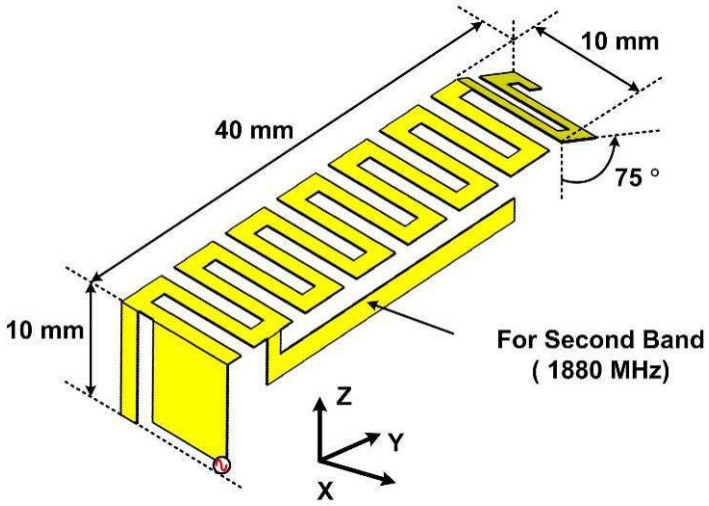
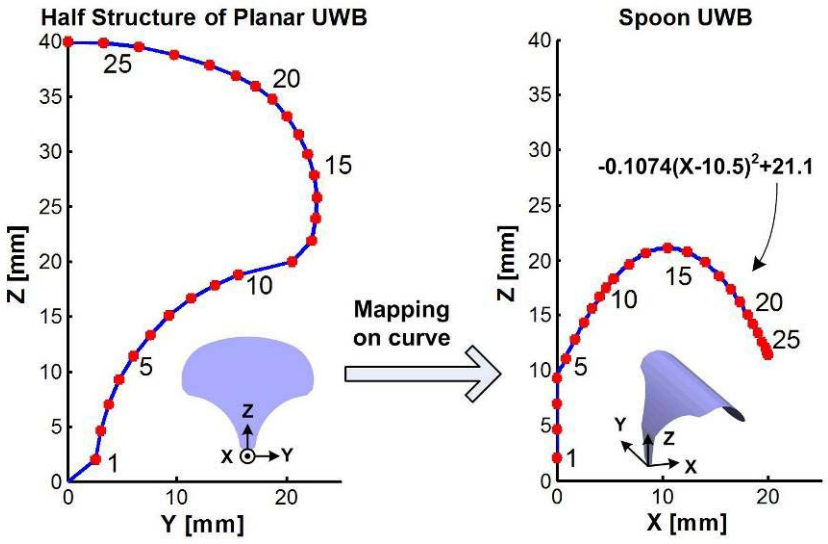
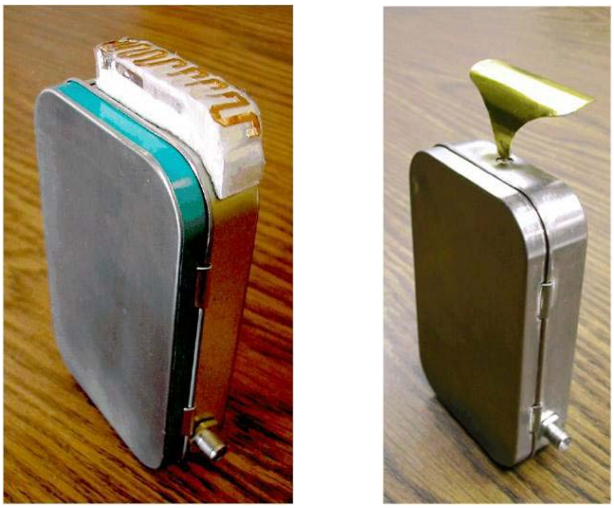


Figure 5.23 Geometry and dimension of designed dual-band PIFA



(a) Planar UWB Antenna (b) Spoon UWB antenna

Figure 5.24 Evolution of the spoon UWB antenna geometry (height of spoon UWB is 21.1 mm) [*Yang et al.*, 2008]. Copyright © 2008 IEEE.



(a) (b)

Figure 5.25 Pictures of designed antennas on a mock cellular phone – (a) Dual-band PIFA and (b) Spoon UWB antenna.

The simulated and measured VSWR values are compared in Figure 5.26. The dual-band PIFA has resonances at 900 and 1880 MHz. The narrow impedance bandwidth about the resonant frequencies suggests a rather high-Q structure. The intended operating frequencies for the spoon UWB antenna are from 900 MHz to 10 GHz (VSWR < 4), providing a low-Q alternative. The dual-band PIFA has a lower VSWR than the spoon UWB at 900 MHz, while the VSWR is similar at 1880 MHz.

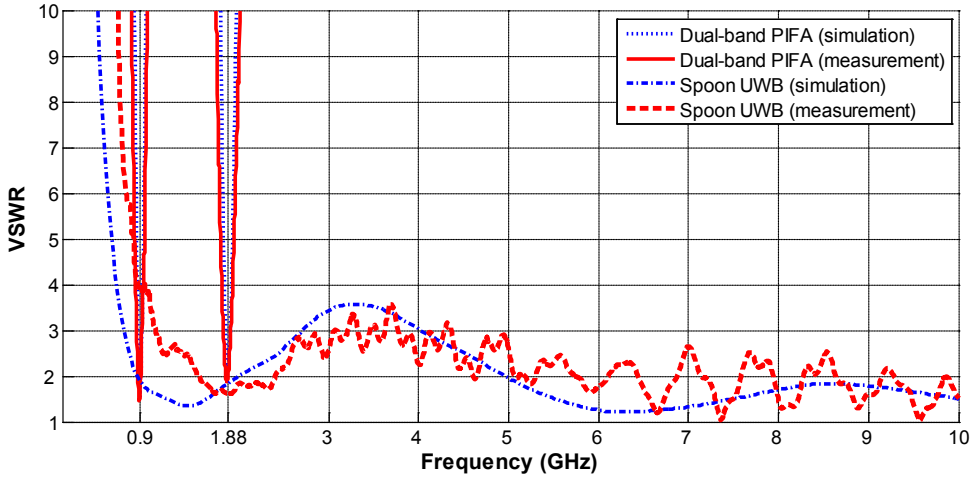
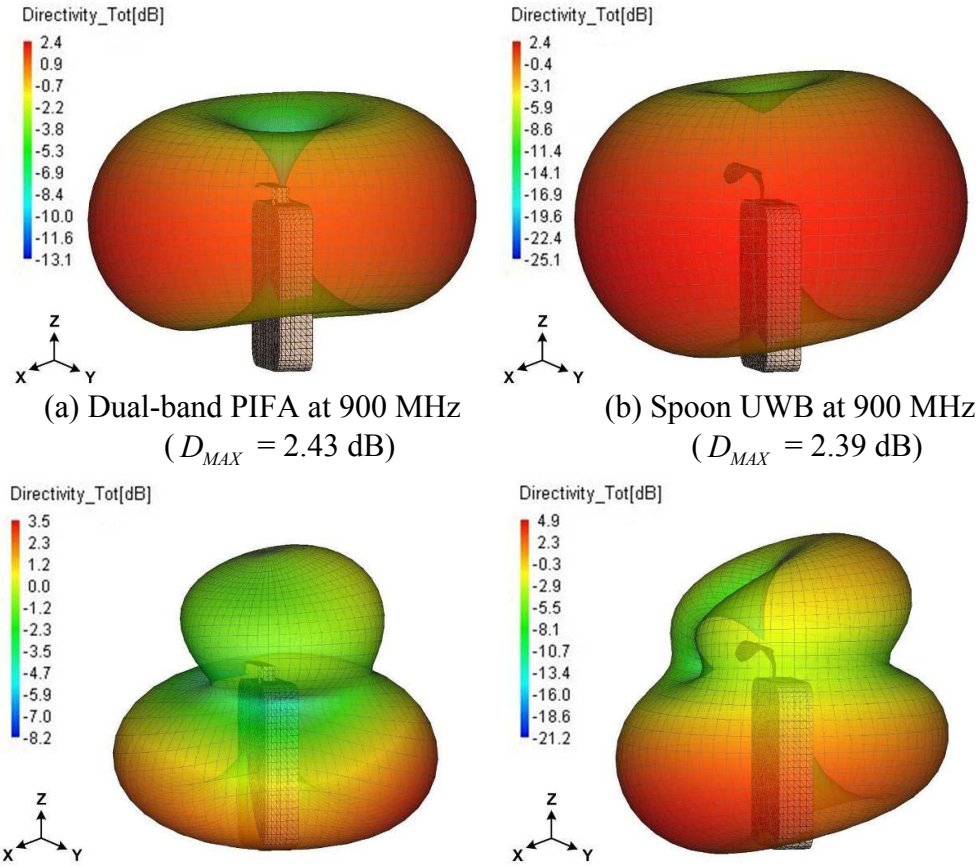


Figure 5.26 Measured and simulated VSWR as a function of frequency for a dual-band PIFA (frequency sweep up to 2.5 GHz) and spoon UWB. Simulations were performed using a commercial method of moments code [*FEKO*, 2006] [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

The simulated radiation patterns of the test antennas are shown in Figure 5.27. At 900 MHz, both test antennas have omni-directional radiation patterns and similar directivity values: D_{MAX} (dual-band PIFA) = 2.43 dB and D_{MAX} (Spoon UWB) = 2.39 dB. However, at 1880 MHz the radiation patterns are bidirectional, resulting from interaction with the cell phone. As seen in [*Chavannes et al.*, 2003], even when the cellular phone with a dual-band antenna is modeled with great detail, the radiation pattern around 1800 MHz is distorted. The whole structure, including the phone body, acts as an unbalanced, pseudo dipole.

The XZ-plane cut of the radiation pattern for each antenna in Figure 5.28 indicates that the far-field patterns are close at 900 MHz and deviate somewhat at 1880 MHz (1.23 dB) for the PIFA and spoon UWB.



(a) Dual-band PIFA at 900 MHz (D_{MAX} = 2.43 dB)

(b) Spoon UWB at 900 MHz (D_{MAX} = 2.39 dB)

(c) Dual-band PIFA at 1880 MHz (D_{MAX} = 3.52 dB)

(d) Spoon UWB at 1880 MHz (D_{MAX} = 4.94 dB)

Figure 5.27 Radiation directivity patterns of the test antennas computed at 900 and 1880 MHz. Simulations were performed using a commercial method of moments code [*FEKO*, 2006] [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

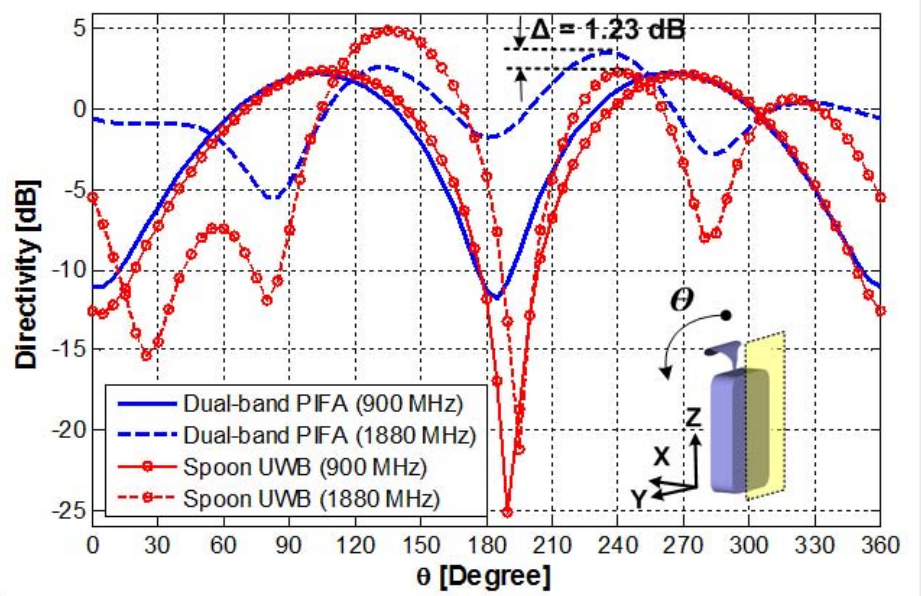


Figure 5.28 Radiation directivity comparison between dual-band PIFA and spoon UWB in the XZ-plane computed at 900 and 1880 MHz. Results are from simulations performed using a commercial method of moments code [*FEKO*, 2006] [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

5.3.2 Comparison between Low and High-*Q* Antennas

An extensive cooperative research program was conducted jointly by Virginia Tech Antenna Group (VTAG) and Sony-Ericsson to investigate HAC [*Yang et al.*, 2008]. For the near-field comparison between low and high-*Q* antennas, VTAG performed the simulation using a commercial method of moments code [*FEKO*, 2006] and Sony-Ericsson made measurements on near-field scan using DASY3 system [*SPEAG*, 2006].

The mock test phones did not have a speaker, so the scan area with 9 grids, as in the HAC standard [*ANSI*, 2006], cannot be defined. Therefore, the near-field strength was scanned over the entire structure at the distance X_{dist}, as shown in Figure 5.29.

In the simulation of the near field scanning of the test antennas, the delivered power was set to 32 dBm at 900 MHz and 30 dBm at 1880 MHz in order to remove the return loss issues in the comparisons. For the near-field measurements, the DASY3 system [*SPEAG*, 2006] was used with the available power at each antenna terminal set to 32 dBm at 900 MHz and 30 dBm at 1880 MHz. A picture of measurement setup is shown in Figure 5.30.

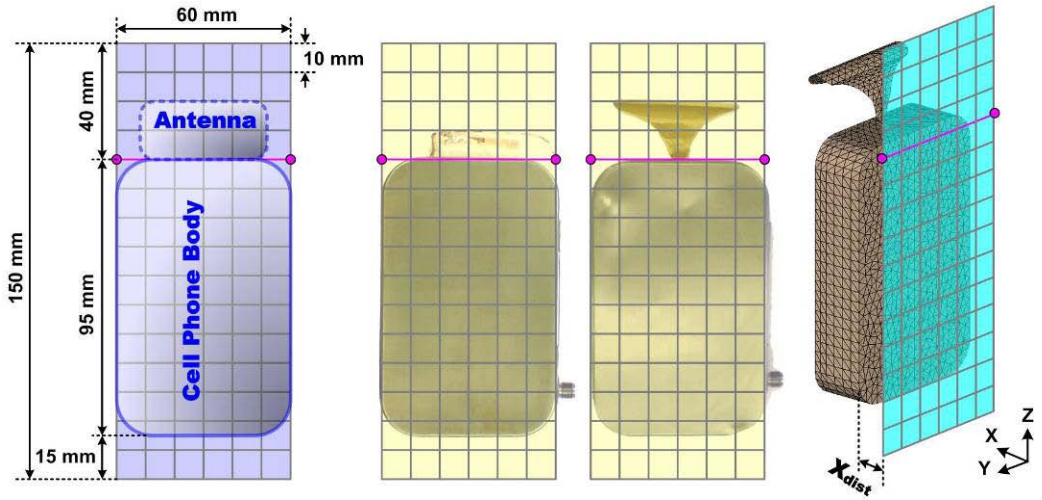
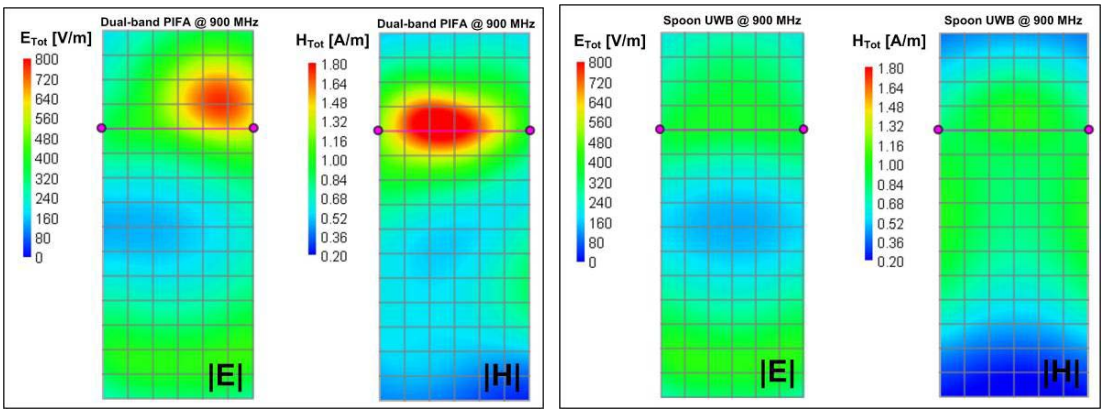


Figure 5.29 Near-field measurement setup [*Yang et al.*, 2008]. (a) Mock cellular phone dimension, (b) Dual-band PIFA mock cellular pone, (c) Spoon UWB mock cellular phone, and (d) Definition of X_{dist}. Copyright © 2008 IEEE.

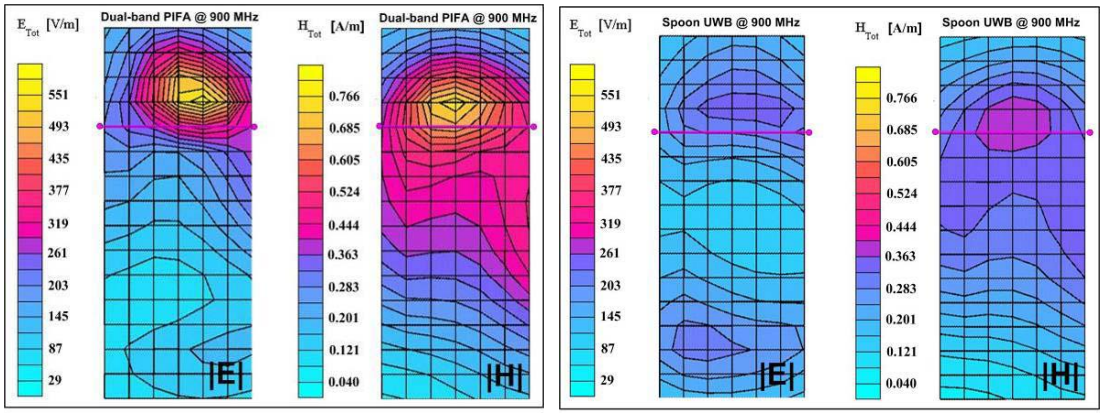
Figure 5.30 Picture of near-field measurement setup using DASY3^{TM} system [*SPEAG*, 2006]. Near-field measurements were performed by Dr. Minh-Chau Huynh, Sony Ericsson.

The simulation and measured near-field scan results are compared in Figure 5.31 through Figure 5.34. In all cases, the spoon UWB antenna shows lower peak near-field strength. The simulated and measured peak field values differ because loss was not considered in the simulation. However, the overall field distributions obtained from the simulations and the measurements are very similar and demonstrate the same improvement using a UWB antenna structure.



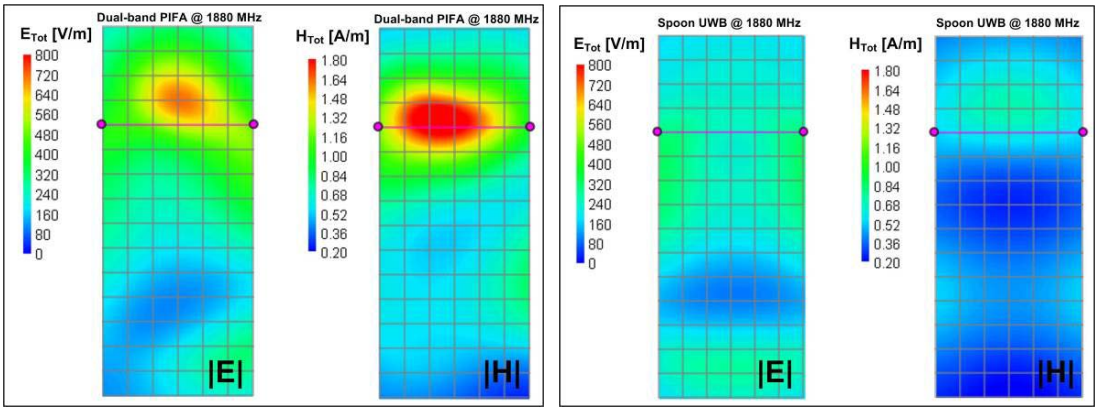
(a) Dual-band PIFA (b) Spoon UWB

Figure 5.31 Simulated near-field distribution at 900 MHz [*Yang et al.*, 2008]. Copyright © 2008 IEEE.



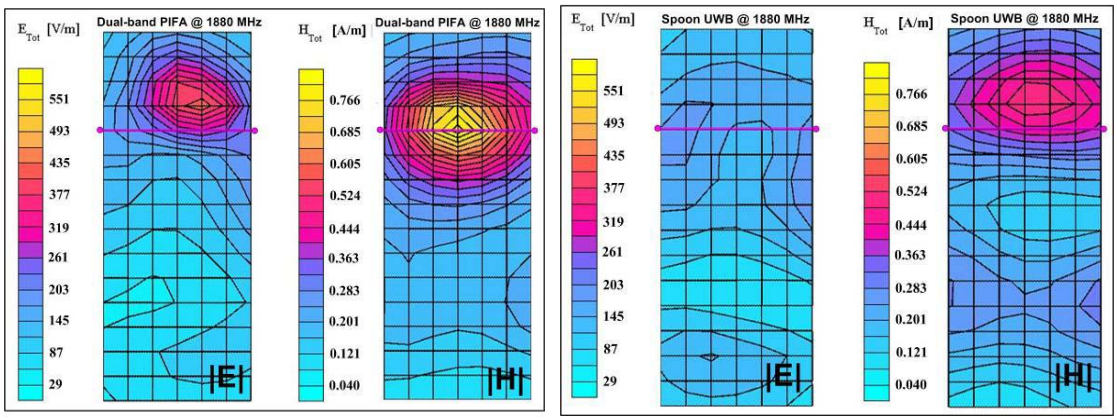
(a) Dual-band PIFA (b) Spoon UWB

Figure 5.32 Measured near-field distribution at 900 MHz [*Yang et al.*, 2008]. Copyright © 2008 IEEE.



(a) Dual-band PIFA (b) Spoon UWB

Figure 5.33 Simulated near-field distribution at 1880 MHz [*Yang et al.*, 2008]. Copyright © 2008 IEEE.



(a) Dual-band PIFA (b) Spoon UWB

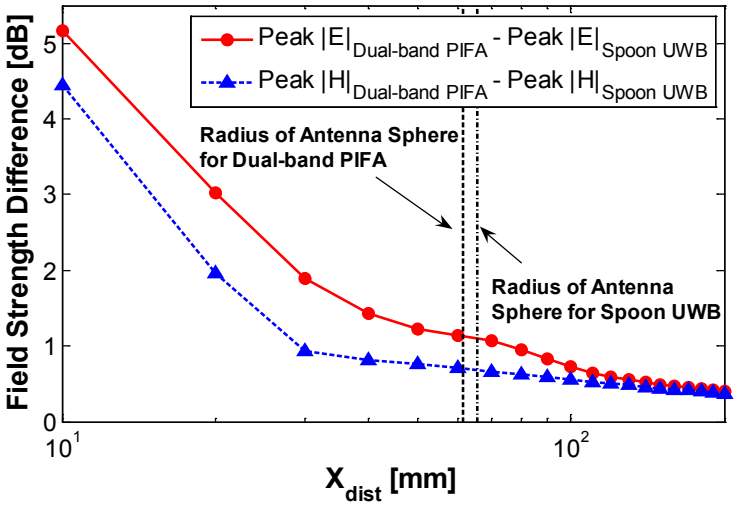
Figure 5.34 Measured near-field distribution at 1880 MHz [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

The differences in the peak field strength are summarized in Table 5-7. At 900 MHz, the peak difference values of the measurements are overestimated because the mismatch loss in the measurement environment was included. The higher VSWR for the measured values at 900 MHz corresponds to about 1.5 dB worse return loss; 1.5 dB corrections is included in Table 5-7 to account for this return loss difference. In addition, as explained earlier, the maximum directivity of the spoon UWB antenna in the measurement plane is lower by 1.23 dB at 1880 MHz than the dual-band PIFA. As a result, the difference of the peak field value at 1880 MHz exceeds that at 900 MHz. In this particular experiment, both simulations and measurements suggest that the peak $|\vec{E}|$ and $|\vec{H}|$ of the spoon UWB antenna (low-Q antenna) is approximately 5 dB and 4 dB, respectively, below that of the dual-band PIFA (high-Q antenna).

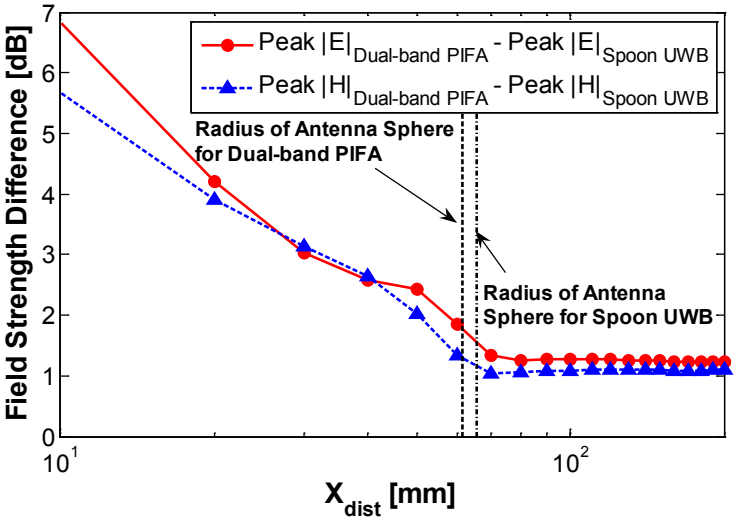
Table 5-7 Comparison of the difference in peak field strengths for dual-band PIFA and spoon UWB antennas (ΔPeak |field| = Peak $|\text{field}|_{\text{Dual-band PIFA}}$-Peak $|\text{field}|_{\text{Spoon UWB}}$) [*Yang et al.*, 2008]

	Simulated		**Measured**	
	900 MHz (P_{del} = 32 dBm)	**1880 MHz (P_{del} = 30 dBm)**	**900 MHz (P_{avail} = 32 dBm)**	**1880 MHz (P_{avail} = 30 dBm)**
ΔPeak $\|\vec{E}\|$ (V/m)	5.2 dB	6.8 dB	7.1 dB – 1.5 dB = 5.6 dB	6.5 dB
ΔPeak $\|\vec{H}\|$ (A/m)	4.4 dB	5.6 dB	5.6 dB – 1.5 dB = 4.1 dB	4.0 dB

The peak field-strength difference versus distance from the antenna (X_{dist}) for both test antennas are plotted in Figure 5.35. As previously explained, the overall structure of the test antennas on the mock phones act as asymmetric, pseudo dipoles. The effective radii of the antenna spheres for the spoon UWB and the dual-band PIFA are 65.6 mm and 61.3 mm, respectively. The effect of using a low-Q antenna is to reduce the non-radiating energy within the antenna sphere. The field strength difference at 1880 MHz does not approach zero as expected due to a 1.23 dB difference in the directive gain off the side of the phone.



(a) 900 MHz



(b) 1880 MHz

Figure 5.35 Computed peak field strength difference comparison between dual-band PIFA and spoon UWB at 900 and 1880 MHz [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

5.4 Radiation Performance in The Presence of The Human Head

5.4.1 Visible Human Head Case

When a cellular phone is used in close proximity to a human head, dielectric loading effects can be expected. There may also be a detuning issue for narrowband antennas (high-Q antennas). Chen [*Chen et al.*, 2006] reported that the impedance performance of an UWB antenna is only slightly affected by the presence of a human head. During normal cellular phone use, the human head is inside the antenna sphere and is exposed to strong near-fields. The strong near-field energy increases the interaction with the human head and can alter the input impedance. It is reasonable to expect that UWB antennas would not experience as much detuning as found with narrowband antennas due to the reduced near-field energy and associated interaction with adjacent materials. Similarly, the efficiency of the low-Q antenna should be higher than that of a high-Q antenna when including human-head effects.

In order to demonstrate improved performance of the low-Q antenna in the presence of a human head, simulations were performed using the test antennas on mock phones in the presence of the NIH head model, called the "Visible Human Head" [*NIH*, 1986]. A commercial moment-method/finite-element-method hybrid code [*FEKO*, 2006; *Meyer et al.*, 2003] was used for the simulations. The male human head is solved with the finite-element method (tetrahedral meshes, 4 mm resolution) using the head CAD model available from EM Software & Systems [*FEKO*, 2006]. As shown in Figure 5.36, the CAD model of the human male head is composed of 24 biological tissues and the electrical properties of the biological tissues can be found at the web site of Italian national research council [*IFAC*]. The specific properties of the biological tissues at 900 and 1880 MHz are summarized in Table 5-8.

The test antennas on the mock phone were modeled with the moment method [*FEKO*, 2006] neglecting ohmic loss. The test antennas were located 10 mm from the male human head as specified in the HAC Standard. The simulated radiation gain patterns at 900 and 1880 MHz, shown in Figure 5.37, are directional due to head interaction. Antenna efficiency is reduced by power absorption in the human head and is summarized in Table 5-9. In this particular configuration of the cellular phone and human head, the spoon UWB antenna shows an 8 and 20 percent improvement in efficiency at 900 and 1800 MHz, respectively.

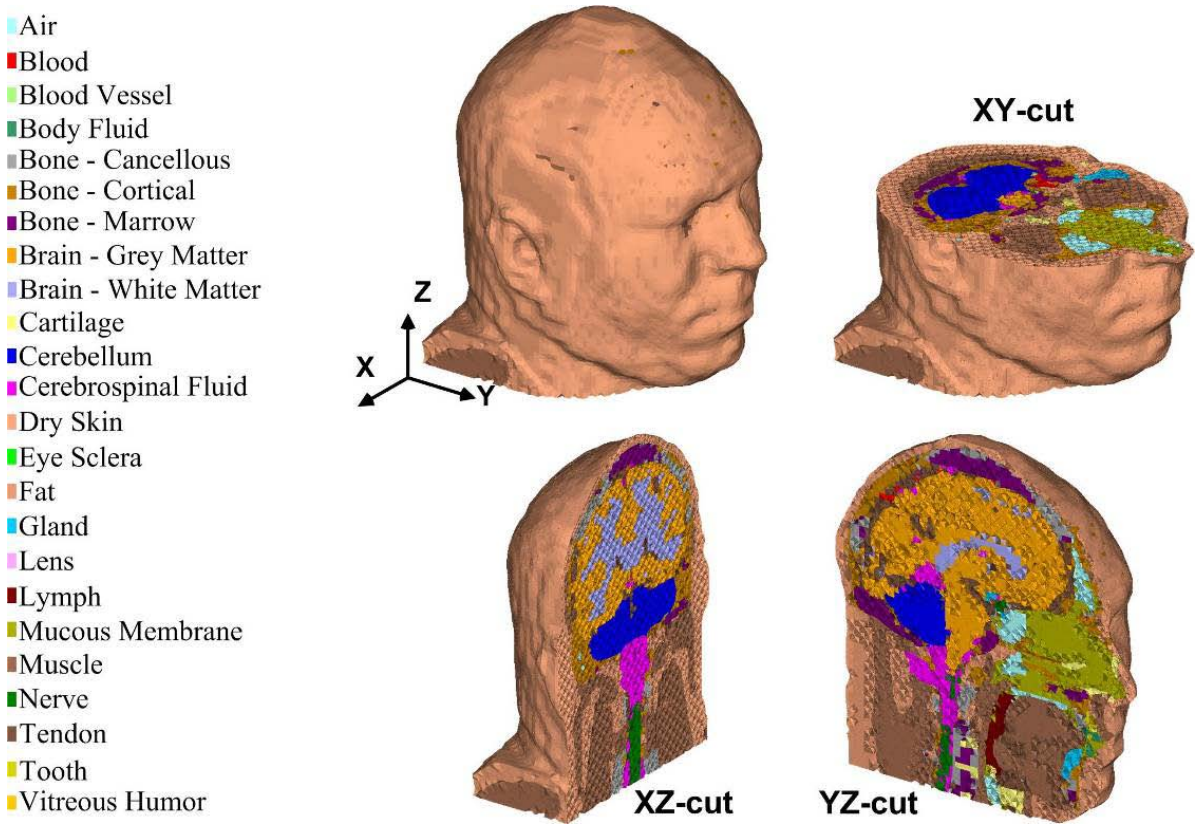
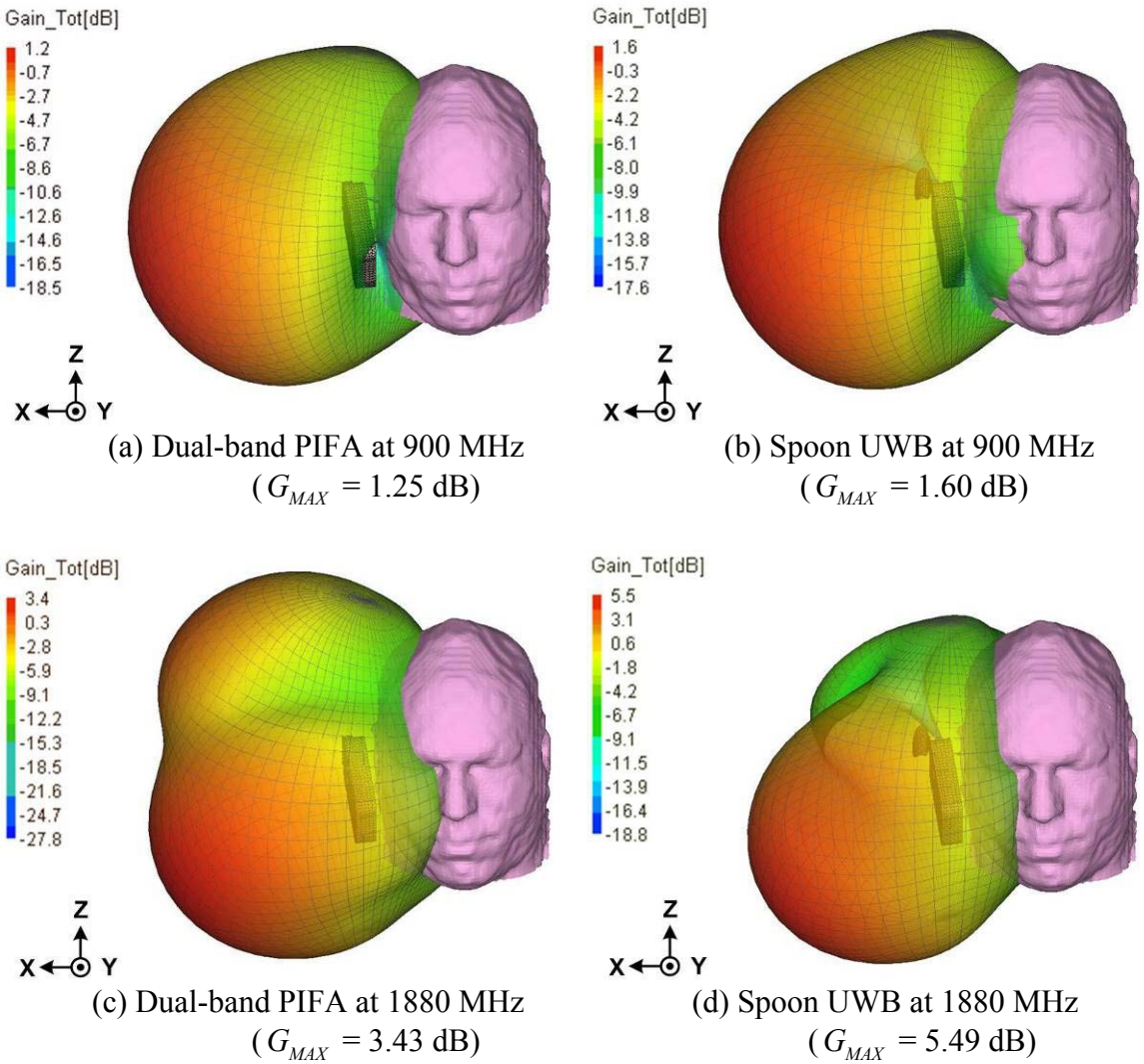


Figure 5.36 Visible human male head model used for simulations. The computer-aided design (CAD) model of Visible human head was provided as a part of a commercial method of moments package [*FEKO*, 2006; *Yang et al.*, 2008]. Copyright © 2008 IEEE.

Table 5-8 Summary of dielectric properties used in Visible human head model [*IFAC*]

Dielectrics		Relative Permittivity (ε_r')		Conductivity (σ_e) [S/m]		Mass Density (ρ) [Kg/ m^3]
		900 MHz	1880 MHz	900 MHz	1880 MHz	
Air		1	1	1	1	1.3
Blood		61.36	59.372	1.5379	2.0435	1058
Blood Vessel		44.775	43.241	0.69612	1.1068	1040
Body Fluid		68.902	68.534	1.6362	2.0802	1010
Bone	Cancellous	20.788	19.239	0.34	0.61351	1920
	Cortical	12.454	11.729	0.14331	0.28895	1990
	Marrow	5.5043	5.362	0.040208	0.071489	1040
Brain	Grey Matter	52.725	49.92	0.94227	1.4383	1038
	White Matter	38.886	36.896	0.59079	0.94887	1038
Cartilage		42.653	40.031	0.78239	1.3404	1097
Cerebellum		49.444	45.93	1.2628	1.7538	1038
Cerebrospinal Fluid		68.638	67.085	2.4126	2.9823	1007
Dry Skin		41.405	38.745	0.86674	1.2165	1125
Eye Sclera		55.271	53.448	1.1668	1.6498	1026
Fat		5.462	5.3406	0.051043	0.081343	916
Gland		59.684	58.025	1.0385	1.5527	1050
Lens		46.573	45.261	0.7934	1.1869	1053
Lymph		59.684	58.025	1.0385	1.5527	1040
Mucous Membrane		46.08	43.715	0.84465	1.2727	1040
Muscle		55.032	53.444	0.94294	1.3851	1047
Nerve		32.531	30.768	0.57369	0.87077	1038
Tendon		45.825	44.114	0.71839	1.2548	1220
Tooth		12.454	11.729	0.14331	0.28895	2160
Vitreous Humor		68.902	68.534	1.6362	2.0802	1009



(a) Dual-band PIFA at 900 MHz (G_{MAX} = 1.25 dB)

(b) Spoon UWB at 900 MHz (G_{MAX} = 1.60 dB)

(c) Dual-band PIFA at 1880 MHz (G_{MAX} = 3.43 dB)

(d) Spoon UWB at 1880 MHz (G_{MAX} = 5.49 dB)

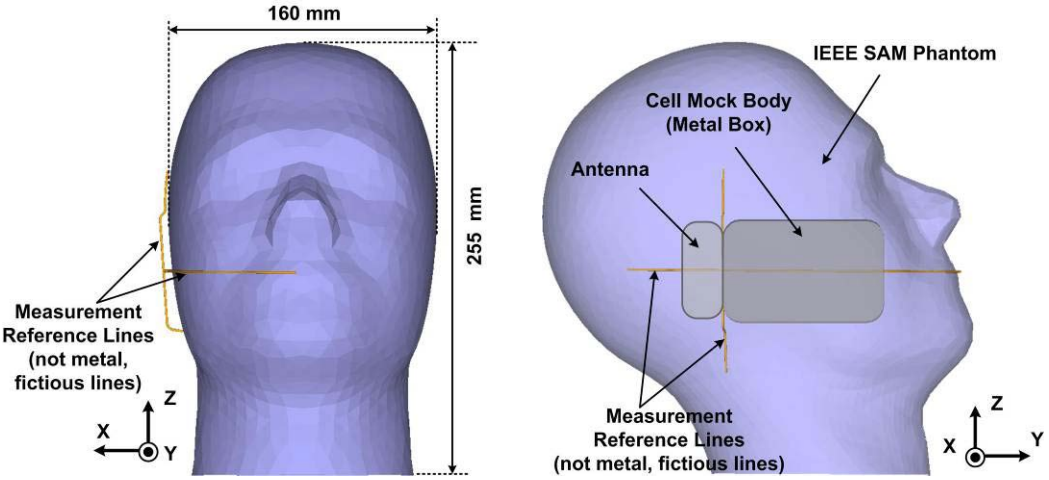
Figure 5.37 Computed radiation gain patterns in the presence of visible human male head at 900 and 1880 MHz [*Yang et al.*, 2008]. Copyright © 2008 IEEE.

Table 5-9 Comparison of radiation performance for dual-band PIFA and spoon UWB antennas at 900 and 1880 MHz obtained by simulations with a human head model [*Yang et al.*, 2008]

	900 MHz		**1880 MHz**	
	Dual-band PIFA	**Spoon UWB**	**Dual-band PIFA**	**Spoon UWB**
Directivity	5.72 dB	5.21 dB	6.97 dB	7.42 dB
Gain	1.25 dB	1.60 dB	3.43 dB	5.49 dB
Efficiency	36 %	44 %	44 %	64 %

5.4.2 SAM Head Case

The radiation performance of both dual-band PIFA (high-Q) and spoon UWB antenna (low-Q) on the mock cellular phone is also investigated through simulations for the case of SAM head model. The overall dimensions and mock cellular phone alignment with the SAM head model are depicted in Figure 5.38. Compared to the commercial SAM phantoms for measurements, our CAD model for simulations does not include the outer shell and detailed ear structures. Since the outer shell is supposed to be made of low dielectric material, simulation results with and without the outer shell would be similar. Dielectric properties of the SAM head model are summarized in Table 5-10. Only the cheek mode (see Section 5.1.2.2) is considered in this study.



(a) (b)

Figure 5.38 Overall dimension of SAM head CAD model and alignment configuration with mock cellular phone (a) SAM head phantom model and (b) mock cellular phone alignment on the SAM head model.

Table 5-10 Dielectric properties of SAM head model for simulations based on [*ANSI/IEEE*, 2003]

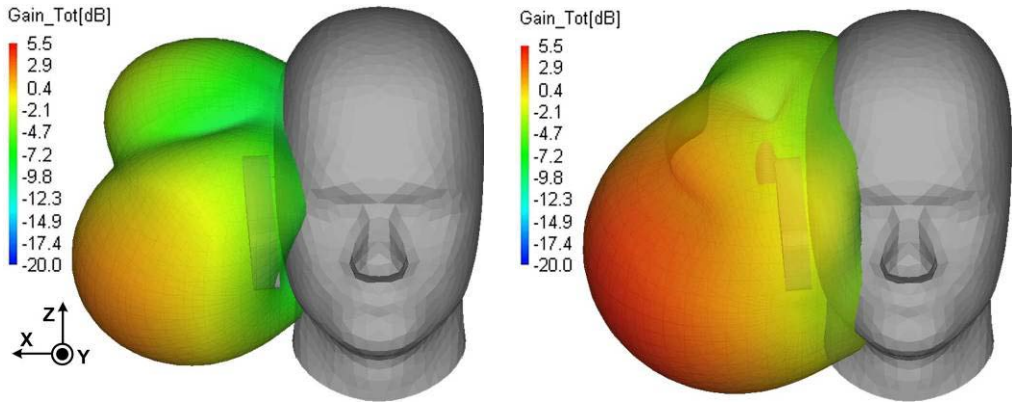
Frequency	**Relative Permittivity (ε_r')**	**Conductivity (σ_e) [S/m]**	**Mass Density (ρ) [Kg/ m^3]**
900 MHz	41.5	0.97	1030
1880 MHz	40.0	1.40	1030

Radiation performance simulations for both antennas were carried out at 900 and 1880 MHz. At 900 MHz, the radiation patterns of both antennas were similar to the case of visible. The

computed maximum gain for the dual-band PIFA and the spoon UWB were 2.0 dB and 2.1 dB, respectively. So, similar to the visible human case, there was no noticeable radiation performance difference between the test cases in the present of SAM head model at 900 MHz.

At 1880 MHz, both test antennas had approximately 7.4 dB directivity, but the computed maximum gain was 2.89 dB for the dual-band PIFA and 5.49 dB for the spoon UWB antenna. Thus, the spoon UWB antenna showed more than 25 % increase in radiation efficiency, compare to the dual-band PIFA antenna in the present of SAM head model.

As summarized in Table 5-11, the radiation efficiency of the low-Q antenna (spoon UWB) clearly outperformed that of the high-Q antenna case (dual-band PIFA) at 1880 MHz for both the visible human head and SAM head cases.



(a) Dual-band PIFA
(G_{MAX} = 2.89 dB, e_r = 35.55 %)

(b) Spoon UWB
(G_{MAX} = 5.49 dB, e_r = 63.82 %)

Figure 5.39 Computed radiation gain patterns in the presence of SAM head model at 1880 MHz.

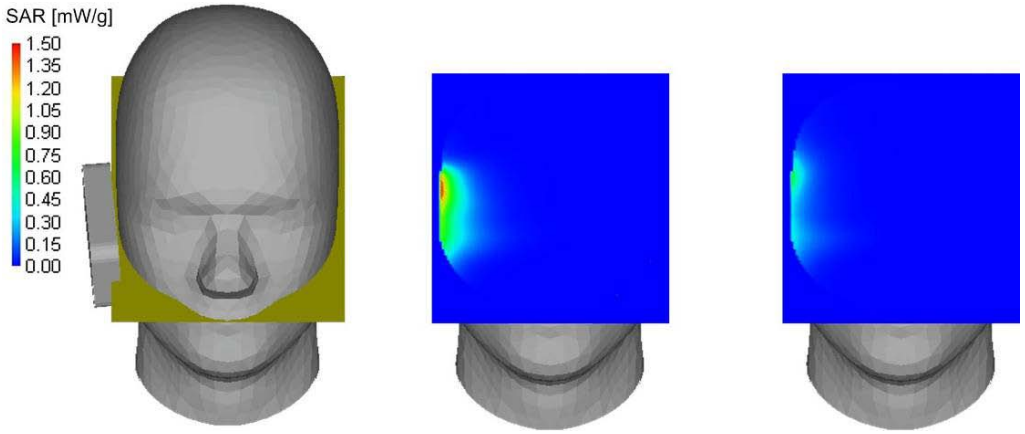
Table 5-11 Summary of radiation performance for dual-band PIFA and spoon UWB antennas at 1880 MHz obtained by simulations with Visible human head and SAM head model.

	Visible Human Head		**SAM Head**	
	Dual-band PIFA (High-Q)	**Spoon UWB (Low-Q)**	**Dual-band PIFA (High-Q)**	**Spoon UWB (Low-Q)**
Directivity	6.97 dB	7.42 dB	7.38 dB	7.44 dB
Gain	3.43 dB	5.49 dB	2.89 dB	5.49 dB
Efficiency	44 %	64 %	36 %	64 %

5.5 Specific Absorption Rate (SAR) Performance of Mock Cellular Phone Antennas in the Presence of SAM head

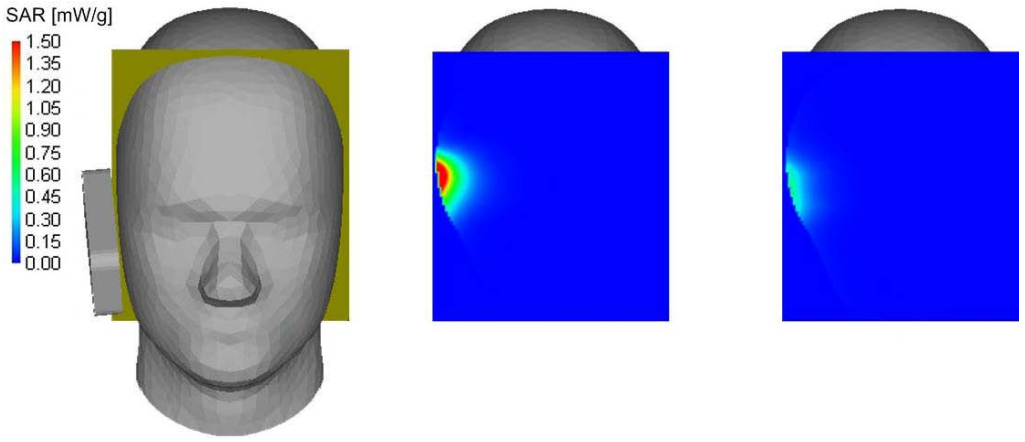
SAR performance was investigated in the presence of the SAM head using the same dual-band PIFA (high-*Q*) and spoon UWB antenna (low-*Q*) on the mock cellular phone.

The computed SAR distributions at 1880 MHz at three principal planes are presented in Figure 5.40 through Figure 5.42. The delivered power was set to 21 dBm in all cases. As predicted from the near-field theories, the spoon UWB (low-*Q*) showed a lower peak SAR than that of the dual-band PIFA (high-*Q*) in all planes. The dual-band PIFA show a strong peak SAR (strong hot spot) on the surface of SAM head model closed to the feed structure while the spoon UWB antenna has a smooth SAR distribution without a strong hot spot.



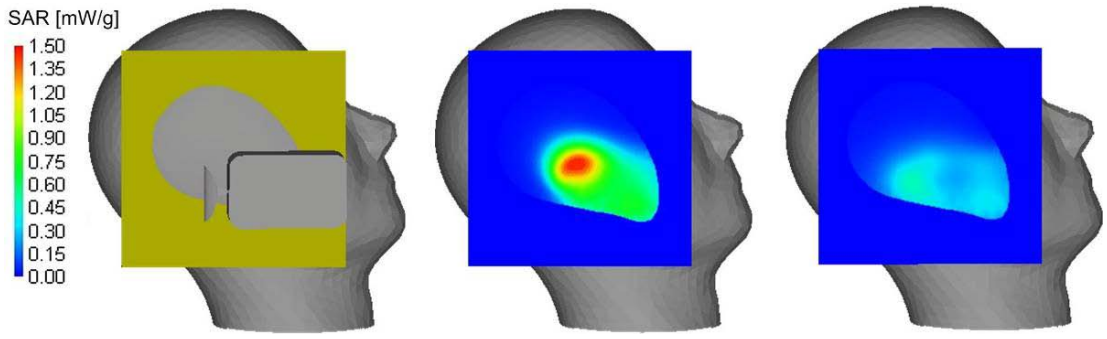
(a) XY-cut (b) Dual-PIFA (c) Spoon UWB

Figure 5.40 XY-cut definition and computed SAR distribution comparison at 1880 MHz between dual-band PIFA (SAR_{Peak} = 1.33 mW/g) and Spoon UWB (SAR_{Peak} = 0.73 mW/g).



(a) XZ-cut (b) Dual-PIFA (c) Spoon UWB

Figure 5.41 XZ-cut definition and computed SAR distribution comparison at 1880 MHz between dual-band PIFA (SAR_{Peak} = 2.96 mW/g) and Spoon UWB (SAR_{Peak} = 0.69 mW/g).



(a) YZ-cut (b) Dual-PIFA (c) Spoon UWB

Figure 5.42 YZ-cut definition and computed SAR distribution comparison at 1880 MHz between dual-band PIFA (SAR_{Peak} = 1.48 mW/g) and Spoon UWB (SAR_{Peak} = 0.49 mW/g).

SAR measurement was carried out by using DASY4^{TM} system [*SPEAG*, 2007]. A picture of the test setup is depicted in Figure 5.43. Typical commercial SAM phantoms are made out of a low-dielectric shell giving a human body shape container. The DASY system has a vertically-dissected half SAM phantom located under a test table. The dissected side of the SAM phantom is open and facing toward the top of the table. A body equivalent fluid is filled inside the half SAM phantom. During the measurement, an automated robot arm scans the entire volume inside the SAM phantom with a small field probe. As explained in Section 5.1.2.2, spatial-average SAR needs to be measured in both cheek and tilt modes. The worst spatial-average SAR value needs to meet the SAR compliance specification (1.6 W/Kg in United States or 2.0 W/Kg in other countries).

The measured SAR distributions at 900 and 1880 MHz are shown in Figure 5.44 and Figure 5.45, respectively. Similar to the simulation results, one can observe that the dual-band PIFA on the mock cellular phone show strong hot spots at both 900 and 1880 MHz, compared to the case of the spoon UWB antenna case with smooth SAR distribution on the surface of the SAM head.

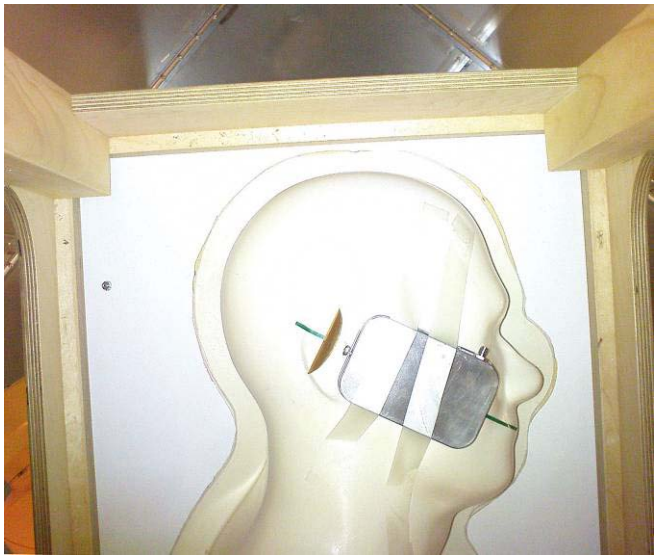
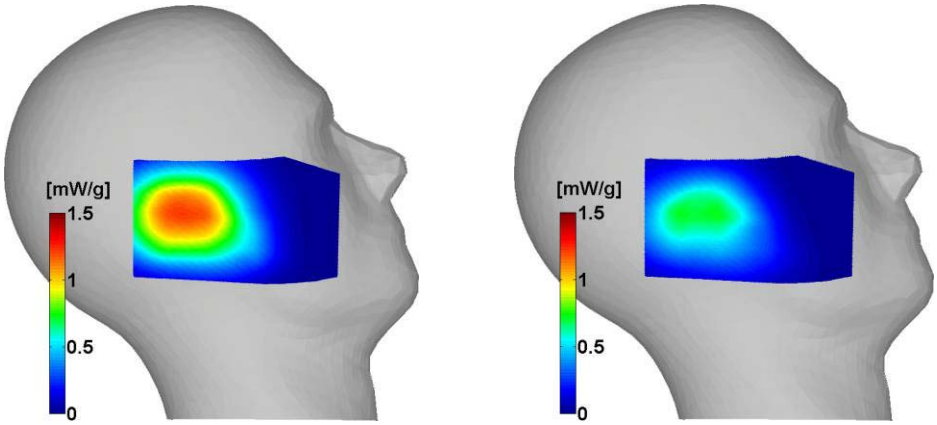
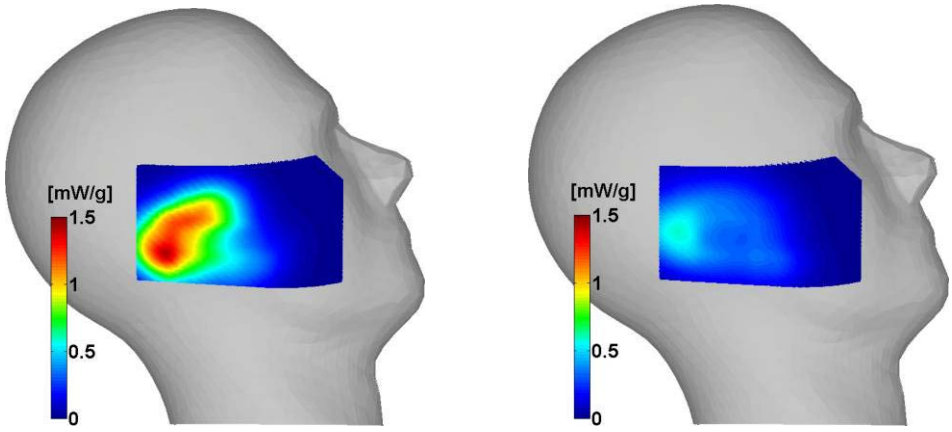


Figure 5.43 Picture of SAR measurement with DASY4™ system [*SPEAG*, 2007] with spoon UWB antenna in the cheek mode. Near-field measurements were performed by Dr. Minh-Chau Huynh, Sony Ericsson.



(a) Dual PIFA @ 900 MHz (b) Spoon UWB @ 900 MHz

Figure 5.44 Comparison of measured SAR distributions between dual-band PIFA (high-*Q*) and Spoon UWB antenna (low-*Q*) on the surface of SAM phantom at 900 Mhz (delivered power = 23 dBm) .



(a) Dual-PIFA @ 1880 MHz (b) Spoon UWB @ 1880 MHz

Figure 5.45 Comparison of measured SAR distributions between dual-band PIFA (high-*Q*) and Spoon UWB antenna (low-*Q*) on the surface of SAM phantom at 1880 MHz (delivered power = 21 dBm).

The simulated and measured SAR test results at 1880 MHz are compared in Table 5-12. In both simulation and measurement cases, the low-*Q* antenna (spoon UWB) has more than 50% lower SAR values than the case of high-*Q* antenna in both simulation and measurement results.

Table 5-12 Comparison between simulated and measured SAR results at 1880 MHz. (Delivered power to the test antennas was 21 dBm)

	Simulated SAR (mW/g)		**Measured SAR (mW/g)**	
	Peak Spatial-Average (1 g)	**Peak**	**Peak Spatial-Average (1 g)**	**Peak**
Dual-PIFA	1.78	2.96	1.26	1.42
Spoon UWB	0.45	0.73	0.54	0.59
SAR Reduction	75 %	75 %	57 %	59 %

5.6 Chapter Summary

A method to reduce near-field electromagnetic energy around a cellular phone was presented to mitigate the interaction between a cellular phone and a hearing aid. Examining near-field terms in field equations showed that current slope discontinuities cause high near field intensities. The

fundamental limit theory of antennas indicates that the near-field interaction problems can be reduced by using low-Q antennas, such as an ultra-wideband (UWB) antenna.

The low-Q antenna approach to reduce near-field strength was demonstrated through both simulations and measurements. Simulations showed that the peak electric and magnetic near-field amplitudes are 7 dB and 4 dB, respectively, lower for the planar UWB fat dipole (low Q) relative to the narrow-band dipole (high Q).

The spherical wave decomposition results showed that both planar UWB fat dipole and narrow-band dipole excite the fundamental spherical mode (TM_{01}) with the same spherical coupling coefficient, which indicates that both antennas will have the same far-field radiation performance. However, the planar UWB fat dipole does not excite higher-order spherical modes strongly, compared to the case of the narrow-band dipole, indicating the generated near-field strength of the UWB fat dipole will be weaker than that of the narrow thin dipole for the same radiated power. Therefore, the spherical mode decomposition results were successfully used to explain why a low-Q antenna has lower near-field strength than a high-Q antenna without sacrificing a far-field performance.

For more practical cellular-phone antennas, the simulation and measurement results for the dual-band, planar inverted-F (900 and 1880 MHz, high Q) and ultra-wideband (900 MHz – 10 GHz, low-Q) test antennas mounted on mock phones showed that the peak electric and magnetic near-field strength is reduced by at least 5 dB and 4 dB, respectively, for the UWB test antenna. Near-field strength reduction was achieved without sacrificing far-field performance.

In the presence of a human head, simulation results also showed that the radiation efficiency of the ultra-wideband test antenna is better than the dual-band PIFA antenna, using both the Visible human and SAM head models. Furthermore, the simulation and measurement results of specific absorption rate (SAR) with a SAM phantom head model revealed that the spoon UWB (low-Q) antenna showed more than a 50% reduction of both peak and spatial-average SAR (1 g) values, compared to the dual-band PIFA (high-Q).

The low-Q antenna approach to reduce near-field strength and spread the current distribution to reduce discontinuity peaks can also be applied to a variety of other applications, including safety issues of RF energy exposure, such as high power radiating systems and medical electronic systems, efficiency of the antenna system, and co-site interference issues (see Figure

5.46). Alternatively, a high-Q antenna approach can be effective when the range of communication link is critical, such as the link between a scanner and RFIDs.

Future research on low-Q antennas is needed to reduce the antenna size while retaining the wide bandwidth and low-Q nature of the antenna, approaching the theoretical limit. This need is critical in an industry that must consider the value of aesthetics and wide-scale production

It is important to note that one will not necessarily realize the reduced near-field strength of a wideband antenna using a continuous multiple resonance approach. That is because each resonance in the continuous band has a high-Q nature.

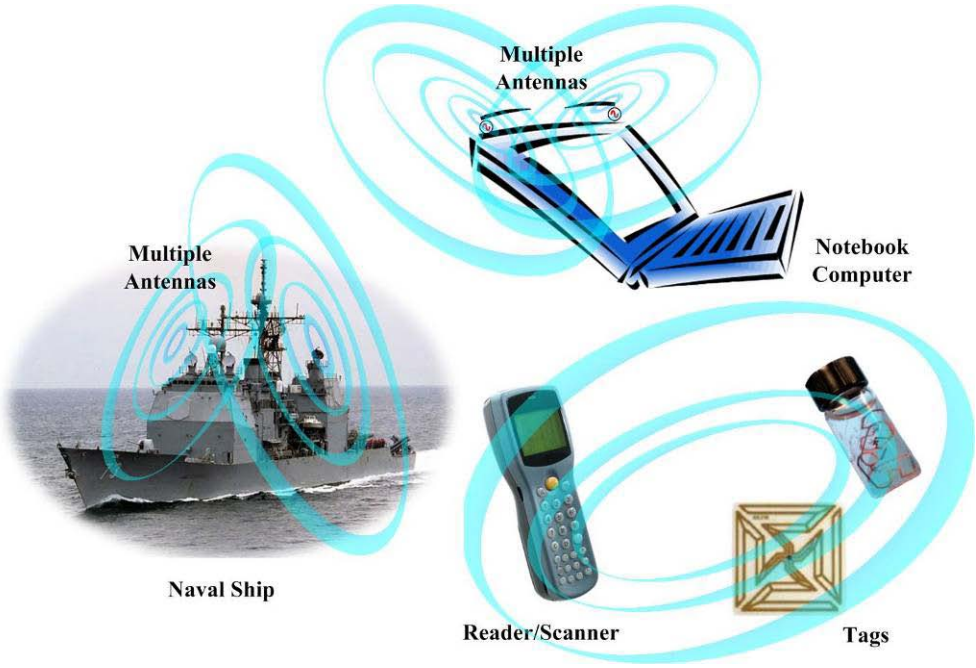


Figure 5.46 Various other near-field interaction problems.

Chapter 6

A Time-Domain Application: Compact Ultra-Wideband Antenna (CUA) Design

Transfer functions of various types of antennas have been investigated in Chapter 3 using the pole-residue model with the spherical wave expansion approach. The transfer function analysis for the canonical antennas helps to understand the correlations between antenna structure and radiation performance in both frequency and time domain, but the size and performance of the canonical antennas are not even close to the theoretical limits, as discussed in Chapter 4.

In this chapter, challenges and trade-offs in a compact ultra-wideband antenna design are addressed. A general design strategy is developed, based on observations from the fundamental-limit theory. An example antenna design for ultra-wideband and software-defined radio applications is introduced.

6.1 Overview of Compact Ultra-Wideband Antenna Design

There are many challenging issues involved in designing ultra-wideband (UWB) systems, such as antenna design, interference, propagation and channel effects, and modulation methods [*Lembrikov*, 2011; *J Reed*, 2005; *Sahinoglu et al.*, 2008; *Stutzman and Thiele*, 2012]. Designing the UWB antenna can be one of the most challenging of these issues. For example, UWB antennas must cover an extremely wide band, 3.1 GHz to 10.6 GHz for indoor and handheld

UWB applications, have compact size, and hold a reasonable impedance match over the band for high efficiency. In addition, they are required to have a non-dispersive characteristic in time and frequency, providing narrow pulse duration to enhance a high data throughput.

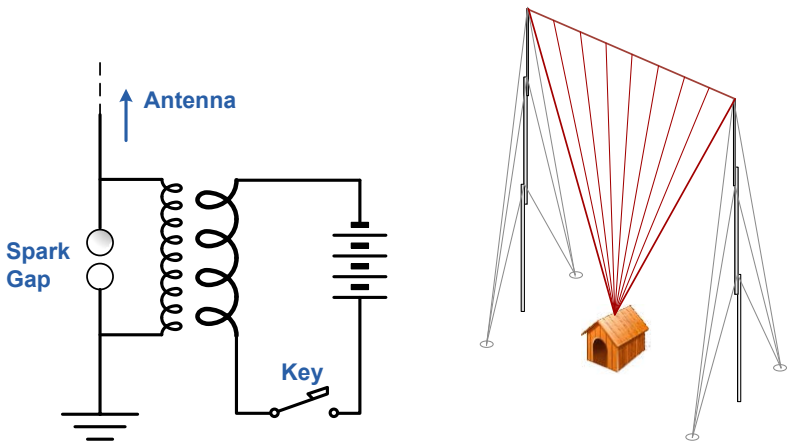
In the following sections, UWB technology is briefly reviewed and challenges in compact UWB antenna design are addressed.

6.1.1 Brief Introduction of Ultra-Wideband Technology

The first ultra-wideband (UWB) communication system may have originated from the historical spark gap transmitter [*H Schantz*, 2005] used in Marconi's transatlantic radio link experiment. A simplified circuit of the spark gap transmitter and its antenna are depicted in Figure 6.1. If the key is closed, the spark gap is charged to a high voltage through the induction coil. When the potential across the spark gap reaches to the sufficiently high to break down the air insulation of the gap, a spark occurs. As the spark gap discharges, it generates a short input pulse. The short pulse radiates through the antenna.

Similar to the spark gap transmitter, impulse UWB (I-UWB) technology also sends a train of very short pulses in the time domain, but one pulse cycle is from several hundred picoseconds to several nanoseconds, which is extremely short and almost like an impulse. Thus, the total energy in each pulse of the pulse train is spread over a very large frequency range and the amount of energy at any particular frequency is very low. That makes it possible to co-exist with existing licensed spectrum without causing a significant amount of interference. Therefore, UWB technology is not a new technology. Former names include carrier-free, baseband, time domain, non-sinusoidal, and impulse radio [*Barrett and Vienna*, 2001].

The United States Federal Communications Commission (FCC) adopted the first UWB Report and Order on February 14, 2002 [*FCC*, 2002]. As summarized in Table 6-1, frequency allocations for various applications were specified in the report and order. Since then, the interest in UWB technology has increased substantially in both academics and the market place. The interest is stimulated by the expectation that UWB can solve the data throughput limit without adding the shortage of the frequency resources.



(a) (b)

Figure 6.1 Illustration of the transmitter and antenna used in Marconi's transatlantic communication experiment – (a) Simplified circuit of spark gap transmitter [*Simons*, 1996] and (b) Fan monopole antenna (mono-cone) [*Stutzman and Thiele*, 2012].

Table 6-1 UWB applications and frequency allocations [*FCC*, 2002]

Applications	Bandwidth
Low frequency imaging systems	Below 960 MHz
Mid frequency imaging systems	1.99 GHz – 10.6 GHz
High frequency imagine systems	3.1 GHz – 10.6 GHz
Indoor UWB systems	3.1 GHz – 10.6 GHz
Handheld UWB systems (outdoor)	3.1 GHz – 10.6 GHz
Vehicular radar systems	22 GHz – 29 GHz

For example, consider an indoor UWB system, compared to traditional wireless local area network (WLAN) systems. Simple theoretical bound of maximum raw data throughput ($T_{\max}$ [bits/sec]) for these systems may be estimated from the Shannon-Hartley theorem [*Proakis*, 2000], i.e.

$$T_{\max}(r) = BW \log_2\left[1 + SNR(r)\right] \tag{6.1}$$

where *BW* is bandwidth in Hertz and *SNR(r)* is a received signal-to-noise power ratio at distance *r* in additive white Gaussian noise (AWGN) environment. Thus, the bound on maximum raw data throughput increases linearly with bandwidth, but only logarithmically with the signal-to-noise power ratio. This is the basic reason that UWB technology can have high data throughput

with the lower signal power (equivalent isotropic radiated power, EIRP < -41.3 dBm/MHz) and co-exist with existing licensed spectrum without causing a significant amount of interference (see Figure 6.2). However, potential interference issues to other communication systems have been raised [*Stumpf*, 2003].

The computed maximum raw data throughputs in a free-space environment are shown in Figure 6.3, after combining with the Friis transmission formula [*Stutzman and Thiele*, 2012] and assuming transmitting power of -41.3 dBm/MHz for UWB and 33 dBm/MHz for WLAN with a unity gain transmitting antenna, receiver noise power spectral density of -114 dBm/MHz, and receiver noise figure of 9 dB with a unity gain receiving antenna. Although the realistic maximum throughputs for the indoor UWB and WLANs will vary with multi-path channel environment and system configurations, Figure 6.3 clearly shows that UWB system outperforms other WLAN systems within a short range.

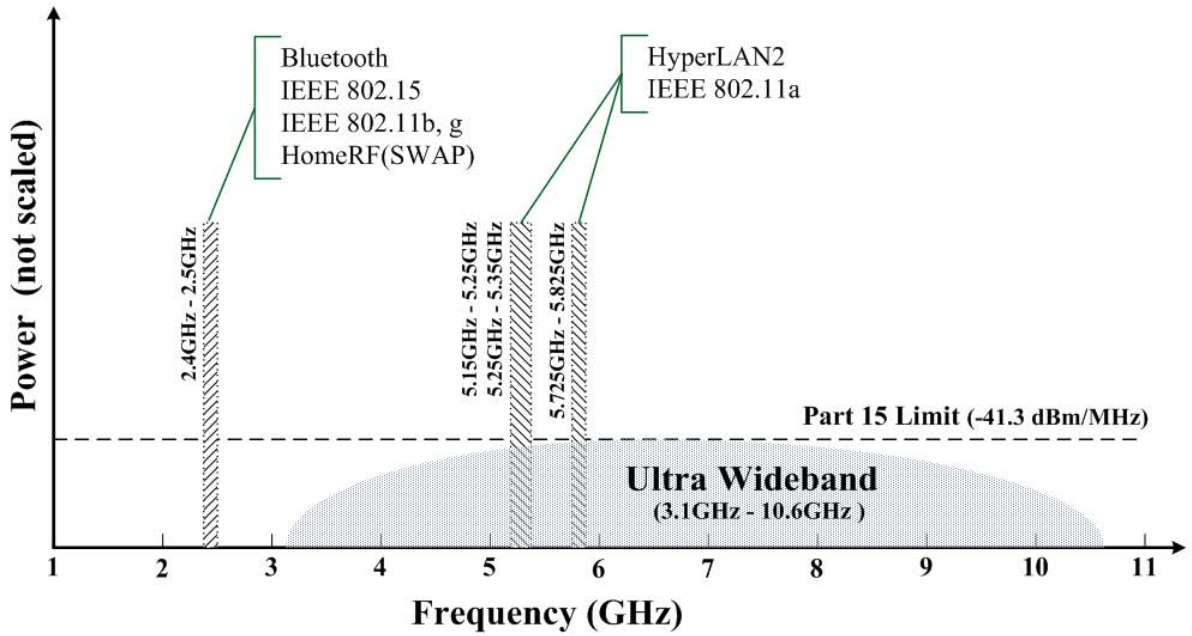


Figure 6.2 Frequency allocation and power level for various wireless network standards and ultra-wideband.

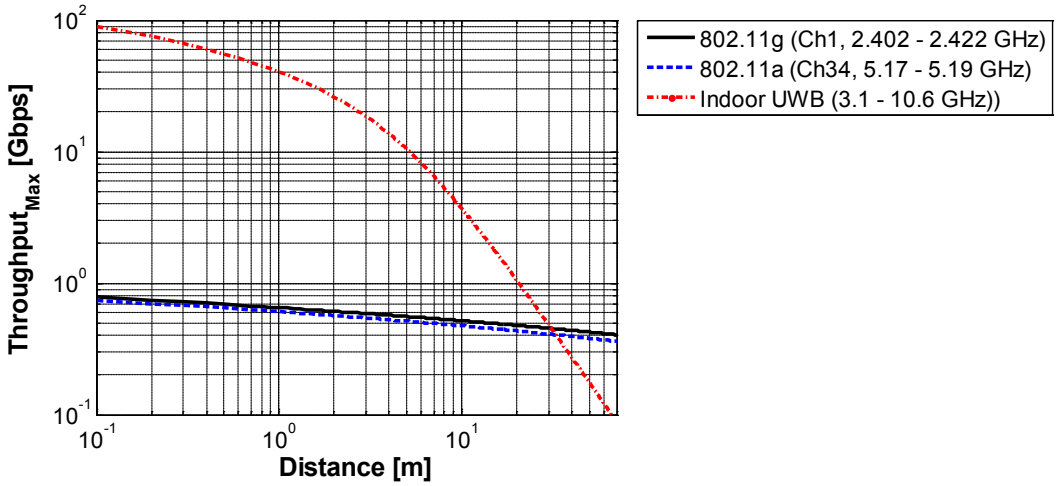


Figure 6.3 Distance versus maximum raw data throughput for IEEE 802.11g WLAN (BW = 20 MHz, EIRP = 33 dBm), IEEE 802.11a WLAN (BW = 20 MHz, EIRP = 33 dBm) and indoor UWB system (BW = 7.5 GHz, EIRP = -41/.3 dBm) in free-space environment. Note that the rated maximum raw data throughput of IEEE 802.11a and IEEE 802.11a are 54 Mbps.

As well as high data transmitting rate, UWB technology can have two additional key advantages over conventional narrow band communication technologies. One advantage is lower cost. Because the UWB transmitter does not need an intermediate frequency stage such that it can modulate baseband signal directly, leading to a low hardware complexity. Also, low cost CMOS technologies can be applied. As a result, UWB systems can have nearly all-digital structure except the conventional RF front-end including the antenna. Thus UWB systems are ideal for realization in software defined radio [*Anderson and Reed*, 2006]. The other key advantage is immunity to multipath and interference. Because the multipath pulses appear later in time, these late-time pulses can be easily eliminated with time gating. In addition, it is extremely difficult to intercept the UWB communication data due to very wide spectrum with a low energy. Thus, it has excellent immunity to interference from other wireless communication system and is inherently secure [*Wilson*, 2002].

With the key and other advantages of the UWB technology over the conventional narrow band communication technologies [*Lembrikov*, 2011; *J Reed*, 2005; *Sahinoglu et al.*, 2008; *Stutzman and Thiele*, 2012], UWB systems have many potential applications to: high-speed wireless personal area network (WPAN) [*Kuang-Hao et al.*, 2008], ground-penetrating radar (GPR) [*Li et al.*, 2012], vehicular radar [*El-Gabaly and Saavedra*, 2011; *Immoreev and Fedotov*,

2002], through-wall radar [*Lanbo et al.*, 2011], surveillance [*Douglas Lynch et al.*, 2007], medical imaging [*Ghavami et al.*, 2012], and high-power microwave [*Giri*, 2004].

Observing the variety of potential applications, WiMedia Alliance (http://www.wimedia.org) and UWB Forum (http://www.uwbforum.org) proposed multi-band orthogonal frequency division modulation (MB-OFDM) UWB [*Bahai et al.*, 2004] and direct-sequence (DS) UWB [*Foerster*, 2002] respectively for IEEE 802.15.3a WPAN standard targeting 480 Mbps. Unfortunately, the IEEE 802.15.3a standardization task group was dissolved because both groups could not converge to same recommendation for the standard. Later, the WiMedia Alliance got approval for WPAN standard, based on their MB-OFDM UWB, from an industry association, ecma international (http://www.ecma-international.org).

Regardless of the standardization issue, the needs for ultra-wideband technology and its applications are keeping researchers and industries motivated to improve the UWB technology and find more exciting applications.

6.1.2 Compact Ultra-Wideband Antenna Design Challenges

Antennas in the frequency domain are typically characterized by radiation pattern, efficiency, gain, and bandwidth [*Stutzman and Thiele*, 2012]. In addition, there exist clear standard definitions related to antenna performance in the frequency domain [*IEEE*, 1993]. However, time domain antenna characterization for I-UWB applications is a different story.

For example, a resonant monopole shown in Figure 6.4a typically would not be considered to be a wideband or ultra-wideband antenna. However, it would be considered to be an UWB antenna in FCC's definition [*FCC*, 2012] of ultra-wideband: more than 500 MHz bandwidth or 20% of fractional bandwidth. The highest and the lowest frequency in the FCC bandwidth definition are referenced to -10 dB below the equivalent isotropic radiated power (EIRP) at a center frequency. As shown in Figure 6.5, the measured EIRP bandwidth of the 3-GHz resonant quarter-wavelength monopole is approximately 1.62 GHz or 52% while the conventional 3-dB gain bandwidth is approximately 0.65 GHz or 22%. Thus, the half-wavelength monopole with 3-GHz resonance is an ultra-wideband antenna in terms of FCC's EIRP bandwidth definition. However, the classification of ultra-wideband antenna does not necessarily mean that the monopole is a good ultra-wideband antenna with good time domain characteristics. The FCC's

ultra-wideband bandwidth definition would come from the conservative viewpoint concerning potential interference issues with existing legacy radio systems.

The shape of the radiated pulse in the time domain is important for an I-UWB antenna. If only the far-field term is considered, the radiated electric field can be written in the time domain as [*Shlivinski et al.*, 1997]:

$$\vec{E}_{rad,FF}(\vec{r},t) = -\frac{\partial}{\partial t}\frac{\mu}{4\pi r}\int_V \vec{J}\left(\vec{r}', t - \frac{\left(r - \hat{r}\cdot\vec{r}'\right)}{c}\right)dv' \tag{6.2}$$

where c is the speed of light. In terms of input pulse of voltage $2V_g$ when matched to the source, (6.2) can re-written as

$$\vec{E}_{rad,FF}(\vec{r},t) = -\frac{\partial}{\partial t}\frac{\mu}{4\pi r}\frac{1}{Z_0}\int V_g(\tau')\vec{h}_{FF}(\hat{r},\tau-\tau')d\tau' \tag{6.3}$$

where $\tau = t - r/c$, Z_0 is a reference impedance, and $\vec{h}_{FF}(\hat{r},\tau)$ is a far-field antenna effective height. Compared to Shlivinski's approach, the time derivative is not included in (6.3), i.e.

$$\vec{h}_{FF}(\hat{r},\tau) = \int_V \vec{J}_{\delta,FF}\left(\hat{r}, \tau + \frac{\hat{r}\cdot\vec{r}'}{c}\right)dv' \tag{6.4}$$

where $\vec{J}_{\delta,FF}$ is the antenna current distribution contributing to far-field radiation for an impulse input. This is because the time derivative should come from a radiation process when the pulse is radiating from reactive near-field to far-field region (see Chapter 2). In order words, the antenna effective height should not provide a time derivative effect for incoming electric fields in a receiving problem. Then, the definition of the effective height in (6.4) can be considered as a far-field antenna transfer function for both transmitting and receiving antenna problems. Further discussion between the effective height and various antenna types can be found in [*Licul and Davis*, 2005].

Due to the time derivative in (6.3), the ideal radiated pulse shape would be a first derivative of an input pulse if the antenna has an impulse-like transfer function. An example of the ideal radiated pulse is shown in Figure 6.6 for a Gaussian input pulse with 3-dB pulse width of 50 pico seconds (corresponding to a 20 GHz wide spectrum).

In the case of the 3-GHz resonant monopole antenna, the measured radiated pulse shape in Figure 6.7a is not a clean doublet because there is sinusoidal damping at the late-time response

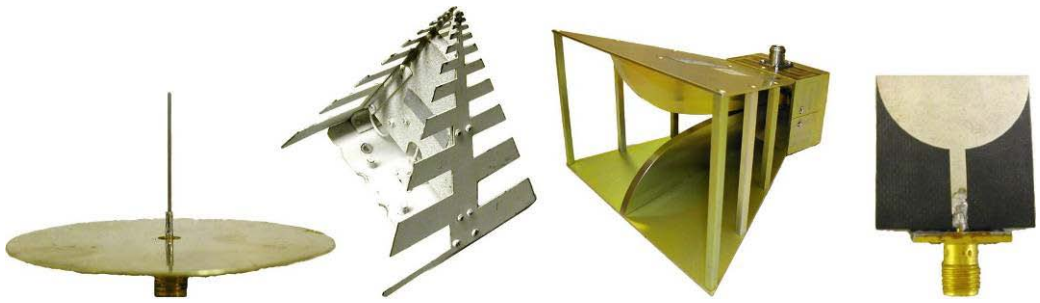
(ringing). However, as the resonant frequency increases, the ringing decays faster with reduced peak amplitude. Alternatively, resistive loading can be considered to convert the resonant antenna to a traveling-wave antenna by sacrificing peak amplitude [*Kanda*, 1978]. Thus, there is a trade-off between the pulse duration and peak amplitude for a given Gaussian pulse input.

On the other hand, frequency independent antennas, such as spiral and log-periodic dipoles, typically have a very large impedance bandwidth with relatively constant gain over frequency. For example, consider the log-periodic trapezoidal antenna in Figure 6.4b. The smallest radiating element contributes to the highest frequency component, and the largest radiating element contributes to the lowest frequency component. Because it is fed at the apex, the smallest radiating element is excited first, and the largest radiation element is excited last. Thus, each element radiates with different time delay, which causes a phase distortion in a frequency domain. The instantaneous bandwidth of frequency-independent antennas is limited due to the phase distortion. As shown in Figure 6.7b, the total response contributed from each element introduces a chirp in the early-time response due to the phase distortion and ringing in the late-time response due to the resonance of overall antenna size. In fact, the phase distortion can be compensated by modifying the input pulse with extra phase delays at the resonant frequency of each element. The resulting input pulse would be like a chirp response, but the radiating pulse shape would be similar to the ideal doublet shape. This concept was demonstrated by Hertel and Smith [*Hertel and Smith*, 2003]. The flip side of this approach is that the input pulse needs to be precisely synthesized. Otherwise, the resulting radiating pulse could experience more severe chirp. In addition, the duration of the both input and radiating pulses increases due to the compensated phase. Thus, achievable overall data throughput using the modified input pulse can be further limited for the I-UWB systems.

TEM horn antennas are a traveling wave antenna and have been used for many years to launch and receive narrow pulses [*Kanda*, 1982]. Unlike the frequency-independent antennas, a TEM horn antenna supports a very large instantaneous bandwidth. The biggest aperture height between the two tapered metallic plates is typically about half wavelength at the lowest operating frequency. The horn aperture width is related to the characteristic impedance of stripline seen at the cross section of the metallic plates. As discussed in Chapter 3, an antenna provides an impedance transformation between the reference impedance (typically 50 Ohms) and the free-

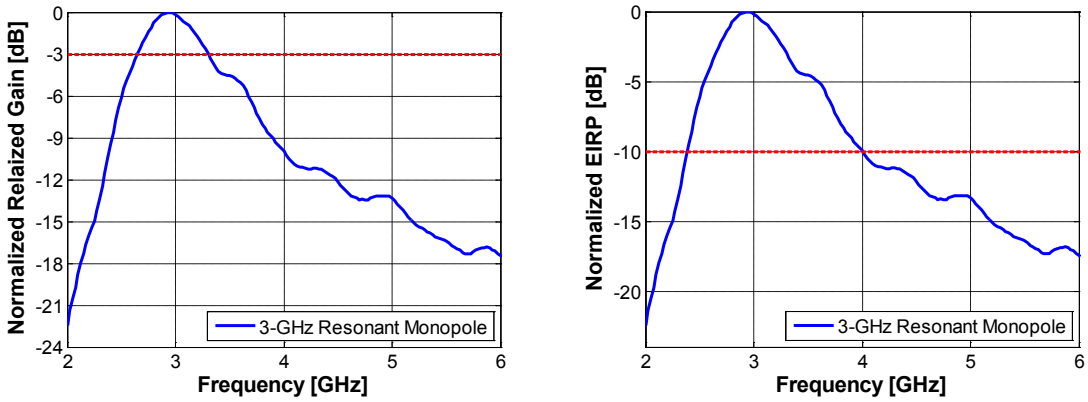
space wave impedance (377 Ohms). Thus, the characteristic impedance of the strip line would increase gradually from 50 Ohm at the feed to 377 Ohms at the open aperture. The overall length of the TEM horn is typically more than one wavelength in order to launch traveling waves. If an exponential tapering is used, the overall length can be effectively reduced. The exponential tapering can be directly applied to the horn plates. Alternatively, exponentially-tapered feed can be integrated with the original TEM horn antenna, as shown in Figure 6.4c. Basically, the exponentially-tapered feed can be viewed as a Vivaldi antenna [*Gibson*, 1979]. The TEM horn with the Vivaldi antenna is referred as *TEM double-ridged horn*. The radiated pulse of the TEM doubly-ridged horn in Figure 6.7c has a doublet close to the ideal response (Figure 6.6b) at early time and a ringing at the late time response. The ringing also comes from the resonance of the overall antenna structure. The resonant frequency is typically located below the operational frequency, but within the frequency spectrum corresponding to the input pulse. Resistive loadings can be applied on the truncated edges in order to suppress the ringing by sacrificing efficiency [*Kanda*, 1982; *Shlager et al.*, 1995].

An alternative way to mitigate the ringing effect of the TEM horn antenna without sacrificing the efficiency would be locating the resonance of the overall structure close to the lowest operating frequency of the tapered slot structure. For example, the planar half-disk antenna in Figure 6.4d acts like an asymmetric dipole at lower operating frequencies and a dual slot at higher operating frequencies [*Yang and Davis*, 2004]. The asymmetric dipole mode comes from the resonance of overall antenna structure. In addition, the bottom structure and feed of the half-disk antenna act as a pseudo balun at the low operational frequencies, where the half-disk antenna is seen as an asymmetric dipole, and a microstrip transmission line at high operating frequencies, where the half-disk is seen as a dual slot. As a result, the planar half-disk antenna supports a short pulse radiation with a minor ringing as shown in Figure 6.7d. Thus, compact antenna size can help to reduce the ringing effect as well as meeting industrial needs for compact antennas. The fringing fields of the dual slot provide reasonable omni-directional coverage in Azimuth even at higher frequencies due to the nature of the planar structure. The half-disk antenna might be classified as a wideband antenna having mixed operational modes (dipole + dual slot) and providing extremely large operational bandwidth, not traveling wave antenna.



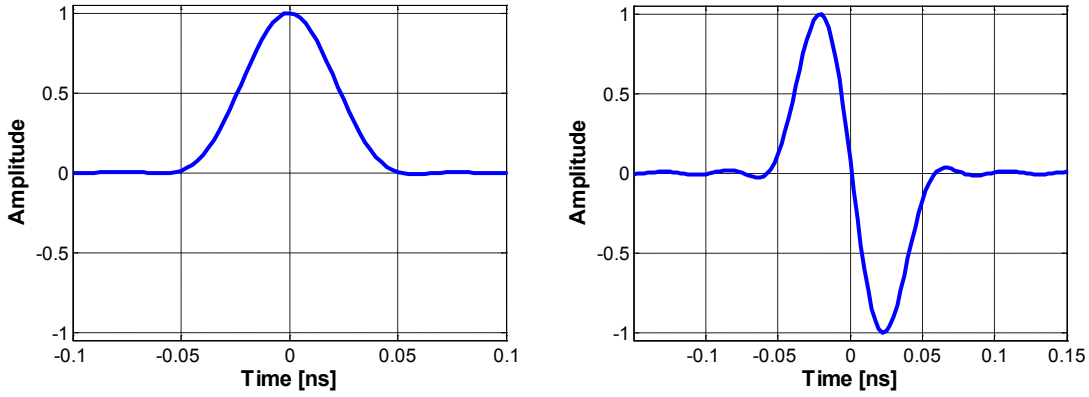
(a) (b) (c) (d)

Figure 6.4 Picture of canonical antennas – (a) Monopole antenna (resonant), (b) Commercial log-periodic trapezoidal (frequency-independent), (c) Commercial TEM double-ridged horn (traveling-wave), and (d) Half-disk (wideband) [*Yang and Davis*, 2004]. Antenna size is not scaled with respect to the size of other antennas.



(a) (b)

Figure 6.5 Radiation performance of 3-GHz resonant quarter-wavelength monopole estimated from the measured link response of two identical antennas – (a) Normalized gain versus frequency and (b) Normalized EIRP versus frequency.



(a) (b)

Figure 6.6 Gaussian input pulse and ideal radiating pulse – (a) Measured Gaussian pulse with 50 pico-second 3-dB pulse width and (b) The first derivative of the Gaussian input pulse. The Gaussion input pulse was measured from Agilent 8510 using the open calibration standard.

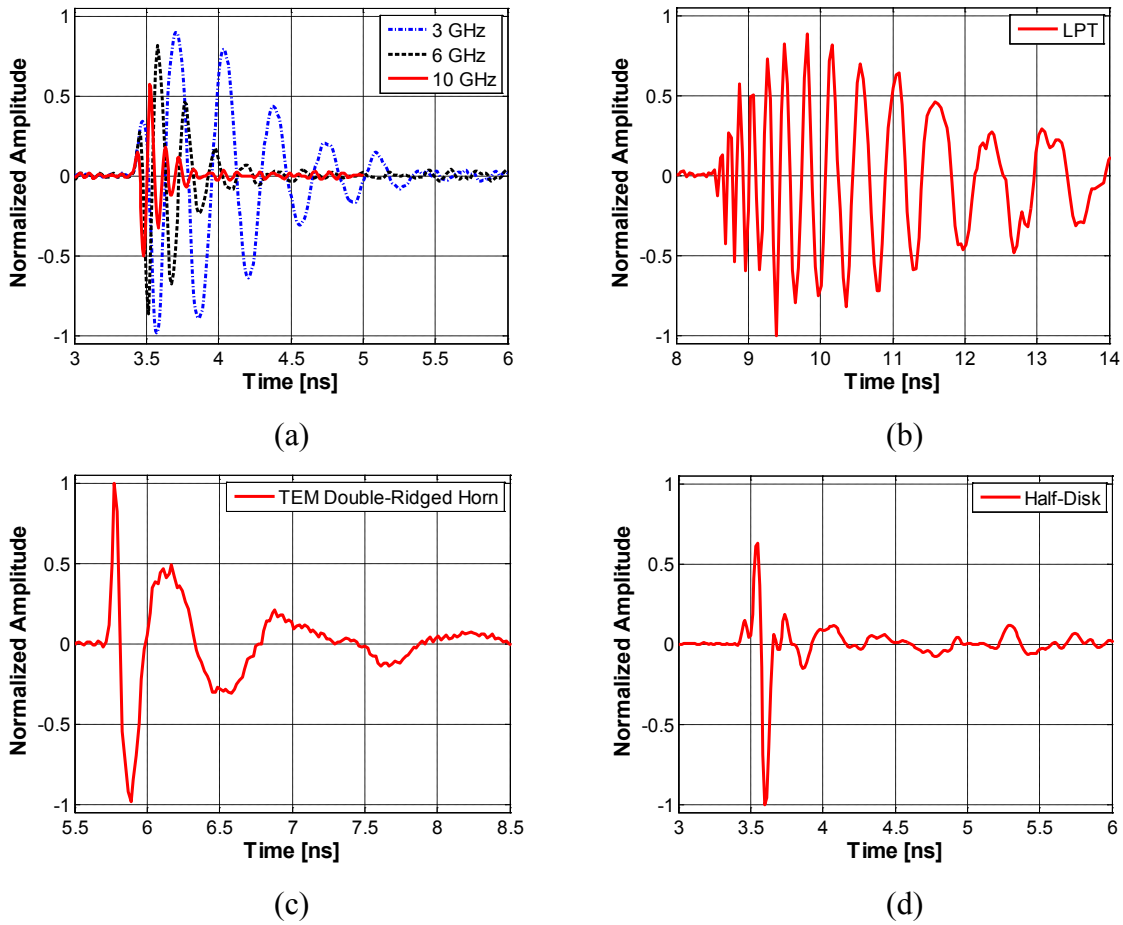


Figure 6.7 Measured radiated pulse of canonical antennas – (a) Monopole antennas (resonant) [*Yang and Davis*, 2005], (b) Log-periodic trapezoidal antenna (frequency-independent) [*Licul*, 2004], (c) TEM double-ridged TEM horn (traveling-wave) [*Licul*, 2004], and (d) Half-disk (wideband) [*Yang and Davis*, 2004].

Phase responses versus frequency for the discussed antennas representing canonical antennas, i.e. resonant, frequency-independent, traveling-wave, and wideband antennas, is compared in Figure 6.8 and their pulse radiation performance is summarized in Table 6-2.

The resonant monopole would be an excellent choice for short-range I-UWB applications if antenna footprint is a critical design factor. However, the resonant monopole would be not the best choice for MB-OFDM UWB applications due to the inherent, limited impedance bandwidth.

The frequency-independent, log-periodic trapezoidal antenna would not be used for I-UWB applications due to the severe phase distortion and extremely long pulse duration. However, frequency-independent antenna can be good candidate for MB-OFDM UWB. MB-OFDM UWB systems use multiple sub-bands and each sub-band has a separate carrier. Thus, the performance

of MB-OFDM UWB systems is not noticeably affected by phase distortion introduced by a frequency-independent antenna.

The traveling-wave, TEM double-ridged antenna can be a good candidate for both I-UWB and MB-OFDM UWB applications due to the narrow pulse width, extremely-wide operational band and relatively-constant gain over frequency. However, efficient traveling antennas often have large size and directivity (limited angular coverage). The operating bandwidth of the TEM double-ridged antenna is practically limited due to the mechanical fabrication precision of the tapered feed structure.

The wideband, half-disk antenna can also be a good candidate for both I-UWB and MB-OFDM UWB applications. Unlike the TEM double-ridged antenna, the half-disk antenna has a reasonable omni-directional angular coverage and a compact planar form factor. Thus, the half-disk antenna would be a good choice for wireless personal area network (WPAN) applications requiring wide angular coverage and compact antenna size.

There might be an optimum choice of antenna types for a specific application and physical-layer configuration (I-UWB or MB-OFDM UWB) to the trade-offs in antenna design and performance. However, in general, an excellent I-UWB antenna with high pulse-amplitude (high efficiency) and narrow pulse-duration (large instantaneous operational bandwidth) also shows excellent frequency-domain antenna performance. In a sense, antenna design for I-UWB applications might be more challenging due to the required phase linearity or large instantaneous bandwidth, compared to the case for MB-OFDM UWB applications. As we discussed in Chapter 4 and Chapter 5, an antenna that has large instantaneous bandwidth has a low radiation Q, which implies that excellent impulse-radiating antennas may experience less interaction issues when the antennas are co-located with other antennas. However, the size of excellent I-UWB antennas is often more than a quarter-wavelength at the lowest operational frequency. Thus, the most challenging design issue in ultra-wideband antenna design would be size miniaturization without sacrificing operational bandwidth and efficiency.

In the next section, we discuss how we can design an ultra-wideband antenna with size and performance close to the fundamental limits for both I-UWB and MB-OFDM UWB applications.

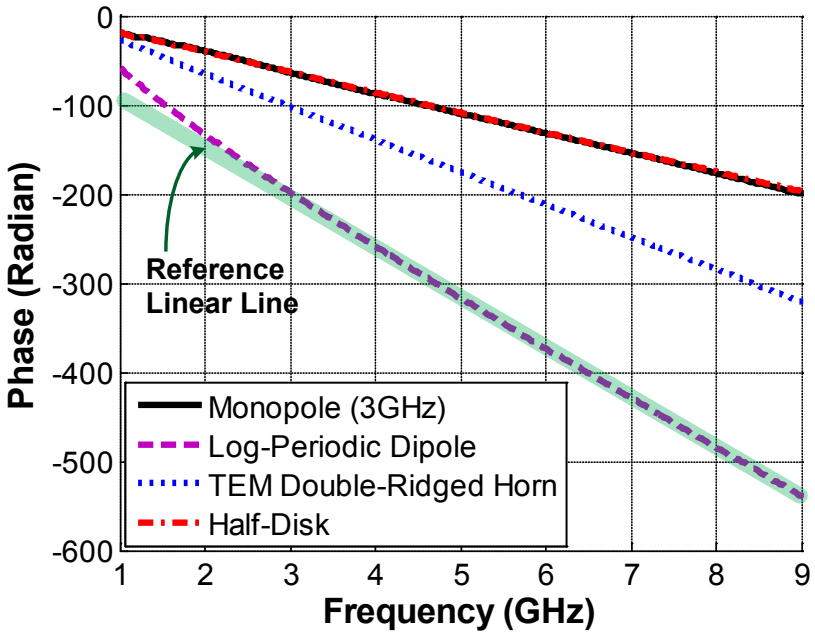


Figure 6.8 Measured phase response versus frequency comparison between the canonical antennas [*Yang and Davis*, 2005].

Table 6-2 Pulse radiation performance of antennas representing canonical antennas for 50 pico-second Gaussian pulse input [*Yang and Davis*, 2005]

Antenna	Pulse Duration [ns]	Pulse Type	Phase
Monopole (3 GHz) [resonant]	2.1	Damped sinusoid	Linear
Log-periodic Trapezoidal antenna [frequency-independent]	11.0	Chirp	Nonlinear at low freq.
Ridged TEM horn [traveling-wave]	2.7	Damped sinusoid	Linear
Half-disk antenna [wideband]	0.8	Damped sinusoid	Linear

6.2 Design Approach of Ultra-Wideband Antenna Close to Fundamental Limits

The concept of the ideal antenna having a high pass characteristic was discussed in Chapter 4. The ideal antenna concept provides insight into antenna design strategies in terms of the trade-off

relationship among antenna size, impedance bandwidth, and gain. The design strategies are illustrated in Figure 6.9 for the spherical TM_{01} mode.

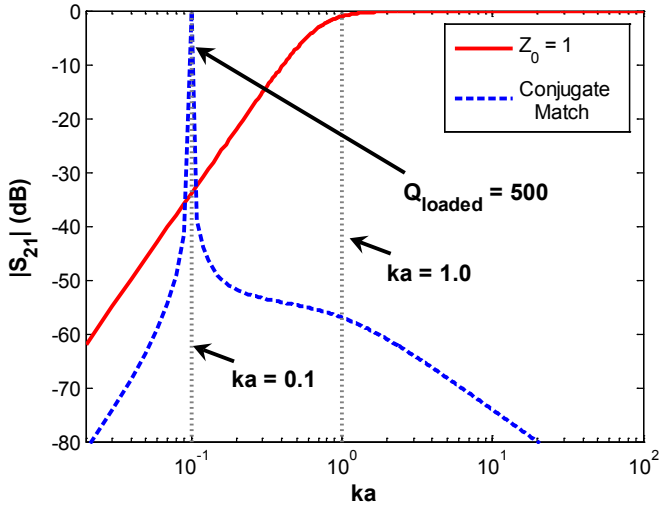


Figure 6.9 Illustration of antenna design strategy in terms of electrical size of antenna and impedance bandwidth (or radiation-Q).

For a given frequency, the product of wave number k and radius a of the antenna represents the antenna electrical size. If the product ka of an antenna is greater than unity, one may choose the system reference impedance equal to the real part of the antenna input impedance in order to maximize the impedance bandwidth of the antenna. For ka smaller than unity (giving an electrically-small antenna), the antenna should be tuned to reduce return loss and maximize realized gain. In the example of Figure 6.9, the antenna was tuned with a conjugate match, such that there is 100 % transmission at the resonant frequency. However, the impedance bandwidth was limited by the tuning requirement compared to the $ka > 1$ case. The maximum achievable impedance bandwidth of the tuned antenna can be found from the classical fundamental-limit theory by using the inverse-proportion relationship between the impedance bandwidth and the radiation-Q. The radiation-Q formulas in the classical fundamental-limit theory literature are for unloaded-Q. So, an achievable impedance bandwidth of a resistively-loaded antenna becomes twice that of the unloaded antenna. This aspect is often neglected by researchers.

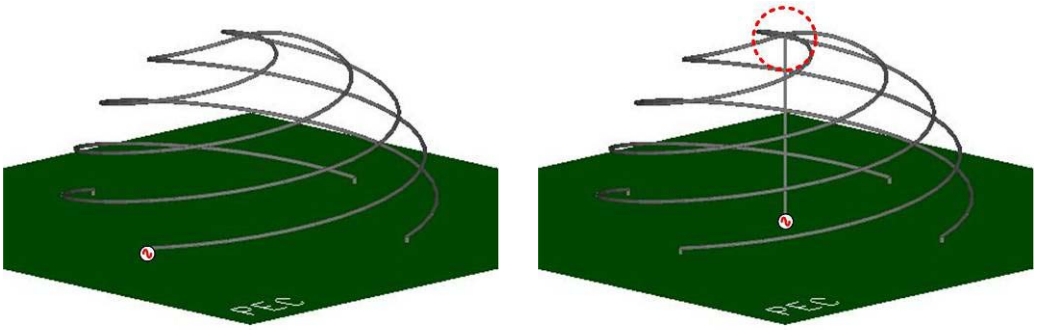
Based on the above antenna design strategy for an omni-directional antenna, it is a reasonable choice to set the size of the antenna to satisfy $k_L a = 1$, where k_L is the wave number

at the lower bound of the operating frequency for an ultra-wideband antenna close to fundamental-limit performance and size. Thus, the size of the antenna is same as the size of the radiation sphere of the spherical TM_{01} mode at the lower bound of operating frequency. For example, if we choose 3 GHz as a lower bound, the radius of the antenna sphere, *a,* will be 15.9 mm.

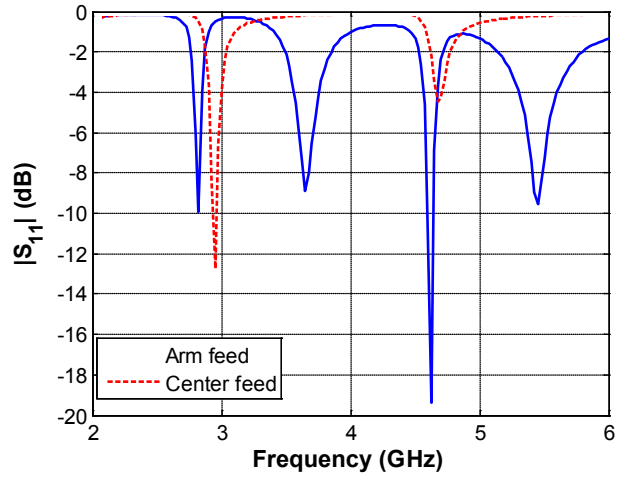
The next question is what shape of antenna will be close to the characteristics of the ideal antenna. One of the reasons that conventional antennas are not close to the limit curve (see Figure 4.22) is that these antennas do not efficiently utilize the volume of the given antenna sphere, reducing non-radiated stored energy and maximizing impedance bandwidth. This suggests a meshed spherical shape as a candidate antenna. A solid spherical shape, however, would cause difficulty in the feed structure. Recently, *Yang et al.* [*Yang et al.*, 2008] showed that an antenna without a slope discontinuity of current distribution on an antenna structure has wideband characteristics and reduces non-radiating near-field energy. This kind of antenna typically has a smooth variation in shape. The meshed spherical structure also satisfies this aspect.

A similar structure to the meshed sphere would be a spherical helix [*Cardoso and Safaai-Jazi*, 1993; *Safaai-Jazi and Cardoso*, 1996]. More recent research produced the multi-arm hemispherical helix over a ground [*Clark and Safaai-Jazi*, 2004]. Their design effort was mainly focused on an axial-mode excitation for global positioning system (GPS) applications. *Best* [*Best*, 2004a] revisited the multi-arm hemispherical helix, but the emphasis on his work was to design an electrically-small, normal-mode helix. Similar to the original hemispherical helix design of Clark and Safaai-Jazi, Best's antenna has multiple arms connected in parallel (see Figure 6.10a). The antenna is excited at the bottom of one of the arms. Best demonstrated that the radiation resistance of the antenna increases as the number of arms increases. Thus, the folded hemispherical helix can easily be tuned at a desired frequency by adjusting the number and length of the arms. As shown in Figure 4.22, size versus radiation-Q (or impedance bandwidth) performance of the antenna is very close to the fundamental-limit curve. Unfortunately, the folded hemispherical helix has multiple resonances over a wide frequency range (see Figure 6.10c). Due to the unique feed location of the folded hemispherical helix, the antenna has resonances at every harmonic frequency. Thus, the current form of the folded

spherical helix cannot be used for ultra-wideband applications. In order to reduce the number of strong resonances, a center-fed version of the folded hemispherical helix antenna (Figure 6.10b) is considered. As compared in Figure 6.10c, the center-fed version has a deep resonance around 3 GHz and the number of the resonance frequencies was reduced significantly compared to Best's original folded hemispherical antenna.



(a) (b)

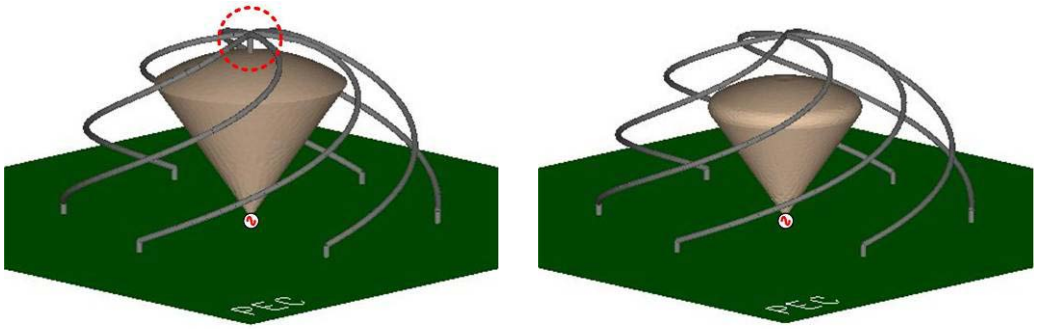


(c)

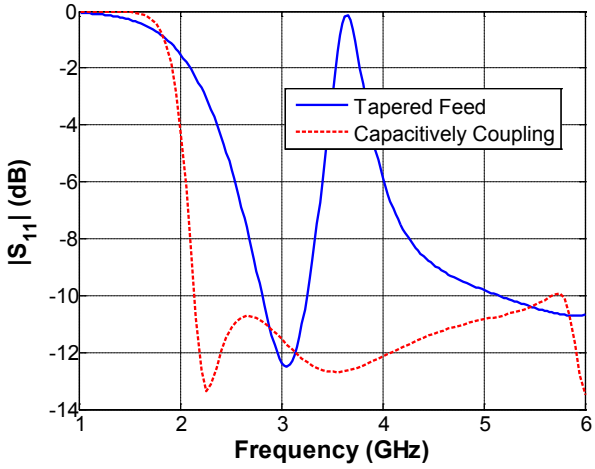
Figure 6.10 Illustration of antenna design evolution process for feed location selection – (a) Half structure of spherical helix [*Best*, 2004a], (b) Center-fed hemispherical helix, (c) Return loss comparison. Infinite PEC ground plane was assumed. Height of antenna was 15.9 mm. A commercial moment method code [*FEKO*, 2007] was used for simulations.

The next step of this design evolution is to increase the impedance bandwidth of the center-fed folded spherical helix. Recall that we assumed the reference impedance for the extended equivalent model (Figure 4.23) to be equal to the normalized intrinsic wave impedance in order to maximize the impedance bandwidth in developing the concept of an ideal antenna in Chapter

4. However, 50 Ohms is preferred as system reference impedance in many applications. This suggests that a distributed impedance transforming structure needs to be included in the center-fed folded hemispherical helix. Thus, a tapered feed structure (cone) is added in the helix. The top of the cone structure is connected to the hemispherical helix through a wire as shown in Figure 6.11.



(a) (b)



(c)

Figure 6.11 Illustration of antenna design evolution process for bandwidth enhancement – (a) Hemispherical helix with a cone feed (the top of feed is connected to the helix), (b) Capacitively-coupled hemispherical helix with a cone feed, (c) Return loss comparison. Infinite PEC ground plane was assumed. The antenna height is 15.9 mm.

The impedance bandwidth of the center-fed folded hemispherical helix with a cone feed noticeably increased, compared to the version without the tapered feed (see Figure 6.11c). However, a frequency notch is observed around 3.6 GHz. The notch splits the impedance bandwidth into two parts. After investigating the current distribution of the antenna, it turned out

that the notch was caused by an equivalent slot stub existing between the top area of the cone and the helical arms on the cone. A solution to eliminate the notch is to remove the wire connection between the helix arms and the top of the cone. Thus, the hemispherical helix is capacitively coupled to the cone (see Figure 6.11b). As shown in Figure 6.11c, the notch was effectively removed.

All-pass-filter characteristics of the extended model of the antenna in Figure 4.23 was implicitly assumed in the classical fundamental-limit theory. However, it seems to be an impossible antenna design. Then, an alternative approach may be to design an antenna that has high-pass filter (HPF) characteristics like the characteristics of the spherical mode. If the cut-off frequency of the HPF antenna is lower than the cut-off frequency of the spherical mode, we may have an antenna design close to the fundamental limit. In fact, the capacitively-coupled hemispherical antenna of this design evolution is one such design. This antenna basically has a high-pass filter structure, i.e. the series capacitance due to the coupling between cone and helical arms and the shunt inductance due to the hemispherical arms connected to the ground.

The simplified circuit model of the capacitively-coupled hemispherical helix with a cone feed is depicted in Figure 6.12. Three cases are compared with this model:

Case 1. Spherical TM_{01} mode only ($Z_0 = 1$, N=1, and $SF = 0$)

Case 2. Spherical TM_{01} mode + HPF-like antenna ($Z_0 = 1$, N=1, and $SF = 1.5$)

Case 3. Spherical TM_{01} mode + HPF-like antenna + Impedance Transformer
($Z_0 = 50/377$, N=0.25, and $SF = 1.5$)

where Z_0 is the system reference impedance, N is the turns ratio of impedance transformer, and SF is the scaling factor to control the cut-off frequency of the antenna. In Figure 6.13, the normalized radiated power ($|s_{21}|^2$) of the cases are compared. The performance of *Case 1* is the same as the ideal antenna. The result of *Case 2* shows that we can lower the lower bound of the impedance bandwidth by adding HPF-like antenna to *Case 1*, though there is some mismatch. In order to mitigate the mismatch issue and change the reference impedance to 50 Ohms, we need to adjust the turns ratio (N) of the impedance transformer, which is *Case 3*. Thus, within the operational bandwidth, the performance of the *Case 3* model, including the HPF-like antenna and impedance transformer is close to the performance of the ideal antenna, which is the fundamental performance limit. However, outside the lower end of the bandwidth, the radiation

performance of the *Case 3* model drops quickly, comparing with the ideal antenna. The minor fluctuation in the response of *Case 3* may be improved by applying an additional matching structure or a higher order HPF-like antenna structure.

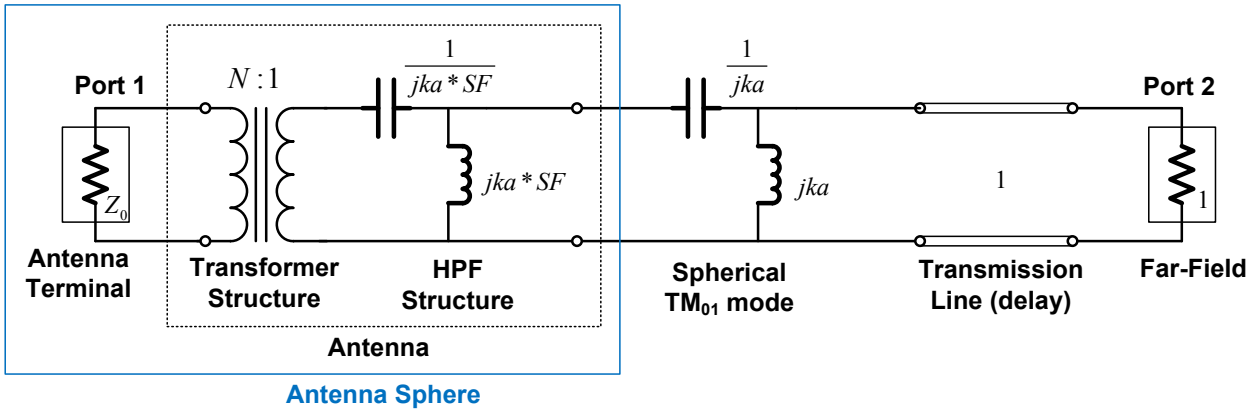


Figure 6.12 Simplified circuit model of the capacitively-coupled hemispherical helix with a tapered feed. *N* is turns ratio of impedance transformer. *SF* is the scaling factor to control cut-off frequency of HPF-like antenna structure.

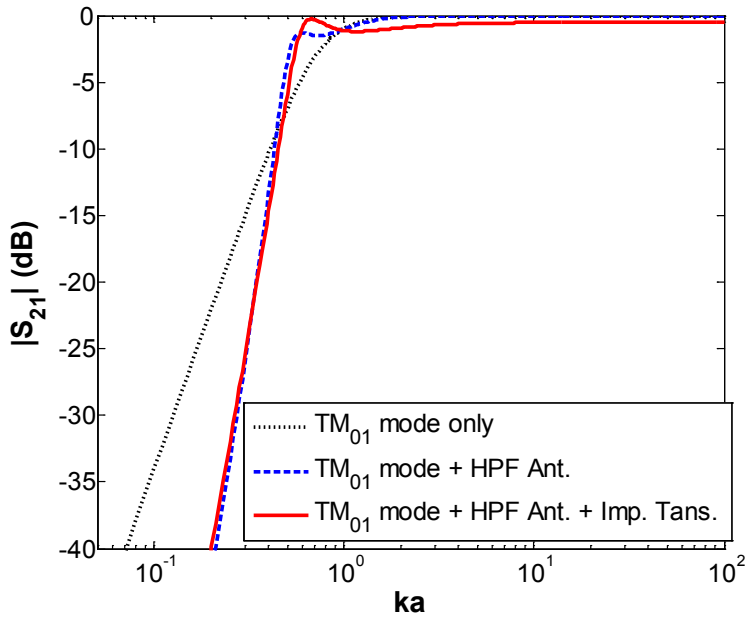


Figure 6.13. $|s_{21}|^2$comparison for various cases in simplified circuit of the capacitively-coupled hemispherical helix with a cone feed: TM_{01} mode only (Z_0 = 1, N=1, and SF = 0), TM_{01} mode + HPF antenna (Z_0 = 1, N=1, and SF = 1.5), and TM_{01} mode + HPF antenna + Impedance Transformer (Z_0 = 50/377, N=0.25, and SF = 1.5). Z_0 is system reference impedance. N is turns ratio of impedance transformer. SF is the scaling factor to control cut-off frequency of antenna. Each $|s_{21}|^2$ data was computed from analytical solution for the circuit in Figure 6.12.

6.3 Wire-Loaded Antenna Design for Ultra-Wideband Applications

The geometry and performance of the designed compact, folded, multi-arm, hemispherical, ultra-wideband antenna are presented in this section. The performance of the designed antenna is characterized in both frequency and time domains.

6.3.1 Antenna Geometry

The overall geometry of the ultra-wideband folded hemispherical helix is depicted in Figure 6.14 with antenna design parameters, including: cone-feed angle (α), cone-feed height (h_c), and overall antenna height (h_a). The cone feed has a tapered top and blended edge in order to mitigate any possible slop-discontinuity in current distribution on the cone. The smoothly-tapered shape also helps to extend the frequency range of impedance transformation. Six wire arms are connected to each other in parallel and the whole antenna is above a ground plane. Elevation angle location (θ, $0 \le \theta \le \pi/2$) for each arm is found from

$$\theta = \cos^{-1}\left(\frac{\phi}{2\pi n}\right) \tag{6.5}$$

where ϕ is the arm rotation angle ($0 \le \phi \le 2\pi n$) and n is the number of turns in azimuth. In Cartesian coordinates, the overall trajectory of each arm is expressed as

$$\begin{bmatrix} x_a \\ y_a \\ z_a \end{bmatrix} = \begin{bmatrix} h_a \sin\theta\cos\phi \\ h_a \sin\theta\sin\phi \\ h_a \cos\theta \end{bmatrix} \tag{6.6}$$

The assigned values for the antenna design parameters are listed in Table 6-3.

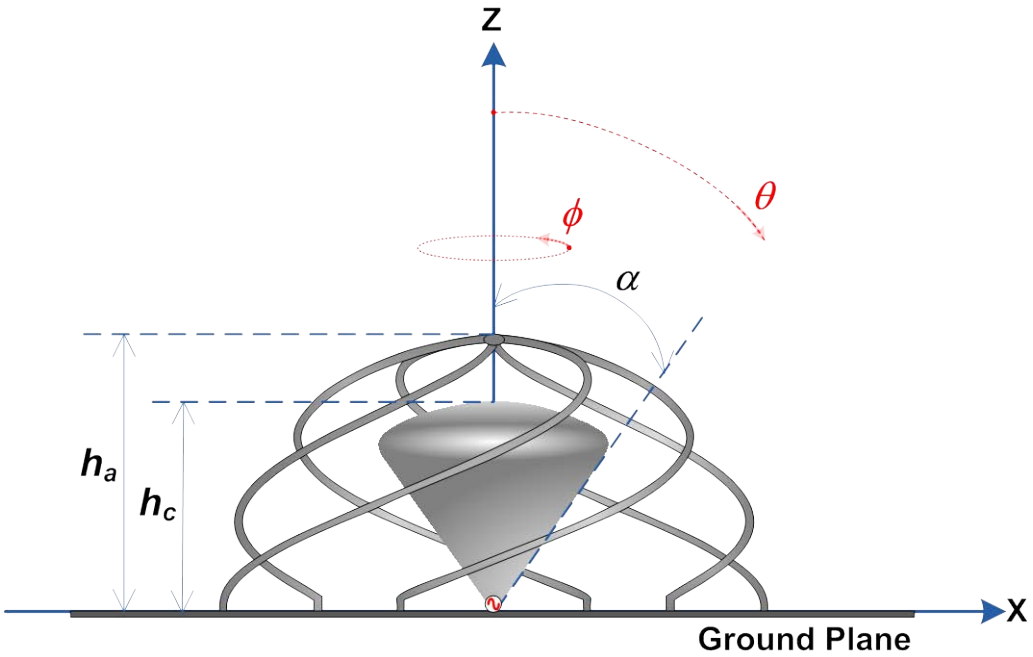


Figure 6.14 Overall geometry of compact ultra-wideband folded hemispherical helix with design parameters.

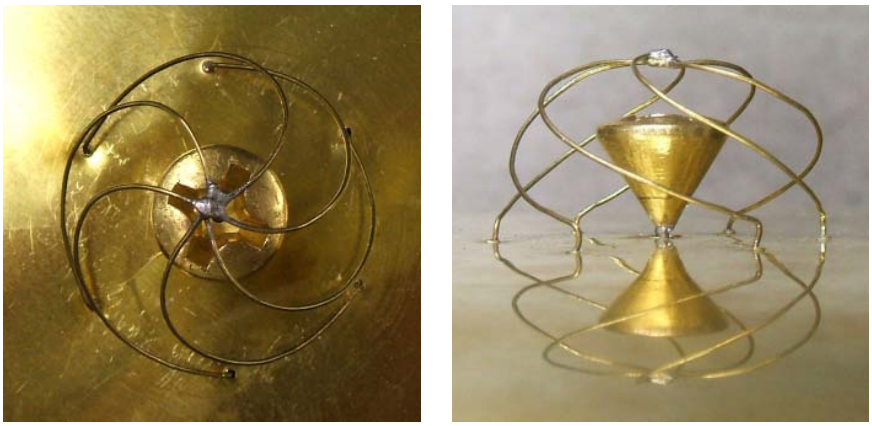
Table 6-3 Values for antenna design parameters

Antenna Design Parameter	Design Value
Overall antenna height (h_a)	15.65 mm
Feed-cone height (h_c)	12.01 mm
Feed-cone angle (α)	0.61 radian
Number of turns of each arm (n)	0.43
Wire diameter of each arm	0.50 mm

6.3.2 Simulated and Measured Performance

A prototype of the designed compact ultra-wideband antenna in the previous section was fabricated out of a brass on a finite, circular ground with 150 mm radius. Pictures of the fabricated prototype and VSWR performance are shown in Figure 6.15 and Figure 6.16, respectively. The simulation results assuming an infinite ground were obtained using two different commercial computational codes, a finite-difference time-domain (FDTD) code [*CST*, 2007] and a method of moments (MoM) code [*FEKO*, 2007]. Both simulated results match very well with the measured results obtained with an Agilent 8510 vector network analyzer (VNA). A small difference between simulation and measurement results at high frequencies are probably

due to fabrication errors. The lower bound of the operational frequency range is about 2.14 GHz. The electrical size of the fabricated antenna is approximately 1/8 wavelength at the lower bound, which is more than 2 times smaller than conventional quarter-wavelength ultra-wideband antennas.



(a) (b)

Figure 6.15 Pictures of the constructed compact ultra-wideband antenna on a finite size ground (round shape, radius of the ground is 150 mm) – (a) Top view and (b) Side view.

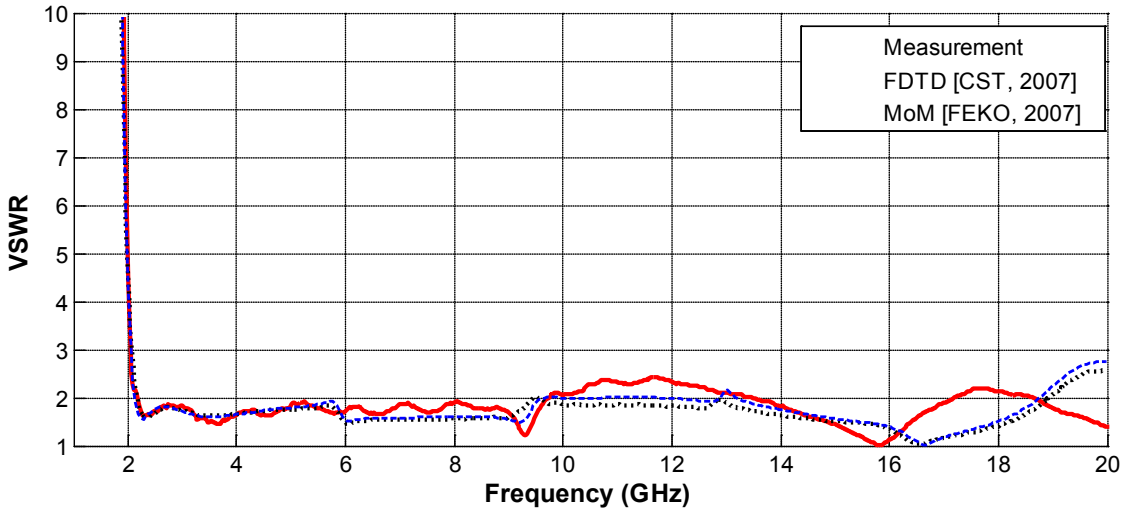
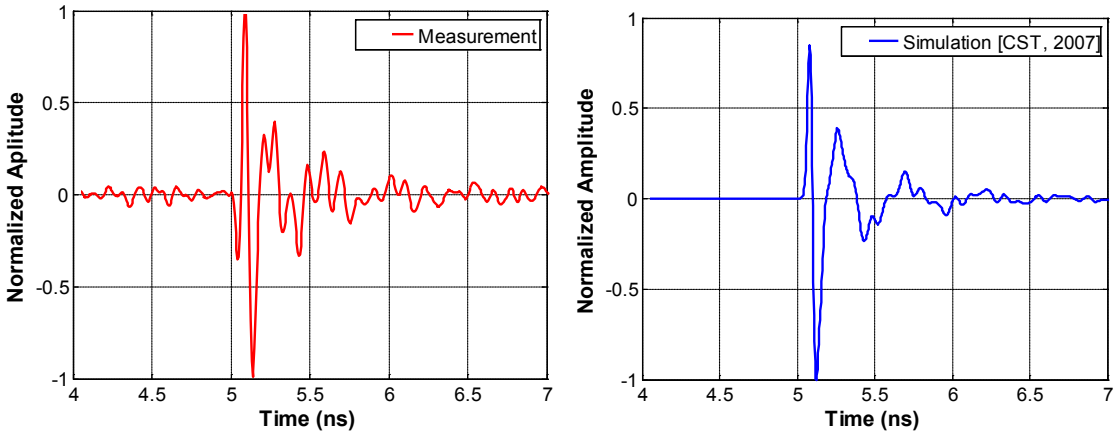


Figure 6.16 VSWR performance comparison between measurement and simulation results for the designed compact ultra-wideband antenna (Z_0 = 50 Ω). Commercial finite-difference time-domain (FDTD) code [*CST*, 2007] and method of moments (MoM) code [*FEKO*, 2007] were used for the simulations.

The radiated pulse of the designed antenna for a Gaussian input pulse with a 3-dB pulse width of 50 pico-second was measured, and the results are shown in Figure 6.17a. Two identical

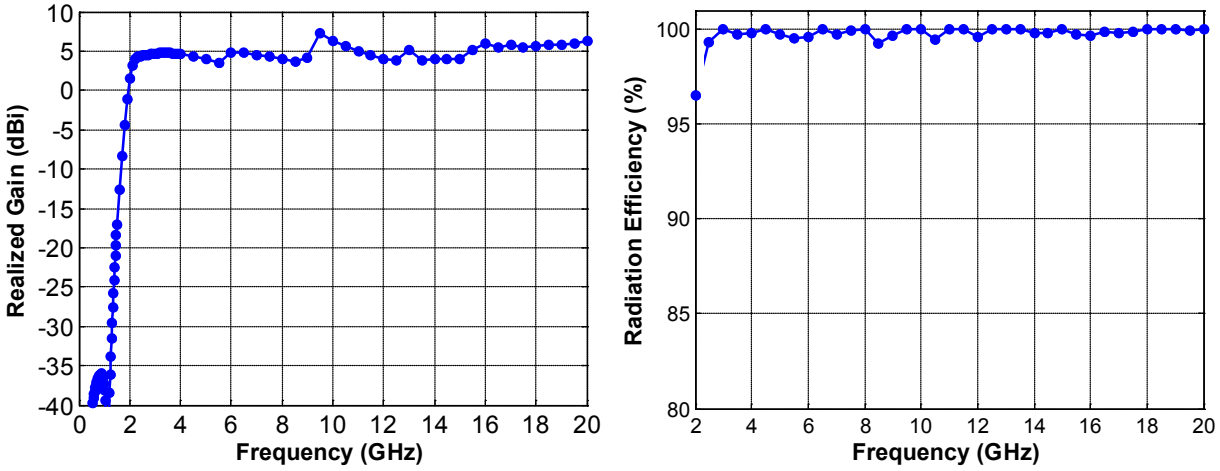
antennas were used for transmit and receive the radiated pulse. These antennas were separated by 1.5 m in order to be approximately in the far-field range. The link response ($|s_{21}|$) of two identical antennas dictates the radiated pulse of the antenna. The radiated pulse showed a doublet waveform with minor ringing, similar to the case of the other excellent ultra-wideband antenna designs discussed in the previous sections, but the electrical size of the designed compact ultra-wideband antenna is half the size of a typical quarter-wavelength size UWB antenna. The measured and simulated radiated pulses in Figure 6.17 agree very well with each other, indicating that the measured result is repeatable. Note that the prototype of the design antenna has a finite-size ground plane while the simulation model has an infinite ground plane.

Simulated realized gain and radiation efficiency versus frequency are shown in Figure 6.18. The gain is relatively constant over the operational frequency range. Radiation efficiency was also close to 100% except the frequencies around the lowest operational frequency. The relatively constant gain implies that that a single spherical mode was strongly excited and dominates the antenna radiation performance. As a result, the radiation pattern would be also relatively consistent over the bandwidth. Indeed, the simulated radiation patterns for the designed compact ultra-wideband antenna over infinite ground plane are fairly consistent over the entire band (see Figure 6.19) except 10 GHz.



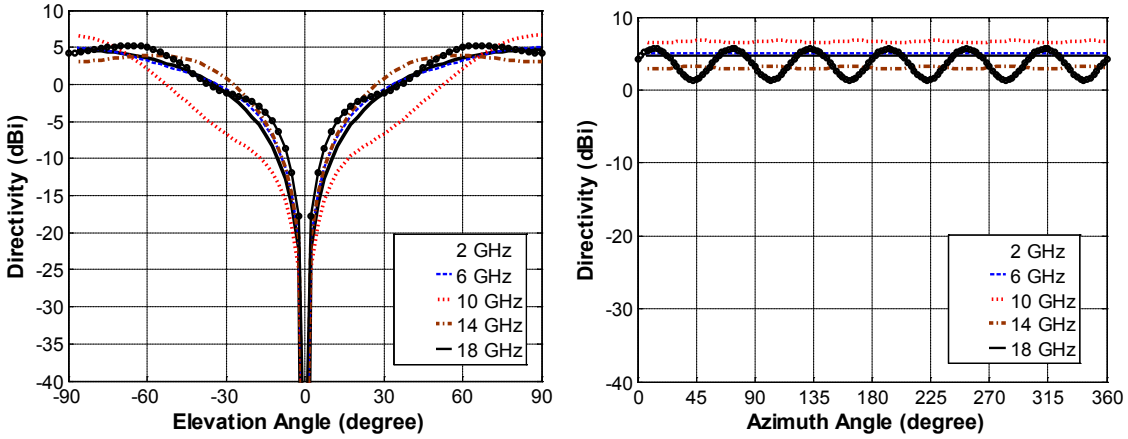
(a) (b)

Figure 6.17 Radiated pulse shape – (a) Measured link response ($|s_{21}|$) from two identical compact ultra-wideband antenna (two antennas are separated by 1.5 m) and (b) Simulated radiated E-field (simulation) at 1.5 m distance from antenna assuming an infinite ground plane. Input Gaussian pulse had 3-dB pulse width of 50 picoseconds. Commercial finite-difference time-domain (FDTD) code [*CST*, 2007] was used for simulation.



(a) (b)

Figure 6.18 Simulated realized gain and radiation efficiency versus frequency – (a) Simulated realized gain and (b) Radiation efficiency. Commercial method of moment (MoM) code [*FEKO*, 2007] was used for the simulations.

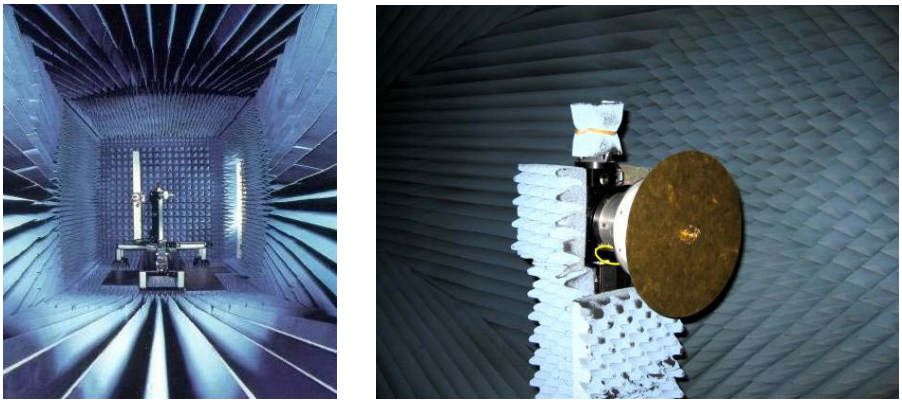


(a) (b)

Figure 6.19 Simulated directivity patterns for some selected frequencies – (a) Elevation patterns and (b) Azimuth patterns. Infinite ground plane is assumed. Commercial method of moment (MoM) code [*FEKO*, 2007] was used for the simulations.

Three dimensional radiation patterns of the designed antenna prototype with a finite-size circular ground plane (150 mm radius) were measured in the indoor near-field antenna range of the Virginia Tech Antenna Group [*VTAG*, 2008]. Pictures of the measurement setup are depicted in Figure 6.20. Two orthogonal electric near-fields were measured at a near-field region in the chamber. The far-field radiation patterns were obtained through the spherical near to far-field

transformation [*Hansen*, 1988]. The measured 3-D radiation patterns are shown at selected frequencies in Figure 6.21. Due to the finite size of the ground plane, the maximum radiation does not occur at the boresight ($\theta = 90°$). Similar to the simulated results in Figure 6.19, the measured radiation patterns are fairly consistent and omni-directional over the entire operating frequency. The complete set of the measured radiation patterns are in Appendix B.



(a) (b)

Figure 6.20 Pictures of measurement setup – (a) Indoor near-field antenna range [*VTAG*, 2008] and (b) Antenna under test mounted on a rotating-axis pole inside the anechoic chamber.

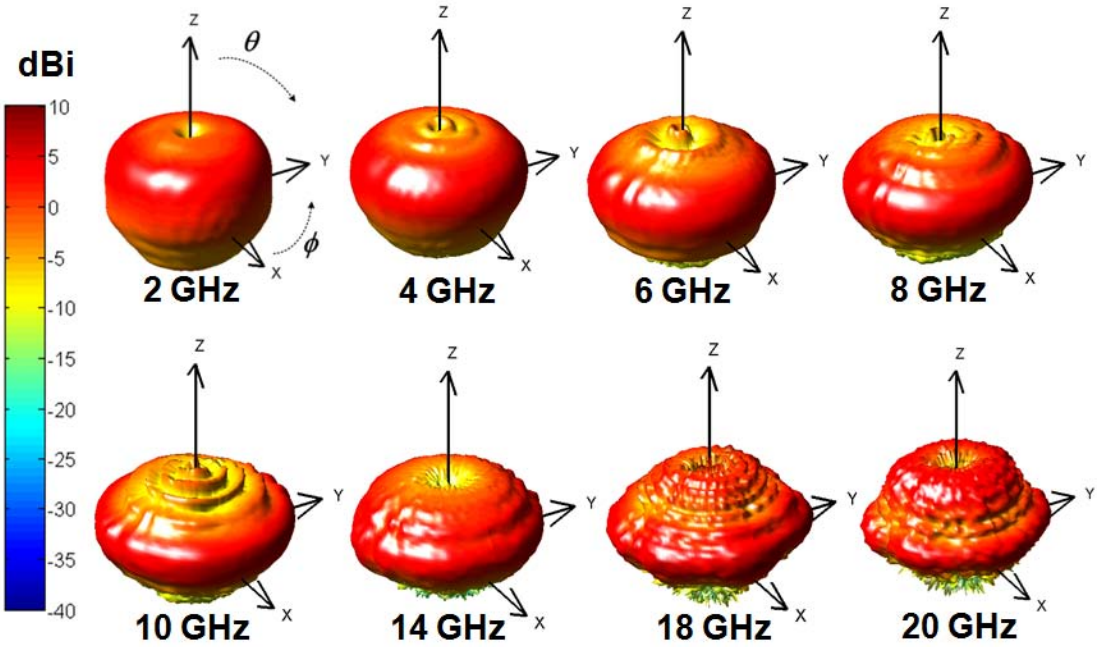


Figure 6.21 Measured 3-dimensional radiation patterns of the designed compact ultra-wideband antenna at some selected frequencies. Complete set of measured patterns are available in Appendix B.

Spherical modal power of the designed compact ultra-wideband antenna was evaluated by using (3.28) and (3.29), and shown in Figure 6.22. As expected from the relatively-constant radiation pattern and directivity over operational frequency range, the fundamental spherical TM_{01} mode is dominant over the entire operating frequency. Notice that the spherical TM_{03} mode is relatively strongly excited around 10 GHz, compared to other frequencies. Thus, the relatively strong excitation of the spherical TM_{03} mode partially explains why the radiation pattern around 10 GHz is a little bit directive in Figure 6.19a.

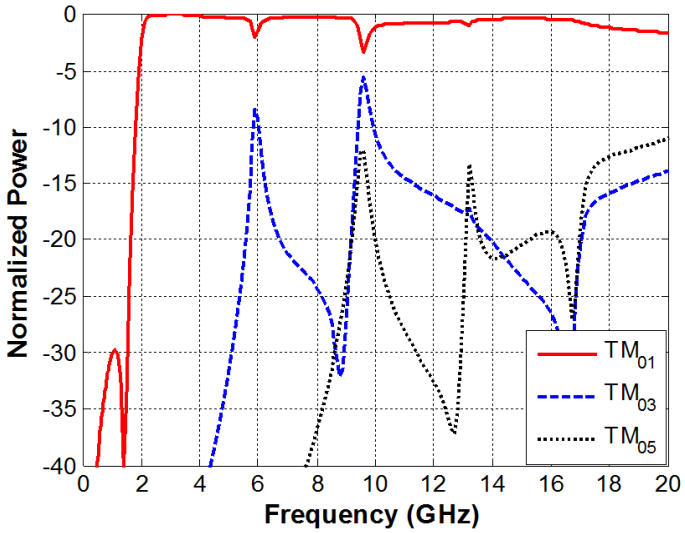
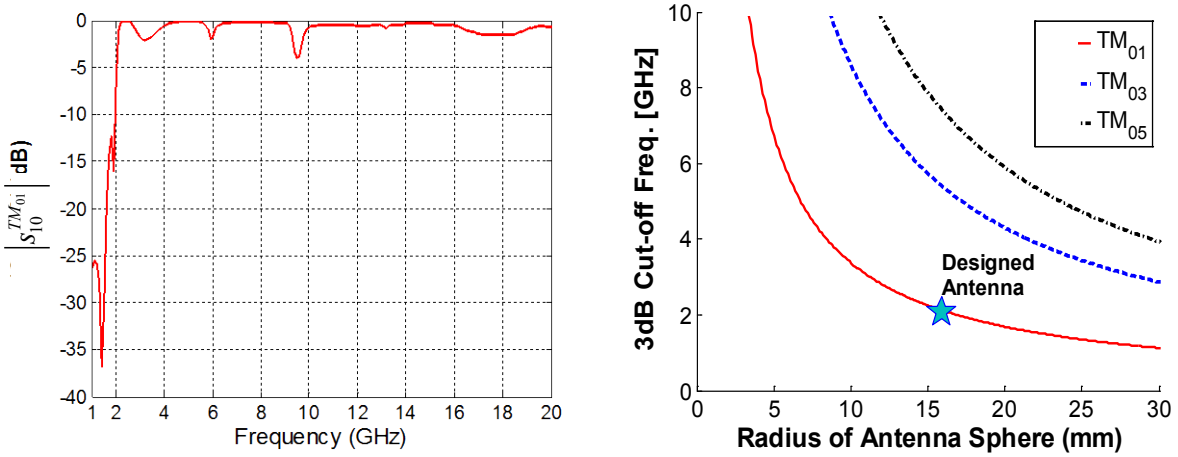


Figure 6.22 Normalized modal power versus frequency for some selected spherical TM modes. Simulated near-field scan data was used for the modal power computation.

As discussed in Chapter 4, the size and performance of an ultra-wideband antenna can be compared to the theoretical 3-dB cut-off frequency limit based on the S-parameter model representing radiating coupling to spherical modes in Figure 3.19. The computed $\left|s_{10}^{TM_{01}}\right|^2$ of the designed compact ultra-wideband antenna is plotted in Figure 6.23a. The sizes versus 3 dB cut-off frequencies for the various spherical TM modes are plotted in Figure 6.23b, compared with the 3 dB cut-off frequency of the designed antenna. The 3-dB cut-off frequency of the designed antenna is approximately 2.048 GHz and slightly lower than the limit curve of the spherical TM_{012} mode. However, this does not mean that the designed antenna broke the limit. As shown

in Figure 6.18b, the radiation efficiency of the antenna around the 3-dB cut-off frequency is approximately 97%. The actual limit curve corresponding to the radiation efficiency would be shifted down slightly. Therefore, the designed compact ultra-wideband antenna is very close to the fundamental limit, in terms of size and operational bandwidth.



(a) (b)

Figure 6.23 Comparison of 3-dB cut-off frequencies between the fundamental-limit theory and the designed compact ultra-wideband antenna – (a) $\left|s_{10}^{TM_{01}}\right|$ of the designed antenna from the method outlined in Chapter 4 and (b) Performance of the design antenna, compared to theoretical fundamental limits of 3 dB cut-off frequencies for various spherical TM_{0n} modes.

6.4 Strip-Loaded Antenna Design with Radome for Software-Defined Radio Applications

Software-defined radios (SDR) offer opportunities for a variety of applications in consumer electronics, public safety, and military [*J H Reed*, 2002]. However, there are practical design issues due to the wide operating frequency range from HF to C-band and beyond. The required operational bandwidth should be considered for analog functions (antennas and RF front ends), digital operations (baseband and software), and transition domains (ADC/DAC) during the implementation of a software-defined radio. Software-defined radios can be configured in several architectures to sustain a frequency-agile capability in wideband operation, but a common bottleneck of the various architectures is the antenna. The stated wide frequency range need not be covered instantaneously for some applications, but it is still very challenging to design a simple, rugged, highly-efficient broadband antenna with a compact footprint. In this

sense, excellent compact ultra-wideband antennas can be a good candidate for software-defined radio applications.

In this section, a compact antenna design for software-defined radio applications is introduced, based on the compact ultra-wideband antenna design in the previous section. The targeted software-defined platform for the antenna design is a direct conversion radio with a flexible radio frequency (RF) front end by Wireless@Virginia Tech. The RF front of the software-defined radio (SDR) uses a Motorola RFIC supporting a 100 MHz - 6 GHz operational frequency range with a 9 kHz to 20 MHz channel bandwidth.

6.4.1 Antenna Geometry

Based on the discussion of antenna design strategy in previous sections, the radius of the antenna sphere covering the lowest operating frequency (100 MHz) of the targeted software-defined radio (SDR) would be about 30 cm without the trade-off in the efficiency for transmission. To avoid this large size, a duplexer or an antenna switch may be used between the antenna and two separate transmitters to separate the operation frequency range (100 MHz – 6 GHz) into two bands below and above 450 MHz. The height of the antenna using a duplexer can be approximately 7.62 cm to meet the high-band needs. Below 450 MHz, an automatic tuner could be connected to the low-band side of the duplexer to provide the appropriate impedance match.

A scaled version of the compact ultra-wideband antenna designed in the previous section is considered as a candidate antenna. Rather than using wire arm loading above the cone feed, the new design in Figure 6.24 uses strip arms printed on the inner layer of the radome. Thus, the antenna structure is protected from potential damage.

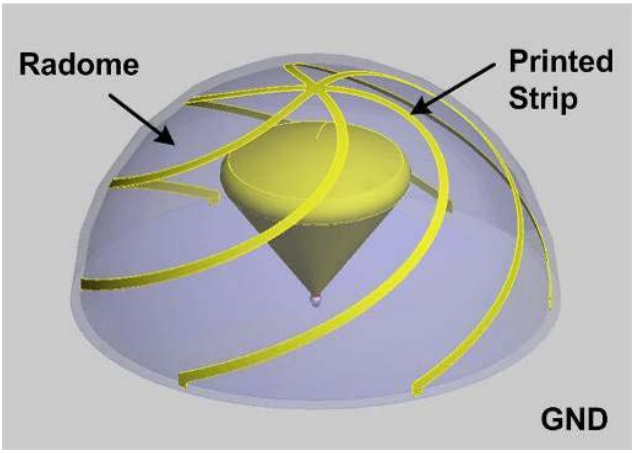
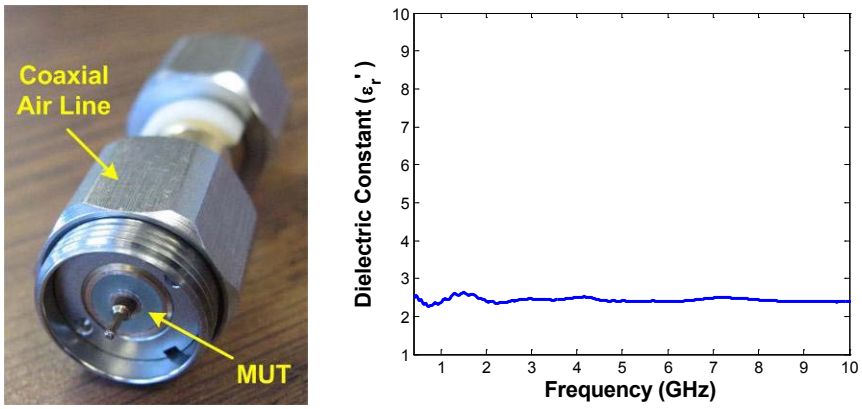


Figure 6.24 Overall geometry of the designed multi-functional compact antenna with the printed strip arms inside a radome for software-defined radio applications. Height of the antenna is approximately 7.62 cm over ground plane.

Because the designed antenna needs to cover a very large bandwidth, the material properties of the radome need to be relatively constant over the operational frequency. Particularly, the dielectric constant (ε_r') of the radome needs to be low and fairly constant due to the potential loading effects to the printed strip arms. If the dielectric constant is low, the effective dielectric constant of the shell radome is lower. Thus, the potential loading effect from the radome can be minimized. The slight loading effect of the radome can be compensated by changing the number of turns for the helical structure without any modification for antenna size. Thus, if the dielectric constant of the shell radome is relatively low, the main role of the radome can be to support/protect the antenna structure. Low loss is also generally preferred for a radome material, but loss would not a primary concern because it will not dominantly affect the equivalent inductance of the printed strip arms.

Rexolite® [*C-Lec Plastics*, 2012] would be an excellent radome material. Rexolite® is a popular material for military antenna designs due to its excellent mechanical and electrical characteristics. In addition, it is well known that the electrical characteristics of Rexolite® (ε_r' = 2.53 and $\tan\delta_e < 0.001$ at 10 GHz) are fairly constant up to a very high frequency (< 50 GHz). Unfortunately, Rexolite® is expensive

Acrylic is often suggested as an alternative due to its cheap price and similar properties to Rexolite®. For our antenna design, a hemispherical acrylic shell with a 7.62 cm radius and 0.95 cm is used as a radome. Since wide-band characterization data of Acrylic is not readily available, the Acrylic radome material was characterized with a reflection and transmission coefficient approach [*Nicolson and Ross*, 1970; *Weir*, 1974] using a 7-mm (2.76-mil) coaxial air line as shown in Figure 6.25a. Overall, the measured dielectric constant was approximately 2.5 up to 10 GHz, which is very similar to Rexolite®.



(a) (b)

Figure 6.25 Material characterization for acrylic-shell radome – (a) Picture of a 7-mm coaxial air line loaded with a material under test (MUT) for measurement setup and (b) Extracted dielectric constant in 0.4 – 10 GHz range from the measured scattering parameters based on reflection / transmission approach.

A picture of the designed antenna prototype is shown in Figure 6.26. The strip arms were painted on the inner surface of the Acrylic shell with a silver-ink paint. The cone feed structure in Figure 6.24 was replaced with multiple planar structures of vertical cross-section shape of the cone in order to avoid the precise 3-D cutting work on a lathe. Compared to the original solid cone feed structure, the equivalent cone feed with multiple planar plates may excite more higher-order spherical modes at the high operational frequencies. However, it is expected that the fundamental spherical mode still will remain dominant, even at these high frequencies.

If the rapid prototyping technology [*Chua et al.*, 2010] is applied, the radome with the strip arms can be easily printed in three dimensions with great precision.

Figure 6.26 Picture of the multi-functional compact antenna prototype for SDR applications.

6.4.2 Simulated and Measured Performance

The performance of the multi-functional compact antenna was evaluated through simulations using a commercial hybrid code (method of moment + finite element method) [*FEKO*, 2010] and measurements using a HP 8510C vector network analyzer (VNA).

The VSWR performance referenced to 50 Ω is shown in Figure 6.27. Both simulated and measured results show a 2:1 VSWR performance up to 6 GHz with similar variation, indicating the measured characteristics of the radome material (acrylic shell) satisfy the needs of the compact multi-functional antenna design for the targeted SDR platform. The measured lowest operating frequency was 433 MHz. Thus, the electrical size of the designed antenna is 0.11 wavelength at 433 MHz.

The plot of the computed realized gain versus frequency is shown in Figure 6.28. Compared to the compact ultra-wideband antenna design in the previous section, the realized gain curve fluctuates over the operational frequency. However, the fluctuation is roughly within ± 0.5 dB. No major radiation pattern distortion is expected in the operational frequency range.

In order to evaluate time-domain performance of the designed antenna, the impulse response of the designed was measured. The overall measurement setup is similar to that of the compact ultra-wideband antenna in the previous section. Two identical antennas were separated by 2 m. An equivalent Gaussian pulse with 50-ps pulse width was used as an input pulse to the designed antenna. As shown in Figure 6.29a, the measured radiated pulse has a very narrow pulse width with minor ringing. As we observed in other ultra-wideband antennas, there is always minor

ringing essentially caused by the finite-size aspect of antenna, which can be mitigated by resistive loading if necessary. The measured phase response of the radiated pulse in Figure 6.29b confirmed that the designed antenna have a linear phase response up to 20 GHz.

Thus, simulated and measured performance results indicate that the designed compact multi-functional antenna is a good candidate for both frequency and time-domain SDR [*Anderson and Reed*, 2006; *Nekoogar and Dowla*, 2009] applications.

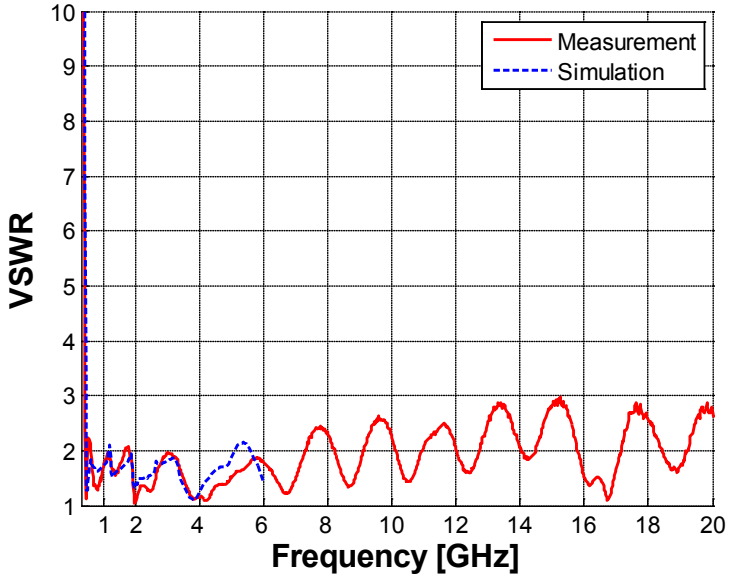


Figure 6.27 Measured and computed VSWR versus frequency for the designed multi-functional antenna for SDR applications (Z_0 = 50 Ω).

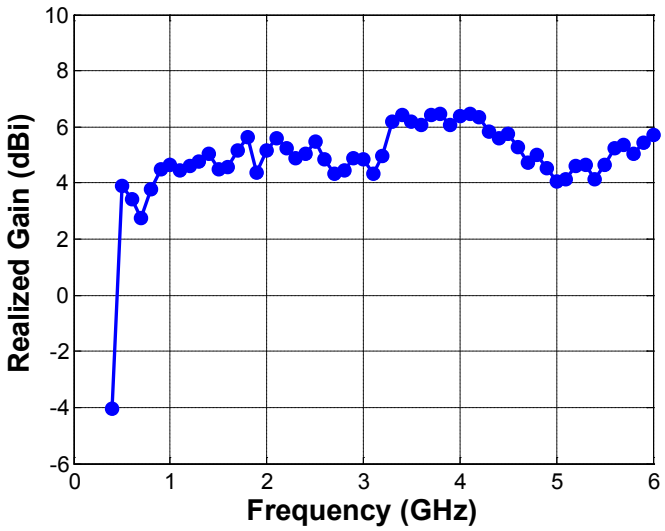
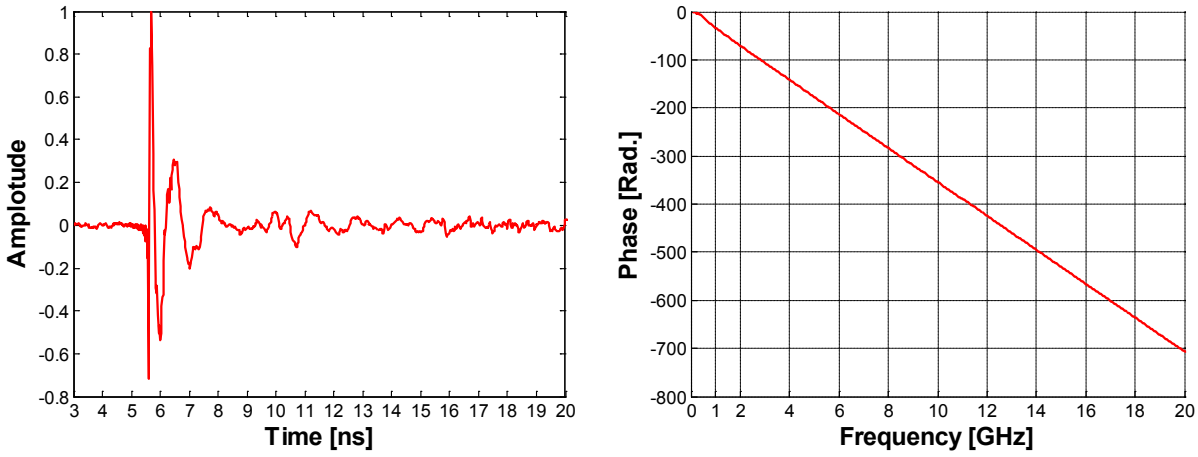


Figure 6.28 Computed realized gain (including return loss) versus frequency for the designed compact multi-functional antenna for SDR applications.



(a) (b)

Figure 6.29 Measured time-domain performance of the designed multi-functional antenna for SDR applications – (a) Radiated pulse for Gaussian input pulse with 3-dB pulse width of 50 pico-second and (b) Phase response versus frequency.

6.5 Chapter Summary

Challenges and trade-offs in a compact ultra-wideband (UWB) antenna design were addressed for canonical types of antennas (resonant, frequency-independent, traveling-wave, and wideband types) after reviewing ultra-wideband technology briefly. For both I-UWB and MB-OFDM UWB applications, the most challenging design issue was a size miniaturization without sacrificing operational bandwidth and efficiency, though the I-UWB technology requires a linear phase response.

A general strategy to design an ultra-wideband antenna close to the theoretical size and performance limit was developed. The design strategy was explained through an evolution process of antenna design. Two example antenna designs for ultra-wideband and software-defined radio applications were introduced and characterized in both the frequency and time domains.

The first example was a compact ultra-wideband antenna (CUA). The designed antenna had an approximate 10:1 impedance bandwidth with fairly-constant radiation patterns and a 5-dBi realized gain over the entire band. Using modal power analysis, the designed antenna dominantly excites the fundamental spherical mode, which explains why the designed antenna has a

relatively constant radiation pattern over the operational frequency range. The size of the designed ultra-wideband antenna was 1/8 wavelength at the lower bound of the band, which is two times smaller than conventional quarter-wavelength ultra-wideband antennas. The radiated pulse of the designed antenna showed a doublet waveform with minor ringing. It was also shown that the designed antenna is very close to the theoretical limit of 3-dB cut-off frequency.

The second example was based on a scaled version of the CUA design for software-defined radio (SDR) applications. A direct conversion radio with flexible radio frequency (RF) front end by Wireless@Virginia Tech was assumed as a target SDR system supporting a 100 MHz - 6 GHz operational frequency range. Based on the developed general antenna design strategy, the operational frequency range was separated into two bands, one below and one above 450 MHz in order to have a compact antenna size. The band separation could be achieved by using a duplexer or a switch. When the antenna is operating below 450 MHz, an automatic tuner could be connected to the low-band side of the duplexer to provide the appropriate impedance match. Above 450 MHz, the antenna provides an instantaneous bandwidth. In this sense, the designed antenna can be considered as a multi-functional antenna. Unlike the CUA, the new antenna design for SDR applications has a shell radome to protect the antenna. The helical strip arms were printed with a silver-ink paint on the inner layer of the radome. The lowest operating frequency was measured as 433 MHz where the new antenna design has the electrical size of 0.11 wavelength. Simulated and measured results in the frequency and time domains showed that the strip-loaded compact antenna with a shell radome can be a good candidate antenna for both frequency and time-domain SDR applications.

The both designed antennas meet normal aerodynamics requirements due to the hemispherical shape. Because the antenna requires only a few small connection points between the antenna and ground plane, the antenna can be easily installed and integrated on a curved surface. Potential applications of the antennas include home wireless systems requiring a high data transferring rate, un-manned air vehicles (UAVs) requiring a small antenna footprint on a curved surface, frequency agile systems requiring a broadband operation, and short range pulsed radar systems requiring a short pulse radiation performance and wide angular coverage.

Chapter 7

Conclusions

This dissertation provides integrated knowledge and understanding of antenna radiation physics and the fundamental limits on antenna size and performance. Antenna design examples were provided to demonstrate the introduced concepts. This chapter summarizes this dissertation and the original contributions. Suggestions for future work are also outlined.

7.1 Summary

The dissertation research was performed in multiple stages; starting with examination of fundamental antenna physics and theoretical limits, and then applications to challenging antenna design problems.

In Chapter 2, antenna radiation models were reviewed to understand radiation physics and energy storage mechanism. Investigation of alternative viewpoints using wave impedance, power, Poynting vector, and energy velocity provided deeper understanding of the antenna radiation process. The investigation indicated that all antenna fields can contribute to both radiation and energy storage. In addition, it was shown that the energy velocity of radiation is not constant. These observations are contradictory to the commonly accepted assumptions that only angular fields with $1/r$ variation contribute to far-field radiation and the propagation speed of fields in free space is constant. Evidence and proofs were provided to show that these widely-accepted facts are not necessary true.

In Chapter 3, modeling approaches for the antenna transfer function were investigated as a tool for the main study of this dissertation. SEM-based antenna transfer functions using effective height [*Licul*, 2004] were reviewed with examples including resonant, frequency-independent, and ultra-wideband antennas. An alternative singularity-expansion-method (SEM)-based antenna transfer function model using the spherical wave expansion (SWE) approach was introduced. The alternative model can have the same advantages that the previous SEM-based antenna transfer function model using the effective height. In addition, the alternative approach can include near-field characteristics of an antenna as well. The usefulness of the alternative model was demonstrated using resonant, frequency-independent, and ultra-wideband antenna examples with simulated near-field scan data.

In Chapter 4, assumptions and errors in the classic fundamental-limit theory were explained, based on the observations and understandings obtained from the antenna radiation physics and antenna transfer function analysis. It has been shown that the frequency-domain approach of Chu [*Chu*, 1948], Collin and Rothschild [*R Collin and Rothschild*, 1964], and McLean [*McLean*, 1996] is essentially based on a far-field approximation extrapolated to the near-field or the assumption that radiating energy travels at the speed of light all the way from the antenna sphere to the far-field. Grimes' the time-domain approach [*Grimes and Grimes*, 1999] provided some insight in terms of the existence of an excess delay in antenna radiation process and a new approach using a non-negative energy condition to obtain the minimum amount of non-radiating energy.

Based on a recent contribution of Davis *et al* [*Davis et al.*, 2011], the classic limit theory was corrected and extended. It was shown that the corrected minimum radiation Q is consistent in both field analysis and Chu's equivalent circuit approaches. The analysis of the static energy of a charge dipole indicated that the minimum non-radiating energy comes from the electric fields bound to the charges and all decoupled energy can contribute to far-field radiation. Our investigation of the minimum radiation Q was extended to consider the spherical $TM_{1,1}$, $TM_{-1,1}$, $TE_{1,1}$, and $TE_{-1,1}$ modes (circular polarization) and combined spherical modes with arbitrary-polarization. In addition, classical fundamental-limit theory on antenna size was interpreted from an ultra-wideband antenna perspective. The frequency response of the extended Chu's equivalent model for spherical TM modes suggests the concept of an ideal antenna, with the entire

frequency response described with only two complex poles and an entire function. It was shown that the ideal antenna naturally has ultra-wideband characteristics.

The size limitations on ultra-wideband antennas in terms of pulse width and 3-dB cut-off frequency were developed and its usefulness was demonstrated with a bicone antenna example. The 3-dB cut-off frequency criterion can also be used to determine which antenna is more suitable in a frequency range of interest for resonant, ultra-wideband, or frequency independent antennas. The demonstrated concepts and approaches can be generalized for spherical $TM_{m,n}$, $TE_{m,n}$, and combined mode cases.

In Chapter 5, a method to reduce the near-field electromagnetic energy around a cellular phone was presented to mitigate the interaction between a cellular phone and a hearing aid. Examining near-field terms in the field equations showed that current slope discontinuities cause high near field intensities. Fundamental limit theory of antennas indicates that the near-field interaction problems can be reduced using low-Q antennas, such as an ultra-wideband (UWB) antenna, in addition to reducing sharp current discontinuities. The low-Q antenna approach to reduce near-field strength was demonstrated through both simulations and measurements. Simulations showed that the peak electric and magnetic near-field amplitudes are 7 dB and 4 dB, respectively, lower for the planar UWB fat dipole (low Q) compared to the narrow-band dipole (high Q). For more practical cellular-phone antennas, the simulation and measurement results for the dual-band, planar inverted-F (900 and 1880 MHz, high Q) and ultra-wideband (900 MHz – 10 GHz, low-Q) test antennas mounted on mock phones showed that the peak electric and magnetic near-field strength of low-Q test antenna is reduced by at least 5 dB and 4 dB, respectively. This near-field strength reduction was achieved without sacrificing far-field performance.

In the presence of a human head, simulation results also showed that the radiation efficiency of the ultra-wideband test antenna is better than the dual-band PIFA antenna, using both Visible human and SAM head models. Furthermore, the simulation and measurement results of specific absorption rate (SAR) with the SAM phantom head model revealed that the spoon UWB (low-Q) antenna showed more than 50% reduction of both peak and spatial-average SAR (1 g) values, compared to the dual-band PIFA (high-Q). The low-Q antenna approach to reduce near-field strength and spread the current distribution to reduce discontinuity peaks can be also applied to a

variety of other applications, including safety issues of RF energy exposure, such as high power radiating systems and medical electronic systems, efficiency of an antenna system, and co-site interference issues. Alternatively, a high-Q antenna approach can be effective when the range of communication link is critical, such as the link between a scanner and RFID devices.

In Chapter 6, a general strategy to design an ultra-wideband antenna that is close to the theoretical size and performance limit was developed. The design strategy was explained through an evolution of antenna designs, based on the concepts of an ideal antenna. Two example antenna designs for ultra-wideband and software-defined radio applications were introduced and characterized in both the frequency and time domains. The first example was the compact ultra-wideband antenna (CUA). The designed antenna had approximately 10:1 impedance bandwidth with fairly constant radiation patterns and 5-dBi realized gain over the entire band. Using modal power analysis, the designed antenna dominantly excites the fundamental spherical modes, which explains why the designed antenna has relatively constant radiation pattern over the operational frequency range. The size of the designed ultra-wideband antenna was 1/8 wavelength at the lower bound of the band, which is two times smaller than conventional quarter-wavelength ultra-wideband antennas. The radiated pulse of the designed antenna showed a doublet waveform with a minor ringing. It was also shown that the designed antenna is very close to the theoretical limit of 3-dB cut-off frequency.

The CUA design also volved for software-defined radio (SDR) applications. A direct conversion radio with flexible radio frequency (RF) front end by Wireless@Vinitia Tech was assumed as the target SDR system to support a 100 MHz - 6 GHz operational frequency range. Based on the developed general antenna design strategy, the operational frequency range was separated into two bands, one below and one above 450 MHz in order to have compact antenna size. The band separation could be achieved using a duplexer or a switch. The new antenna design for SDR applications has a shell radome added to the original CUA to protect the antenna. Helical strip arms were printed with a silver-ink paint on inner layer of the radome. The lowest operating frequency was measured as 433 MHz, where the new antenna design has the electrical size of 0.11 wavelength. Simulated and measured results in frequency and time domains showed that the strip-loaded compact antenna with a shell radome can be a good candidate antenna for both frequency and time-domain SDR applications.

Both designed antennas meet normal aerodynamics requirements due to the hemispherical shape. Because the antenna requires only a few small connection points between the antenna and ground plane, the antenna can be easily installed and integrated on a curved surface. Potential applications of the antennas include home wireless systems requiring a high data transfer rate, un-manned air vehicles (UAVs) requiring a small antenna footprint on a curved surface, and frequency agile systems requiring a broadband operation, and short range pulsed radar systems requiring a short pulse radiation performance and wide angular coverage.

These results lead to some unique contributions, which are outlined in the next section.

7.2 Contributions

This dissertation provides the integrated knowledge and understandings on antenna radiation physics, fundamental limits on antenna size and performance, and challenging real-world antenna design problems. The specific original contributions are detailed in this section.

The first contribution of the research is the correction and extension of classic fundamental limit theory on antenna size, and is based on the recent work of Davis *et al* [*Davis et al.*, 2011] and the observations and understanding from the antenna radiation physics and antenna transfer function analysis (see Chapters 2 and 3). It was shown that the result of the corrected minimum radiation Q is consistent with both field analysis and Chu's equivalent circuit approache. Through static energy analysis using charge dipole, the minimum non-radiating energy comes from the fields bound to charges (static energy) and all decoupled energy can contribute to far-field radiation. Thus, it was pointed out that the corrected minimum radiation Q can also be conveniently obtained from a low-frequency approximation. Using the determined amount of radiating and non-radiating energy, it was demonstrated that both radial and angular fields contribute to radiating and non-radiating energies. The investigation of the minimum radiation Q was extended to consider the spherical $TM_{1,1}$, $TM_{-1,1}$, $TE_{1,1}$ and $TE_{-1,1}$ modes (circular polarization) and combined spherical modes with arbitrary polarization. The size limitations of ultra-wideband antennas in terms of pulse width and 3-dB cut-off frequency were developed and its usefulness was demonstrated with a bicone antenna example. The new radiation-Q limits and the alternative 3-dB cut-off frequency limits provide antenna designers with a simple guide to estimate the minimum required antenna size for a given bandwidth and efficiency, thereby avoiding the fruitless search for an impractical antenna design and specification.

The second contribution is the demonstration of the low-Q antenna approach to mitigate near-field interaction issues, based on fundamental-limit theory on antenna size (see Chapter 5). Further investigation on near-field theory showed that antenna current slope discontinuities cause high near field intensities. For the interaction problem between a cell phone and a hearing aid, a low-Q test antenna mounted on mock cell phones showed that the peak electric and magnetic near-field strength is reduced by at least 5 dB and 4 dB, respectively, compared to that of high-Q antenna. Near-field strength reduction was achieved without sacrificing far-field performance. In terms of interactions between a cell phone and a human head, the simulated and measured specific absorption rate (SAR) with SAM phantom head model revealed that the low-Q antenna shows more than 50% reduction of both peak and spatial-average SAR (1 g) values, compared to the high-Q antenna. The same low-Q antenna solution can be applied to various near-field interaction problems including safety issues of RF energy exposure, such as high power radiating systems and medical electronic systems, efficiency of the antenna system, and co-site interference issues.

The third contribution is design of a compact ultra-wideband antenna (CUA) close to the size and performance of theoretical limit, based on the developed general antenna design strategy for an ultra-wideband antenna. The CUA excited the dominant fundamental spherical mode and had approximately 10:1 impedance bandwidth with fairly constant radiation patterns and 5-dBi realized gain over the entire band (see Chapter 6). The size of the CUA was 1/8 wavelength at the lower end of the band, which is half the size of conventional quarter-wavelength ultra-wideband antennas. The radiated pulse of the designed antenna showed a doublet waveform with minor ringing, which suggests that the CUA can be a good candidate for I-UWB as well as frequency-domain applications. It was also shown that the designed antenna is very close to the theoretical limit of 3-dB cut-off frequency. The CUA design was evolved for software-defined radio (SDR) applications. The developed general antenna design strategy was used to determine the optimum antenna size, frequency coverage, and number of operational band. The new CUA design was integrated with a shell radome to protect the antenna. A part of the CUA was printed with a silver-ink paint on inner layer of the radome. The lowest operating frequency was measured as 433 MHz where the new antenna design has the electrical size of 0.11 wavelength. Simulated and measured results in frequency and time domains showed that the new CUA can

also be a good candidate antenna for both frequency and time-domain SDR applications. Both CUA designs can find a variety of applications, including home wireless systems requiring a high data transferring rate, un-manned air vehicles (UAVs) requiring a small antenna footprint on a curved surface, and frequency agile systems requiring a broadband operation, and short range pulsed radar systems requiring a short pulse radiation performance and wide angular coverage. In addition, the developed general antenna design strategy for an ultra-wideband antenna can be applied to obtain other ultra-wideband antenna designs and the development of other antenna design strategies.

Some of the work in this dissertation was published in journals and magazines [*Davis et al.*, 2011; *T. Yang et al.*, 2009; *Yang et al.*, 2008; *Yang et al.*, 2004; *Yang et al.*, 2011], presented at multiple national conferences, and also recognized through multiple national and international paper and research-excellence competitions.

7.3 Suggested Future Work

The provided integrated knowledge and understandings on antenna physics and size limit through the presented doctoral work can open research opportunities in various areas, as now described.

In the presented work, time-harmonic solutions for ideal dipoles were used to understand the antenna radiation physics and energy storage mechanism. However, even deeper understanding of the radiation process through transient solutions. In particular, better understanding of the energy storage mechanism may provide a clue to an antenna design below the size limit of classic limit theory. For example, the impedance modulation approach through a switching has been explored in the literature to break the classic limit theory [*Manteghi*, 2009; *Merenda*, 2006; 2008; *Xiaojing and Wang*, 2007], but it has been known that the precise control of switching speed is key to mitigate spurious radiation to improve radiation efficiency. Because the excess delay can be seen as a required time to store the energy, it seems that the energy storage mechanism is closely related to the required switching speed in the impedance or current modulation for better radiation efficiency. In our opinion, we reached the modeling limit of Maxwell's equations for further investigation on the antenna radiation mechanism. An alternative modeling tool could be a quantum electrodynamics to overcome the macroscopic

nature of Maxwell's equation and to obtain a solution for improving the efficiency of the impedance or current modulation approach.

The decomposed, singularity-expansion-method (SEM)-based antenna transfer-function model was introduced in Chapter 3. The model was used to understand the correlation between antenna physical structure and performance parameters for various types of antennas including resonant, frequency-independent, and ultra-wideband. In addition, the model was also used to define the size and performance limits of frequency-independent and ultra-wideband antennas. Unlike the previous models for far-field antenna characteristics, the introduced model can model antenna near-field characteristics as well. However, in this dissertation, the investigation was carried out only through simulation results. Thus, the introduced SEM-based antenna transfer function model decomposed with the spherical wave expansion (SWE) method can be further refined and verified through experiments with actual antenna designs. The advance in computing technology and fast measurement with a multi-probe near-field scanning system makes the decomposed, SEM-based antenna transfer function more attractive and opens other related research opportunities, such as near-field communication, near-field interactions in multi-input multi-output (MIMO) antennas, radar target detection, and tumor detection.

Although an antenna design strategy to design an antenna that is close to the theoretical size and performance limits was presented and demonstrated through a compact ultra-wideband antenna design example in this work, challenging real-world antenna design problems still beg for a systematic approach to antenna element design for given size and performance specifications. Because characteristic-mode solutions [*Garbacz and Turpin*, 1971; *R. Harrington and Mautz*, 1971] are directly correlated to the impedance matrix of physical antenna structure, fundamental-limit theory using the characteristic-mode analysis may provide a clue on how to obtain a systematic antenna design approach.

The direct relationship between radiation Q and the amount of non-radiating energy not only suggests a low-Q antenna approach to mitigate near-field interaction problems, but also suggests a high-Q antenna approach to extend the communication range or increase efficiency of wireless power transfer. Although these benefits are expected for the high-Q antenna approach, there can be sensitivity and loss increase issues due to the high-Q nature of the antenna. Thus, there might

be an optimum radiation Q to meet the required range specification and to minimize the loss and sensitivity concerns.

The presented work on fundamental-limit theory and radiation physics focused on transmitting antenna problems. As a result, operational bandwidth, size and radiation efficiency were the main design parameters of the presented work. However, signal-to-noise bandwidth is the primary performance parameter for receiving antennas. Thus, investigation of fundamental-limit theory and antenna receiving physics can be an excellent research opportunity. Combining limit theories for transmitting and receiving antenna problems may help to identify new size and performance trade-offs for a communication link.

Non-radiating energy is ultimately limits the antenna size and performance. However, the figure of merit for the limit basically focused on far-field radiation performance. For near-field communication, far-field radiation is not a required performance criterion. In addition, the non-radiating energy may increase performance and reduce antenna size. In this sense, one can develop fundamental limit theory on antenna size and performance for near-field communications.

References

Ahlbom, A., U. Bergqvist, J. Bernhardt, J. Cesarini, M. Grandolfo, M. Hietanen, A. Mckinlay, M. Repacholi, D. Sliney, and J. Stolwijk (1998), Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz). International Commission on Non-Ionizing Radiation Protection, *Health Phys*, *74*(4), 494-522.

Anderson, C. R., and J. H. Reed (2006), A software-defined ultra-wideband transceiver testbed for communications, ranging, and imaging.

ANSI (2006), American National Standard for Method of Measurement of Compatibility between Wireless Communications Devices and Hearing Aids (ANSI C63.19-2006), edited, American National Standards Institute, New York.

ANSI (2011), American National Standard for Method of Measurement of Compatibility between Wireless Communications Devices and Hearing Aids (ANSI C63.19-2011), edited, American National Standards Institute.

ANSI/IEEE (2003), IEEE Recommended Practice for Determining the Peak Spatial-Average Specific Absorption Rate (SAR) in the Human Head From Wireless Communications Devices: Measurement Techniques, *IEEE Std 1528-2003*, 0_1-149.

ANSI/IEEE (2006), IEEE Standard for Safety Levels With Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz, *IEEE Std C95.1-2005 (Revision of IEEE Std C95.1-1991)*, 0_1-238.

Arlinghaus, S. L. (1994), *Practical handbook of curve fitting*, 249 p. pp., CRC Press, Boca Raton, Fla.

Bahai, A. R. S., B. R. Saltzberg, and M. Ergen (2004), *Multi-carrier digital communications : theory and applications of OFDM*, 2nd ed., xix, 411 p. pp., Springer, New York.

Balanis, C. A. (2005), *Antenna theory : analysis and design*, 3rd ed., xvii, 1117 p. pp., John Wiley, Hoboken, NJ.

Barrett, T. W., and V. Vienna (2001), Technical Features History of Ultra Wideband Communications and Radar: Part I, UWB Communications, *Microwave journal*.

Baum, C. E. (1971), On the singularity expansion method for the solution of electromagnetic interaction problems*Rep.*, DTIC Document.

Baum, C. E. (1976), The singularity expansion method, *Transient Electromagnetic Fields*, 129-179.

Baum, C. E. (2012), *The Singularity Expansion Method in Electromagnetics: A Summary Survey and Open Questions*, 104 pp., Lulu Enterprise, Inc.

Baum, C. E., E. G. Farr, and D. V. Giri (1999), Review of impulse-radiating antennas, *Review of Radio Science*, 403-439.

Baum, C. E., E. J. Rothwell, K. M. Chen, and D. P. Nyquist (1991), The singularity expansion method and its application to target identification, *Proceedings of the IEEE*, *79*(10), 1481-1492.

Berger, H. S. (2001), Hearing aid and cellular telephone compatibility: working toward solutions, *Journal of the American Academy of Audiology*, *12*(6), 309.

Best, S. R. (2004a), The radiation properties of electrically small folded spherical helix antennas, *Antennas and Propagation, IEEE Transactions on*, *52*(4), 953-960.

Best, S. R. (2004b), The Foster reactance theorem and quality factor for antennas, *Antennas and Wireless Propagation Letters, IEEE*, *3*(1), 306-309.

Bolomey, J. C., B. J. Cown, G. Fine, L. Jofre, M. Mostafavi, D. Picard, J. P. Estrada, P. G. Friederich, and F. L. Cain (1988), Rapid near-field antenna testing via arrays of modulated scattering probes, *Antennas and Propagation, IEEE Transactions on*, *36*(6), 804-814.

Brillouin, L. o. (1960), *Wave propagation and group velocity*, 154 p. pp., Academic Press, New York,.

Budko, N. V. (2009), Observation of locally negative velocity of the electromagnetic field in free space, *Physical review letters*, *102*(2), 20401.

Budko, N. V. (2010), Superluminal, subluminal, and negative velocities in free-space electromagnetic propagation, *Arxiv preprint arXiv:1006.5576*.

C-Lec Plastics, I. (2012), Characteristics of Rexolite, edited.

Capps, C. (2001), Near field or far field?, *EDN (Electrical Design News)*, *46*(18), 95-102.

Caputa, K., M. Stuchly, M. Skopec, H. Bassen, P. Ruggera, and M. Kanda (2000), Evaluation of electromagnetic interference from a cellular telephone with a hearing aid, *Microwave Theory and Techniques, IEEE Transactions on*, *48*(11), 2148-2154.

Cardoso, J. C., and A. Safaai-Jazi (1993), Spherical helical antenna with circular polarisation over a broad beam, *Electronics Letters*, *29*(4), 325-326.

CENELEC (1995), CENELEC ENV 50166-2: Human Exposure to Electromagnetic Fields High Frequency : (10 kHz to 300 GHz), edited, European Committee for Electrotechnical Standardization.

Chavannes, N., R. Tay, N. Nikoloski, and N. Kuster (2003), Suitability of FDTD-based TCAD tools RF design of mobile phones, *Antennas and Propagation Magazine, IEEE*, *45*(6), 52-66.

Chen, Z. N., A. Cai, T. S. P. See, X. Qing, and M. Y. W. Chia (2006), Small planar UWB antennas in proximity of the human head, *Microwave Theory and Techniques, IEEE Transactions on*, *54*(4), 1846-1857.

Cheng, D. K. (1989), *Field and wave electromagnetics*, xvi, 703 p. pp., Addison-Wesley, Reading, Mass.

Chu, L. J. (1948), Physical Limitations of Omni-Directional Antennas, *Journal of applied physics*, *19*(12), 1163-1175.

Chua, C. K., K. F. Leong, and C. S. Lim (2010), *Rapid prototyping : principles and applications*, 3rd ed., xxv, 512 p. pp., World Scientific, Singapore ; Hackensack, NJ.

Clark, J., and A. Safaai-Jazi (2004), Multifilar hemispherical helical antennas, paper presented at Antennas and Propagation Society International Symposium, 2004. IEEE, 20-25 June 2004.

Cleveland, R. F., R. E. Fields, and J. L. Ulcek (1999), Questions and answers about biological effects and potential hazards of radiofrequency electromagnetic fields.

Collin, R., and S. Rothschild (1964), Evaluation of antenna Q, *Antennas and Propagation, IEEE Transactions on*, *12*(1), 23-27.

Collin, R. E. (2001), *Foundations for microwave engineering*, 2nd ed., xix, 924 p. pp., IEEE Press, New York.

Collin, R. E., and F. J. Zucker (1969), *Antenna theory*, McGraw-Hill, New York,.

CST (2007), CST Microwave Studio, edited, CST.

Daniel, C., F. S. Wood, and J. W. Gorman (1999), *Fitting equations to data : computer analysis of multifactor data, second edition*, Wiley classics library ed., xviii, 458 p. pp., Wiley, New York.

Davis, W., T. Yang, E. Caswell, and W. Stutzman (2011), Fundamental limits on antenna size: a new limit, *Microwaves, Antennas & Propagation, IET*, *5*(11), 1297-1302.

de Salles, A. A. A. (1999), Biological effects of microwave and RF, paper presented at Microwave and Optoelectronics Conference, 1999. SBMO/IEEE MTT-S, APS and LEOS - IMOC '99. International, 1999.

Dillon, H. (2001), *Hearing aids*, xviii, 504 p. pp., Thieme Medical Publishers, Incorporated, New York.

Douglas Lynch, A. E., B. Gerard Genello, and C. Michael Wicks (2007), UWB Perimeter Surveillance, *Aerospace and Electronic Systems Magazine, IEEE*, *22*(1), 8-10.

El-Gabaly, A. M., and C. E. Saavedra (2011), A Quadrature Pulse Generator for Short-Range UWB Vehicular Radar Applications Using a Pulsed Oscillator and a Variable Attenuator, *Circuits and Systems I: Regular Papers, IEEE Transactions on*, *58*(10), 2285-2295.

Elliott, R. S. (2003), *Antenna theory and design*, Revised ed., xxi, 594 p. pp., John Wiley & Sons, Hoboken, N.J.

FCC (2002), First report and order, revision of Part 15 of Commission's rule regarding ultra-wideband transmission system FCC 02-48, *Apr*, *22*, 2.

FCC (2012), Ultra-wideband Operation, in *U.S. 47 C.F.R Part 15 Subpart F*, edited by FCC.

FEKO (2006), FEKO Suite v5.2, edited, EM Software & Systems.

FEKO (2007), FEKO Suite v5.3, edited, EM Software & Systems.

FEKO (2010), FEKO Suite v5.5, edited, EM Software & Systems

Feynman, R. P. (1961), *Quantum electrodynamics; a lecture note and reprint volume*, 198 p. pp., W.A. Benjamin, New York,.

Foerster, J. R. (2002), The performance of a direct-sequence spread ultrawideband system in the presence of multipath, narrowband interference, and multiuser interference, paper presented at Ultra Wideband Systems and Technologies, 2002. Digest of Papers. 2002 IEEE Conference on, 2002.

Garbacz, R., and R. Turpin (1971), A generalized expansion for radiated and scattered fields, *Antennas and Propagation, IEEE Transactions on*, *19*(3), 348-358.

Geyi, W. (2003), A method for the evaluation of small antenna Q, *Antennas and Propagation, IEEE Transactions on*, *51*(8), 2124-2129.

Ghavami, N., G. Tiberi, D. J. Edwards, and A. Monorchio (2012), UWB Microwave Imaging of Objects With Canonical Shape, *Antennas and Propagation, IEEE Transactions on*, *60*(1), 231-239.

Gibson, P. J. (1979), The Vivaldi Aerial, paper presented at Microwave Conference, 1979. 9th European, 17-20 Sept. 1979.

Giri, D. V. (2004), *High-power electromagnetic radiators : nonlethal weapons and other applications*, xiii, 198 p. pp., Harvard University Press, Cambridge, Mass.

Goldberg, R. B. (1996), Literature resources for understanding biological effects of EM fields, *Engineering in Medicine and Biology Magazine, IEEE*, *15*(4), 96-101.

Golub, G. H., and C. F. Van Loan (1996), *Matrix computations*, 3rd ed., xxvii, 694 p. pp., Johns Hopkins University Press, Baltimore.

Grimes, D. M., and C. A. Grimes (1999), Radiation Q of dipole-generated fields, *Radio Science*, *34*(2), 281-296.

Guyton, A. C., and J. E. Hall (2006), *Textbook of medical physiology*, 11th ed., xxxv, 1116 p. pp., Elsevier Saunders, Philadelphia.

Hansen, J. E. (1988), *Spherical near-field antenna measurements*, xiv, 387 p. pp., P. Peregrinus on behalf of the Institution of Electrical Engineers, London, U.K.

Harrington, R. (1993), Field Computation by Moment Methods, *IEEE/OUP*.

Harrington, R., and J. Mautz (1971), Theory of characteristic modes for conducting bodies, *Antennas and Propagation, IEEE Transactions on*, *19*(5), 622-628.

Harrington, R. F. (2001), *Time-harmonic electromagnetic fields*, xiii, 480 p. pp., IEEE Press : Wiley-Interscience, New York.

Heaviside, O. (1894), *Electromagnetic theory*, "The Electrician" printing and publishing company, London,.

Heaviside, O. (1925), *Electromagnetic theory*, E. Benn, London,.

Hertel, T. W., and G. S. Smith (2003), On the dispersive properties of the conical spiral antenna and its use for pulsed radiation, *Antennas and Propagation, IEEE Transactions on*, *51*(7), 1426-1433.

Hertz, H. (1893), *Electric waves; being researches on the propagation of electric action with finite velocity through space*, xv p., 1 l., 278 p., 271 l. pp., and New York, Macmillan and co., London,.

Hertz, H. (1962), *Electric waves, being researches on the propagation of electric action with finite velocity through space*, 278 p. pp., Dover Publications, New York,.

Hombach, V., K. Meier, M. Burkhardt, E. Kuhn, and N. Kuster (1996), The dependence of EM energy absorption upon human head modeling at 900 MHz, *Microwave Theory and Techniques, IEEE Transactions on*, *44*(10), 1865-1873.

Hua, Y., and T. K. Sarkar (1990), Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise, *Acoustics, Speech and Signal Processing, IEEE Transactions on*, *38*(5), 814-824.

IEC (2005), Standard IEC 62209-1, Human exposure to radio frequency fields from hand-held and body-mounted wireless communication devices–human models, instrumentation, and procedures, edited, International Electrotechnical Commission.

IEEE (1993), IEEE Standard Definitions of Terms for Antennas, *IEEE Std 145-1993*, i.

IFAC Dielectric Properties of Body Tissues, edited, ITALIAN NATIONAL RESEARCH COUNCIL.

Immoreev, I. I., and P. G. S. D. V. Fedotov (2002), Ultra wideband radar systems: advantages and disadvantages, paper presented at Ultra Wideband Systems and Technologies, 2002. Digest of Papers. 2002 IEEE Conference on, 2002.

James, J., and L. Longdon (1969), Prediction of arbitrary electromagnetic fields from measured data, *Alta Frequenza*, *38*, 286-290.

Jefimenko, O. D. (1966), *Electricity and magnetism; an introduction to the theory of electric and magnetic fields*, xii, 591 p. pp., Appleton-Century-Crofts, New York,.

Joshi, G. (2006), Ultra-wideband Channel Modeling using Singularity Expansion Method, Virginia Polytechnic Institute and State University.

Kanda, M. (1978), A relatively short cylindrical broadband antenna with tapered resistive loading for picosecond pulse measurements, *Antennas and Propagation, IEEE Transactions on*, *26*(3), 439-447.

Kanda, M. (1982), The Effects of Resistive Loading of "TEM" Horns, *Electromagnetic Compatibility, IEEE Transactions on*, *EMC-24*(2), 245-255.

Kangwook, K., and W. R. Scott (2005), Impulse-radiating antenna with an offset geometry, *Antennas and Propagation, IEEE Transactions on*, *53*(5), 1738-1744.

Kitra, M. I., C. J. Panagamuwa, P. McEvoy, J. Vardaxoglou, and J. R. James (2007), Low SAR ferrite handset antenna design, *Antennas and Propagation, IEEE Transactions on*, *55*(4), 1155-1164.

Kraus, J. D., and R. J. Marhefka (2002), *Antennas for all applications*, 3rd ed., xviii, 938 p. pp., McGraw-Hill, New York.

Kuang-Hao, L., L. Xinhua, X. S. Shen, and J. W. Mark (2008), Performance Analysis of Prioritized MAC in UWB WPAN With Bursty Multimedia Traffic, *Vehicular Technology, IEEE Transactions on*, *57*(4), 2462-2473.

Kuk, F., and K. Nielsen (1997), Factors affecting interference from digital cellular telephones, *Hearing Journal*, *50*, 32-35.

Kuster, N., and Q. Balzano (1992), Energy absorption mechanism by biological bodies in the near field of dipole antennas above 300 MHz, *Vehicular Technology, IEEE Transactions on*, *41*(1), 17-23.

Kusuma, A., A. Sheta, I. Elshafiey, Z. Siddiqui, M. Alkanhal, S. Aldosari, S. Alshebeili, and S. Mahmoud (2011), A new low SAR antenna structure for wireless handset applications, *Progress In Electromagnetics Research*, *112*, 23-40.

Lanbo, L., L. Zijian, and B. E. Barrowes (2011), Through-Wall Bio-Radiolocation With UWB Impulse Radar: Observation, Simulation and Signal Extraction, *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, *4*(4), 791-798.

Lembrikov, B. I. (2011), Novel Applications of the UWB Technologies, edited, InTech, Croatia.

Levis, C. A. (1957), A Reactance Theorem for Antennas, *Proceedings of the IRE*, *45*(8), 1128-1134.

Li, L., A. E. C. Tan, K. Jhamb, and K. Rambabu (2012), Buried Object Characterization Using Ultra-Wideband Ground Penetrating Radar, *Microwave Theory and Techniques, IEEE Transactions on*, *60*(8), 2654-2664.

Licul, S. (2004), Ultra-wideband antenna characterization and measurements, Virginia Tech.

Licul, S., and W. A. Davis (2005), Unified frequency and time-domain antenna modeling and characterization, *Antennas and Propagation, IEEE Transactions on*, *53*(9), 2882-2888.

Liénard, A. (1898), *Champ électrique et magnétique produit par une charge électrique concentrée en un point et animée d'un mouvement quelconque*, G. Carré et C. Naud.

Lin, J. C. (2002), Cellular mobile telephones and children, *Antennas and Propagation Magazine, IEEE*, *44*(5), 142-145.

Liu, G., C. A. Grimes, and D. M. Grimes (1999), A time-domain technique For determining antenna Q, *Microwave and Optical Technology Letters*, *21*(6), 395-398.

Manteghi, M. (2009), Antenna miniaturization beyond the fundamental limits using impedance modulation, paper presented at Antennas and Propagation Society International Symposium, 2009. APSURSI '09. IEEE, 1-5 June 2009.

Manteghi, M., and Y. Rahmat-Samii (2006), Improved feeding structures to enhance the performance of the reflector impulse radiating antenna (IRA), *Antennas and Propagation, IEEE Transactions on*, *54*(3), 823-834.

Matlab (2007), The MathWorks, *Inc., Natick, MA*, *5*.

McLean, J. S. (1996), A re-examination of the fundamental limits on the radiation Q of electrically small antennas, *Antennas and Propagation, IEEE Transactions on*, *44*(5), 672.

Merenda, J. T. (2006), Digital wideband small antenna systems, *BAE Systems*.

Merenda, J. T. (2008), Waveform-synthesis method that reduces battery power in an electrically small wideband radiating system, *Microwaves, Antennas & Propagation, IET*, *2*(1), 59-65.

Meyer, F. J. C., D. B. Davidson, U. Jakobus, and M. A. Stuchly (2003), Human exposure assessment in the near field of GSM base-station antennas using a hybrid finite element/method of moments technique, *Biomedical Engineering, IEEE Transactions on*, *50*(2), 224-233.

Miller, E. K. (1998a), Model-based parameter estimation in electromagnetics. I. Background and theoretical development, *Antennas and Propagation Magazine, IEEE*, *40*(1), 42-52.

Miller, E. K. (1998b), Model-based parameter estimation in electromagnetics. II. Applications to EM observables, *Antennas and Propagation Magazine, IEEE*, *40*(2), 51-65.

Miller, E. K. (1998c), Model-based parameter estimation in electromagnetics. III. Applications to EM integral equations, *Antennas and Propagation Magazine, IEEE*, *40*(3), 49-66.

Miller, E. K. (2001), Electromagnetics without equations, *Potentials, IEEE*, *20*(2), 16-20.

Mugnai, D., A. Ranfagni, and R. Ruggeri (2000), Observation of superluminal behaviors in wave propagation, *Physical review letters*, *84*(21), 4830-4833.

Nekoogar, F., and F. Dowla (2009), Software-Defined Ultra-wideband Radio Communications: A New RF Technology for Emergency Response Applications*Rep.*, Lawrence Livermore National Laboratory (LLNL), Livermore, CA.

Nicolson, A. M., and G. F. Ross (1970), Measurement of the Intrinsic Properties of Materials by Time-Domain Techniques, *Instrumentation and Measurement, IEEE Transactions on*, *19*(4), 377-382.

NIDCD (2006), Statistics about Hearing, Balance, Ear Infections, and Deafness, edited, p. National Institute on Deafness and Other Communication Disorders.

NIH (1986), Visible Human Project, edited, U.S. National Library of Medicine.

Okoniewski, M., and M. A. Stuchly (1998), Modeling of interaction of electromagnetic fields from a cellular telephone with hearing aids, *Microwave Theory and Techniques, IEEE Transactions on*, *46*(11), 1686-1693.

Ollikainen, J., O. Kivekäs, A. Toropainen, and P. Vainikainen (2000), Internal dual-band patch antenna for mobile phones.

Oppenheim, A. V., and R. W. Schafer (2010), *Discrete-time signal processing*, 3rd ed., xxviii, 1108 p. pp., Pearson, Upper Saddle River.

Oppenheim, A. V., A. S. Willsky, and S. H. Nawab (1997), *Signals & systems*, 2nd ed., xxx, 957 p. pp., Prentice Hall, Upper Saddle River, N.J.

Panofsky, W. K. H., and M. Phillips (2005), *Classical electricity and magnetism*, 2nd ed., xvi, 494 p. pp., Dover Publications, Mineola, N.Y.

Potter, P. (1967), Application of spherical wave theory to cassegrainian-fed paraboloids, *Antennas and Propagation, IEEE Transactions on*, *15*(6), 727-736.

Pozar, D. M. (2012), *Microwave engineering*, 4th ed., xvii, 732 p. pp., Wiley, Hoboken, NJ.

Proakis, J. G. (2000), *Digital communications*, 4th ed., xxi, 1002 p. pp., McGraw-Hill, Boston.

Prony, G. (1795), Essai Experimental et Analytique, etc, *J. l'Ecole Polytechnique*, *1*(2), 24-7676.

Ramo, S., J. R. Whinnery, and T. Van Duzer (1994), *Fields and waves in communication electronics*, 3rd ed., xix, 844 p. pp., Wiley, New York.

Reed, J. (2005), *Introduction to ultra wideband communication systems, an*, Prentice Hall Press.

Reed, J. H. (2002), *Software radio a modern approach to radio engineering*, xx, 567 p. ill. 525 cm. pp., Prentice Hall,, Upper Saddle River, NJ.

Rhodes, D. R. (1976), Observable stored energies of electromagnetic systems, *Journal of the Franklin Institute*, *302*(3), 225-237.

Safaai-Jazi, A., and J. C. Cardoso (1996), Radiation characteristics of a spherical helical antenna, *Microwaves, Antennas and Propagation, IEE Proceedings*, *143*(1), 7-12.

Sahinoglu, Z., S. Gezici, and I. Guvenc (2008), *Ultra-wideband positioning systems*, Cambridge University Press Cambridge, UK:.

Sarkar, T. K., and O. Pereira (1995), Using the matrix pencil method to estimate the parameters of a sum of complex exponentials, *Antennas and Propagation Magazine, IEEE*, *37*(1), 48-55.

Sarkar, T. K., and A. Taaghol (1999), Near-field to near/far-field transformation for arbitrary near-field geometry utilizing an equivalent electric current and MoM, *Antennas and Propagation, IEEE Transactions on*, *47*(3), 566-573.

Schantz, H. (2005), *The art and science of ultrawideband antennas*, xix, 331 p. pp., Artech House, Boston.

Schantz, H. G. (2001), Electromagnetic energy around Hertzian dipoles, *Antennas and Propagation Magazine, IEEE*, *43*(2), 50-62.

Schelkunoff, S. A. (1952), *Advanced antenna theory*, 216p. pp., Wiley, New York,.

Schönborn, F., M. Burkhardt, and N. Kuster (1998), Differences in energy absorption between heads of adults and children in the near field of sources, *Health Physics*, *74*(2), 160.

Seabury, D. (2005), An Update On SAR Standards And The Basic Requirements For SAR Assessment, *Conformity Magazine, April*.

Shlager, K. L., G. S. Smith, and J. G. Maloney (1995), TEM horn antenna for pulse radiation: an optimized design, paper presented at Antennas and Propagation Society International Symposium, 1995. AP-S. Digest, 18-23 Jun 1995.

Shlivinski, A., E. Heyman, and R. Kastner (1997), Antenna characterization in the time domain, *Antennas and Propagation, IEEE Transactions on*, *45*(7), 1140-1149.

Simons, R. (1996), Guglielmo Marconi and Early Systems of Wireless Communication, *GEC Review*, *11*(1), 37-55.

Skopec, M. (1998), Hearing aid electromagnetic interference from digital wireless telephones, *Rehabilitation Engineering, IEEE Transactions on*, *6*(2), 235-239.

Smith, G. S. (1997), *An introduction to classical electromagnetic radiation*, xvi, 653 p. pp., Cambridge University Press, Cambridge, U.K. ; New York, NY, USA.

Smith, G. S. (2002), Teaching antenna reception and scattering from a time-domain perspective, *American Journal of Physics*, *70*, 829.

Smith, G. S., and T. W. Hertel (2001), On the transient radiation of energy from simple current distributions and linear antennas, *Antennas and Propagation Magazine, IEEE*, *43*(3), 49-62.

SPEAG (2006), DASY3 System, edited.

SPEAG (2007), DASY4 System, edited.

Stratton, J. A. (1941), *Electromagnetic theory*, 1st ed., xv, 615 p. pp., McGraw-Hill book company, inc., New York, London,.

Stumpf, W. (2003), Ultrawideband: New promises, new problems, *RF design*, 56-60.

Stutzman, W. L., and G. A. Thiele (2012), *Antenna theory and design*, 3rd ed., xviii, 822 p. pp., Wiley, Hoboken, NJ.

Tarter, M. E., and M. D. Lock (1993), *Model-free curve estimation*, x, 290 p. pp., Chapman & Hall, New York.

Tesche, F. (1973), On the analysis of scattering and antenna problems using the singularity expansion technique, *Antennas and Propagation, IEEE Transactions on*, *21*(1), 53-62.

Van Blaricum, M., and R. Mittra (1975), A technique for extracting the poles and residues of a system directly from its transient response, *Antennas and Propagation, IEEE Transactions on*, *23*(6), 777-781.

Victorian, T. (1998), An update on digital cellular telephone interference and hearing aid compatibility, *Hearing Journal*, *51*, 53-61.

Victorian, T. (2004), Hearing Aid Compatibility: Technical Update, edited, AudiologyOnline.

Volakis, J. L. (2007), *Antenna engineering handbook*, 4th ed., McGraw-Hill, New York.

VTAG (2008), Near-field, indoor anechoic chamber at Virginia Tech Antenna Group, edited.

Weir, W. B. (1974), Automatic measurement of complex dielectric constant and permeability at microwave frequencies, *Proceedings of the IEEE*, *62*(1), 33-36.

Weiss, L., and R. McDonough (1963), Prony's method, z-transforms, and Pade approximation, *Siam Review*, *5*(2), 145-149.

Wheeler, H. A. (1947), Fundamental limitations of small antennas, *Proceedings of the IRE*, *35*(12), 1479-1484.

Wheeler, H. A. (1959), The radiansphere around a small antenna, *Proceedings of the IRE*, *47*(8), 1325-1331.

Wiechert, E. (1901), Elektrodynamische Elementargesetze, *Annalen der Physik*, *309*(4), 667-689.

Wilson, J. M. (2002), Ultra-wideband/a disruptive rf technology?, *Version 1.3, Intel Research and Development*.

Witters, D. (2000), Medical Devices and EMI: The FDA Perspective.

Wynne, K. (2002), Causality and the nature of information, *Optics communications*, *209*(1-3), 85-100.

Xiaojing, X., and Y. E. Wang (2007), Beyond the efficiency bandwidth limit with switched electrically small antennas, paper presented at Antennas and Propagation Society International Symposium, 2007 IEEE, 9-15 June 2007.

Yaghjian, A. D., and S. R. Best (2005), Impedance, bandwidth, and Q of antennas, *Antennas and Propagation, IEEE Transactions on*, *53*(4), 1298-1324.

Yang, T., and W. A. Davis (2004), Planar half-disk antenna structures for ultra-wideband communications, paper presented at Antennas and Propagation Society International Symposium, 2004. IEEE, 20-25 June 2004.

Yang, T., and W. A. Davis (2005), Mitigating the Transient Chirp of Conical Log-Sprical Antennas, *USNC/URSI National Radio Science Meeting*.

Yang, T., and W. Davis (2006), Miniaturization of Planar Two-Arm Spiral Antennas Using Slow-Wave Enhancements.

Yang, T., and W. A. Davis (2009), Minimum radiation-Q of polarized antennas with various source configurations, paper presented at Antennas and Propagation Society International Symposium, 2009. APSURSI '09. IEEE, 1-5 June 2009.

Yang, T., and W. A. Davis (2011), A direct matrix approach to 3-D antenna radiation-pattern estimation from partially-scanned spherical near-field data, paper presented at General Assembly and Scientific Symposium, 2011 XXXth URSI, 13-20 Aug. 2011.

Yang, T., W. A. Davis, and W. L. Stutzman (2009), Fundamental-limit perspectives on ultrawideband antennas, *Radio Science*, *44*(1), RS1S91.

Yang, T., W. A. Davis, and W. L. Stutzman (2009), Multipath mitigation in pattern measurement of small directive antennas based on the singularity expansion method (SEM), paper presented at Antenna Technology, 2009. iWAT 2009. IEEE International Workshop on, 2-4 March 2009.

Yang, T., W. Davis, W. Stutzman, and M. C. Huynh (2008), Cellular-phone and hearing-aid interaction: An antenna solution, *Antennas and Propagation Magazine, IEEE*, *50*(3), 51-65.

Yang, T., S. Y. Suh, R. Nealy, W. A. Davis, and W. L. Stutzman (2004), Compact antennas for UWB applications, *Aerospace and Electronic Systems Magazine, IEEE*, *19*(5), 16-20.

Yang, T., W. A. Davis, W. L. Stutzman, S. Shajedul Hasan, R. Nealy, C. B. Dietrich, and J. H. Reed (2011), Antenna design strategy and demonstration for software-defined radio (SDR), *Analog Integrated Circuits and Signal Processing*, 1-11.

Appendix A

Proof of Othorgonality for Associated Legendre Polynomials

The associated Legendre equation can be written as

$$\frac{d}{du}\sin^2\theta\frac{d}{du}P_n^m(u)+\left[n(n+1)-\frac{m^2}{\sin^2\theta}\right]P_n^m(u)=0 \tag{A.1}$$

Multiply (A.1) by $P_q^m(u)$ and integrate it over u to have

$$\begin{aligned}&\int_{-1}^{1}\left\{P_q^m(u)\frac{d}{du}\sin^2\theta\frac{d}{du}P_n^m(u)\right\}du\\&+n(n+1)\int_{-1}^{1}\left\{P_n^m(u)P_q^m(u)\right\}du-\int_{-1}^{1}\left\{\frac{m^2}{\sin^2\theta}P_n^m(u)P_q^m(u)\right\}du=0\end{aligned} \tag{A.2}$$

The first integral term in (A.2) can be written as

$$\begin{aligned}&\int_{-1}^{1}\left\{P_q^m(u)\frac{d}{du}\sin^2\theta\frac{d}{du}P_n^m(u)\right\}du\\&=\left[P_q^m(u)\sin^2\theta\frac{d}{du}P_n^m(u)\right]_{-1}^{1}-\int_{-1}^{1}\left\{\sin^2\theta\frac{d}{du}P_n^m(u)\frac{d}{du}P_q^m(u)\right\}du\\&=-\int_{-1}^{1}\left\{\sin^2\theta\frac{d}{du}P_n^m(u)\frac{d}{du}P_q^m(u)\right\}du\end{aligned} \tag{A.3}$$

Then, (A.2) becomes

$$\int_{-1}^{1}\left\{\frac{m^2}{\sin^2\theta}P_n^m(u)P_q^m(u)+\sin^2\theta\frac{d}{du}P_n^m(u)\frac{d}{du}P_q^m(u)\right\}du$$
$$=n(n+1)\int_{-1}^{1}\left\{P_n^m(u)P_q^m(u)\right\}du \tag{A.4}$$

Due to the orthogonal property of the Associated Legendre function, (A.4) is evaluated as

$$\int_{-1}^{1}\left\{\frac{m^2}{\sin^2\theta}P_n^m(u)P_q^m(u)+\sin^2\theta\frac{d}{du}P_n^m(u)\frac{d}{du}P_q^m(u)\right\}du=n(n+1)\frac{2}{2n+1}\frac{(n+m)!}{n-m)!}\delta_{nq} \tag{A.5}$$

or

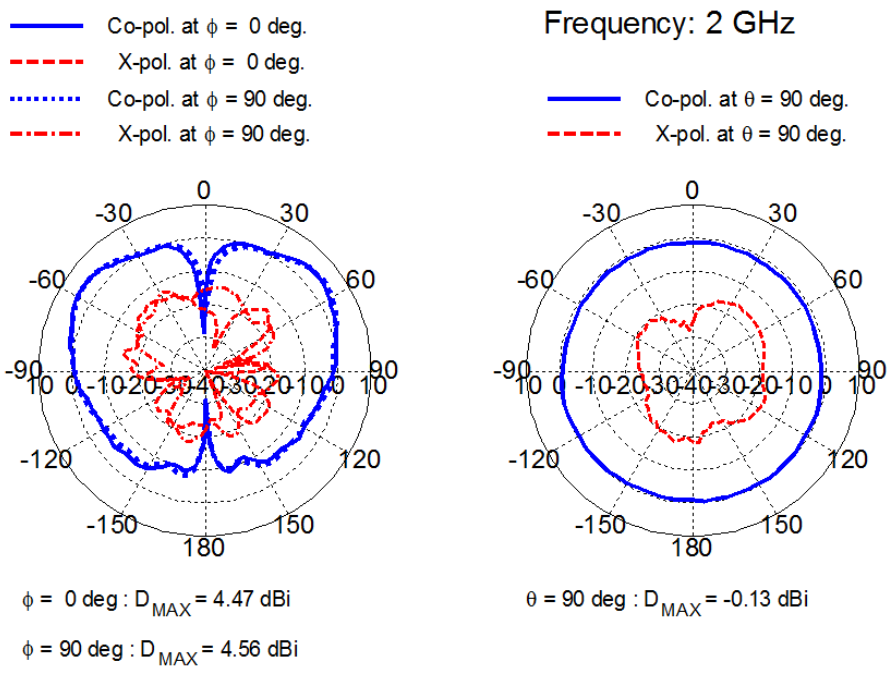
$$\int_{0}^{\pi}\left\{\frac{m^2}{\sin^2\theta}P_n^m(\cos\theta)P_q^m(\cos\theta)+\frac{d}{d\theta}P_n^m(\cos\theta)\frac{d}{d\theta}P_q^m(\cos\theta)\right\}\sin\theta d\theta$$
$$=n(n+1)\frac{2}{2n+1}\frac{(n+m)!}{n-m)!}\delta_{nq} \tag{A.6}$$

where $\delta_{nq}=1$ if $n=q$, otherwise $\delta_{nq}=0$.

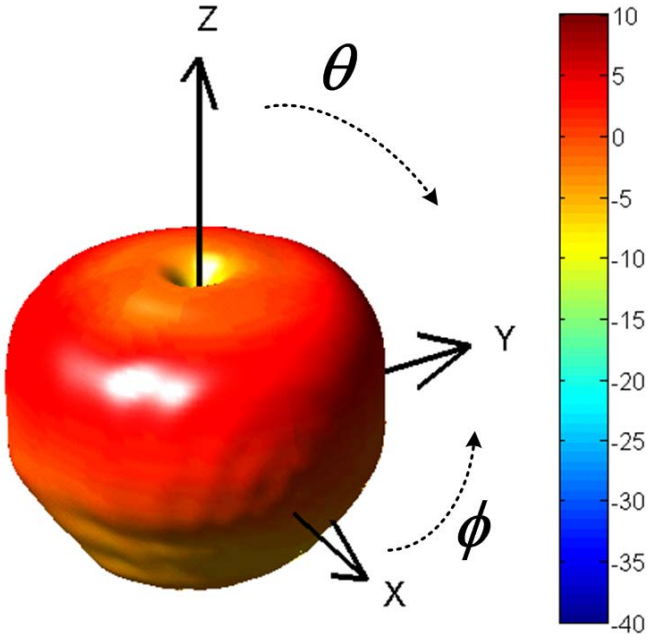
Appendix B

Measured Radiation Patterns of the Designed Compact Ultra-Wideband Antenna (CUA)

Three dimensional radiation patterns of the designed antenna prototype (Figure 6.15) with a finite-size circular ground plane (150 mm radius) were measured in the indoor near-field antenna range of the Virginia Tech Antenna Group [*VTAG*, 2008]. Two orthogonal electric near-fields were measured at a near-field region in the chamber. The far-field radiation patterns were obtained through the spherical near to far-field transformation [*Hansen*, 1988]. The measured 3-D radiation patterns are shown at selected frequencies in Figure A.1 through Figure A.19.

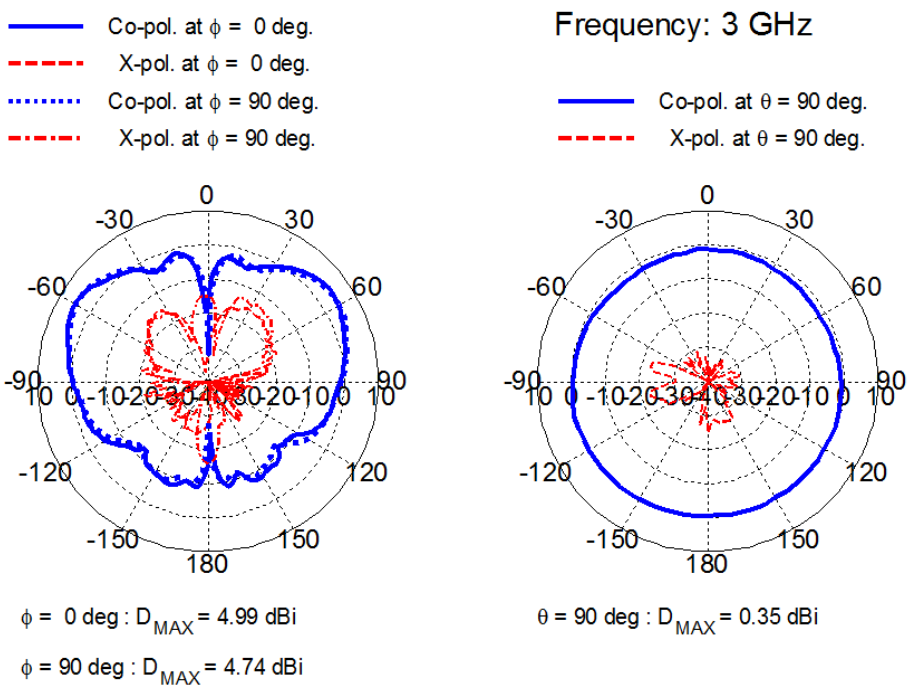


(a)

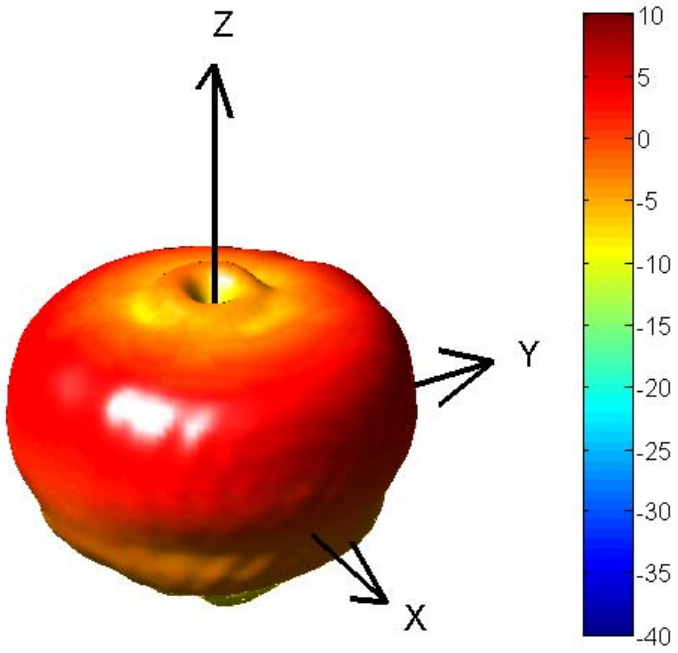


(b)

Figure A.1 Measured radiation patterns of the designed compact UWB antenna at 2 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

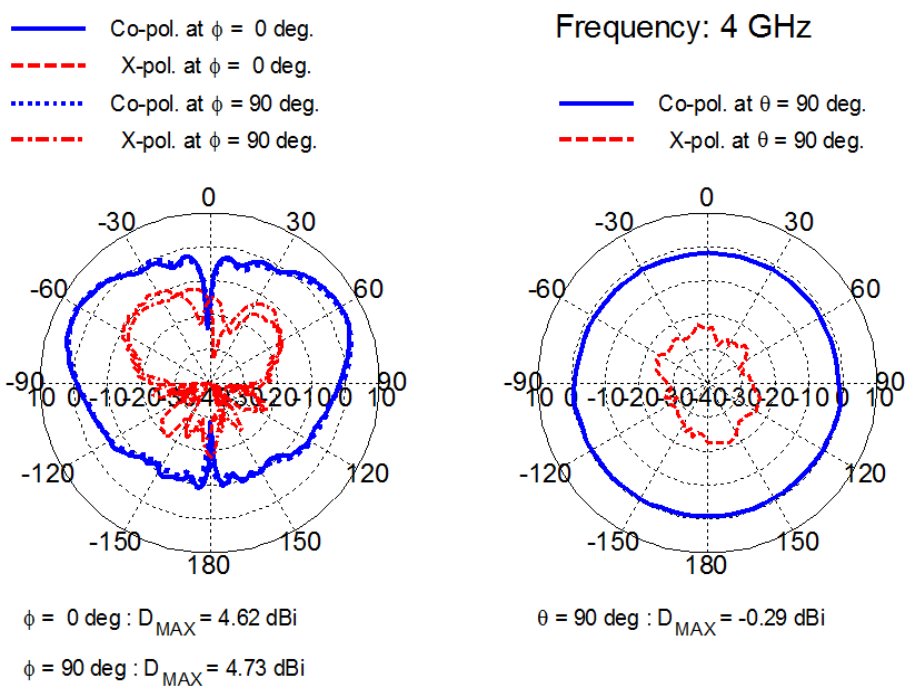


(a)

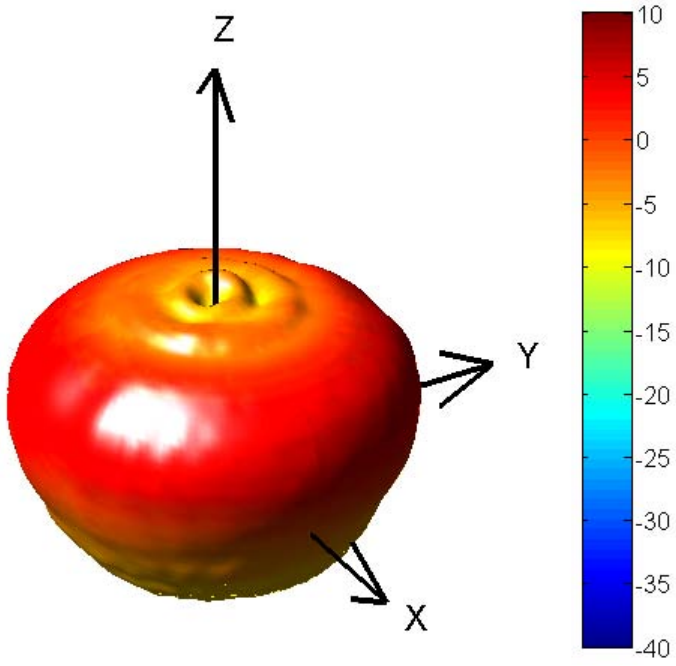


(b)

Figure A.2 Measured radiation patterns of the designed compact UWB antenna at 3 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

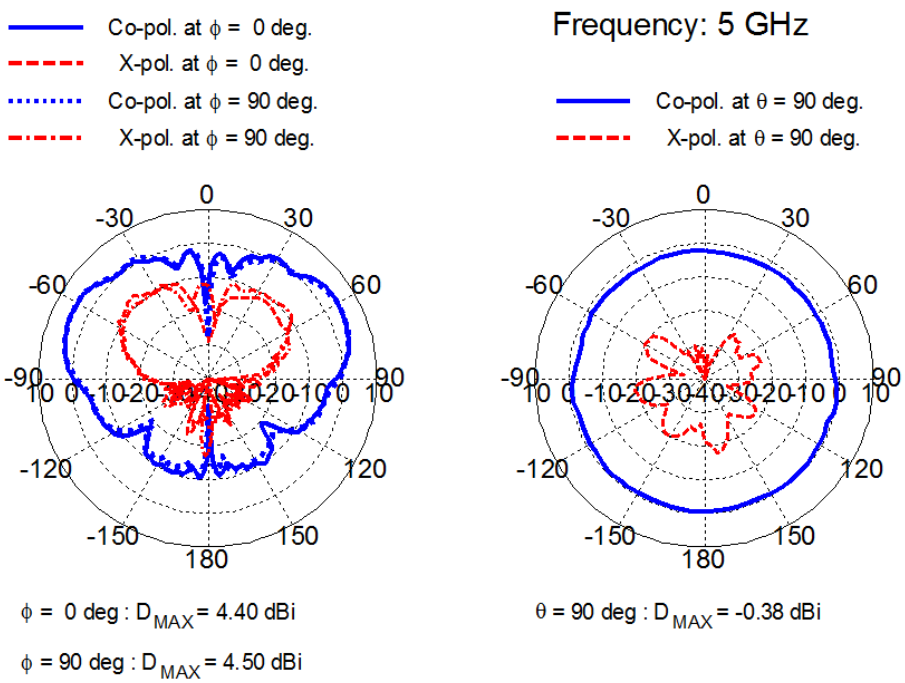


(a)

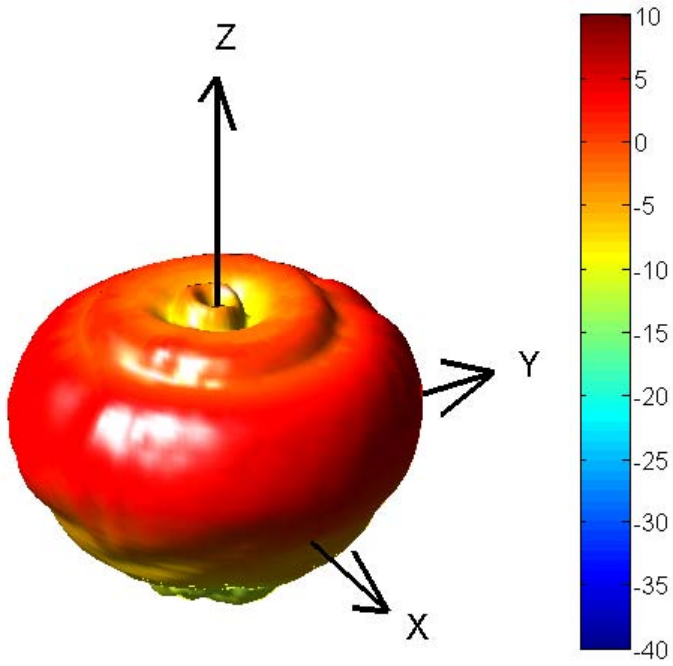


(b)

Figure A.3 Measured radiation patterns of the designed compact UWB antenna at 4 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

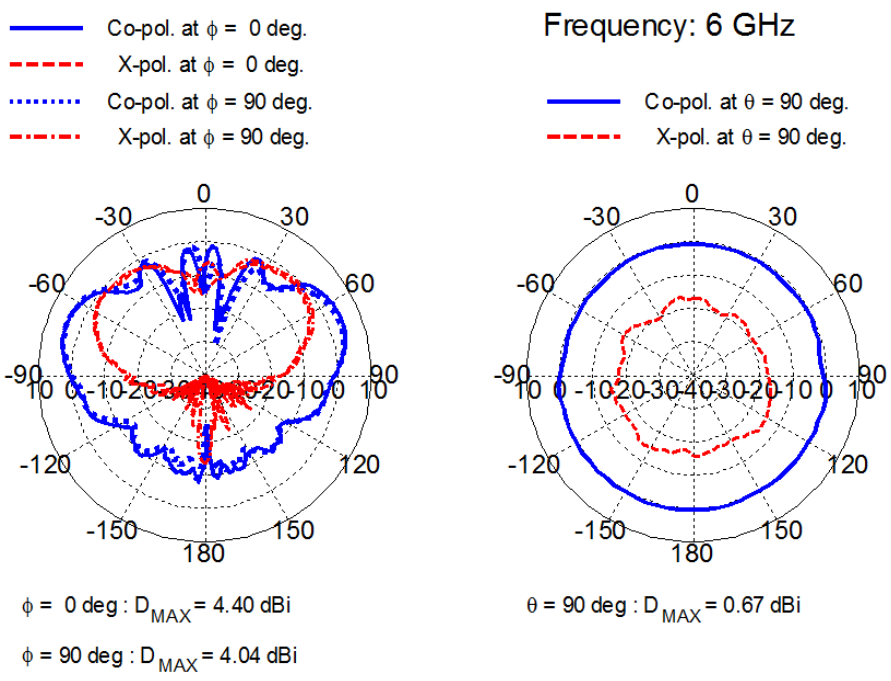


(a)

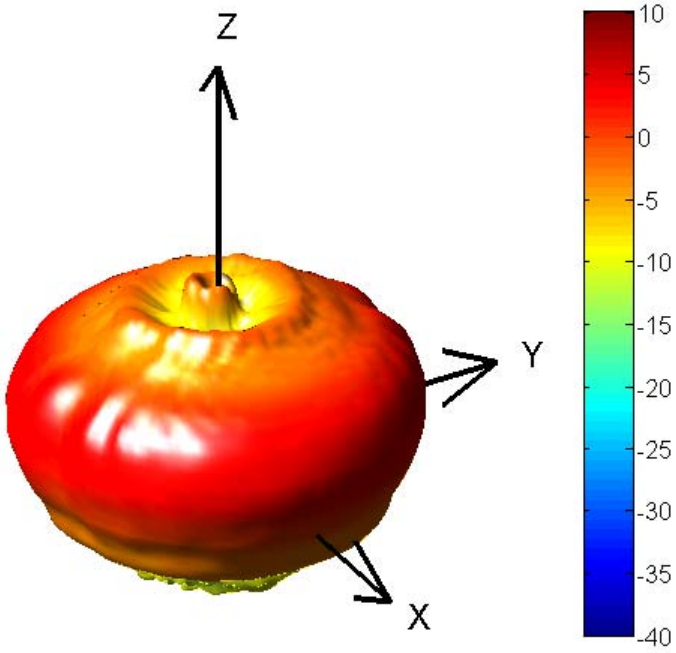


(b)

Figure A.4 Measured radiation patterns of the designed compact UWB antenna at 5 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

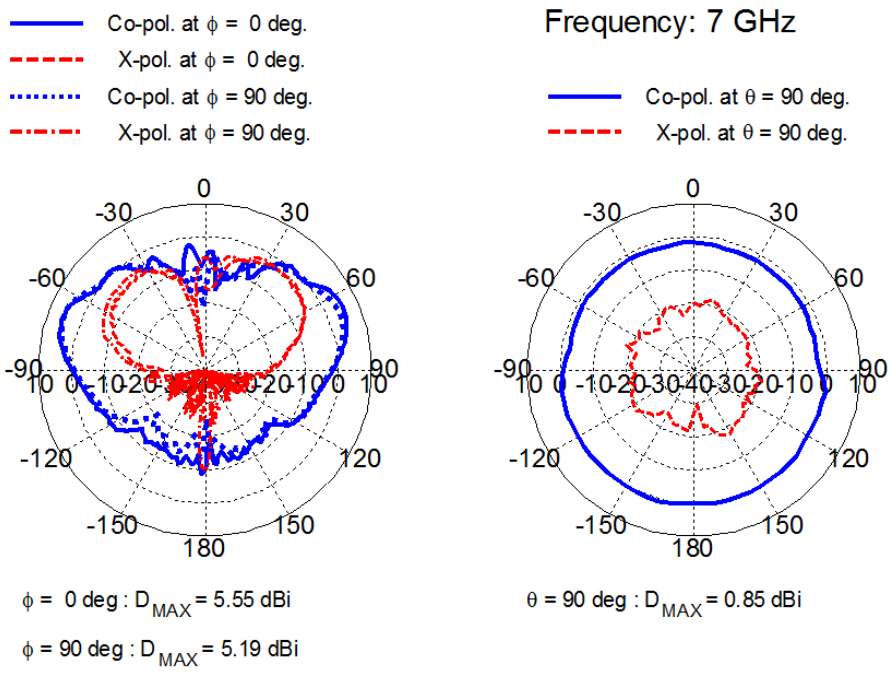


(a)

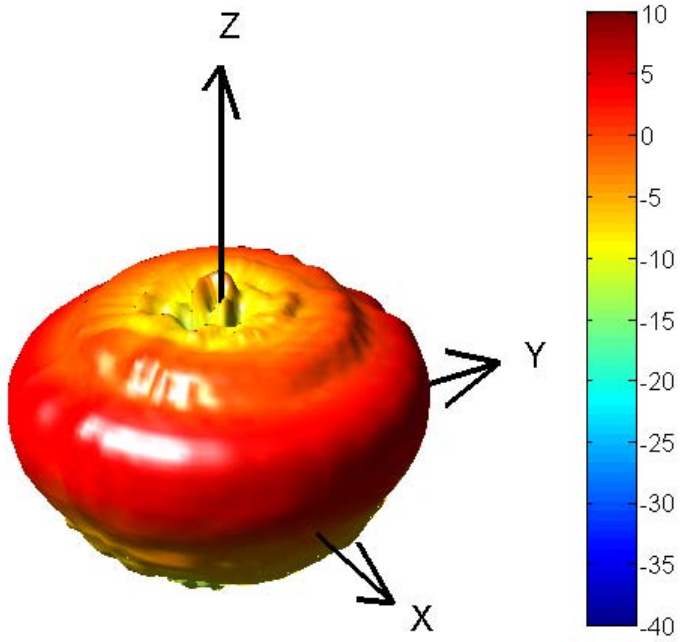


(b)

Figure A.5 Measured radiation patterns of the designed compact UWB antenna at 6 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

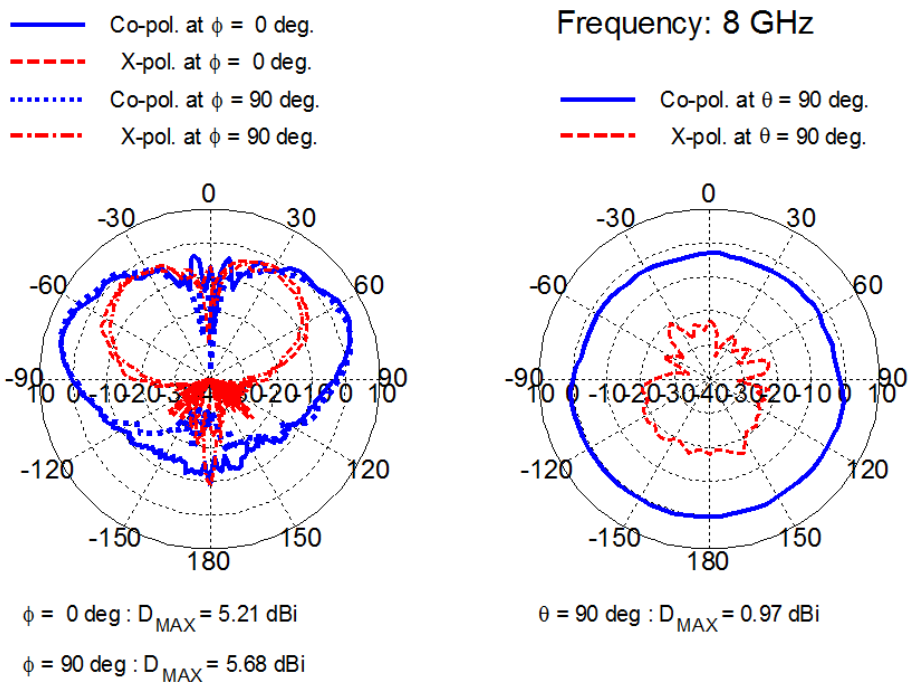


(a)

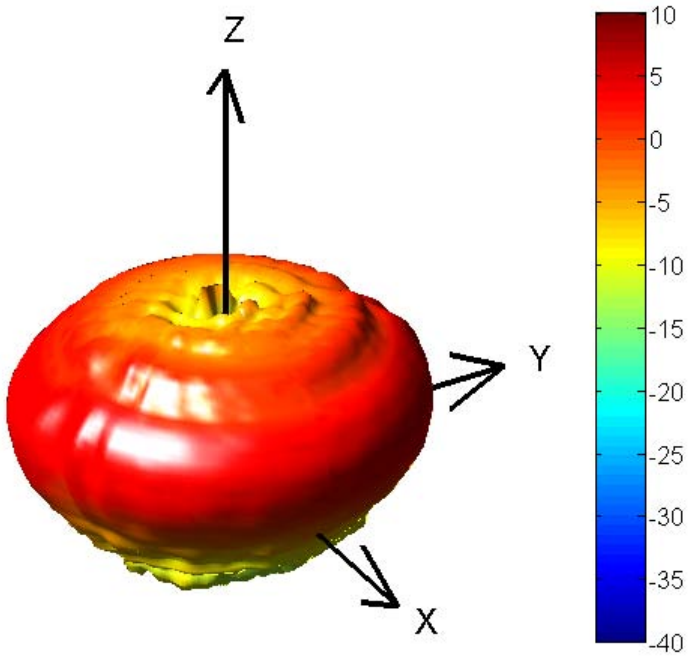


(b)

Figure A.6 Measured radiation patterns of the designed compact UWB antenna at 7 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

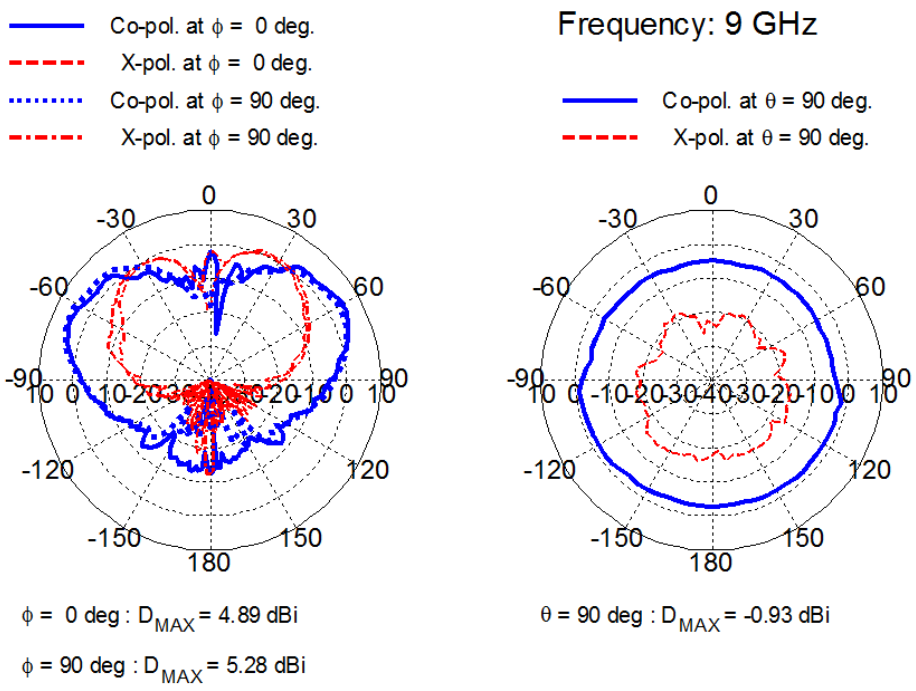


(a)

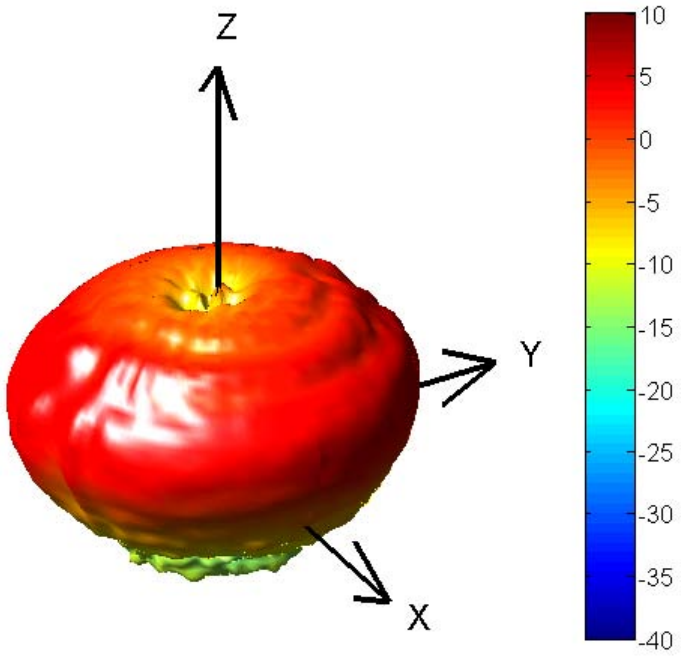


(b)

Figure A.7 Measured radiation patterns of the designed compact UWB antenna at 8 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

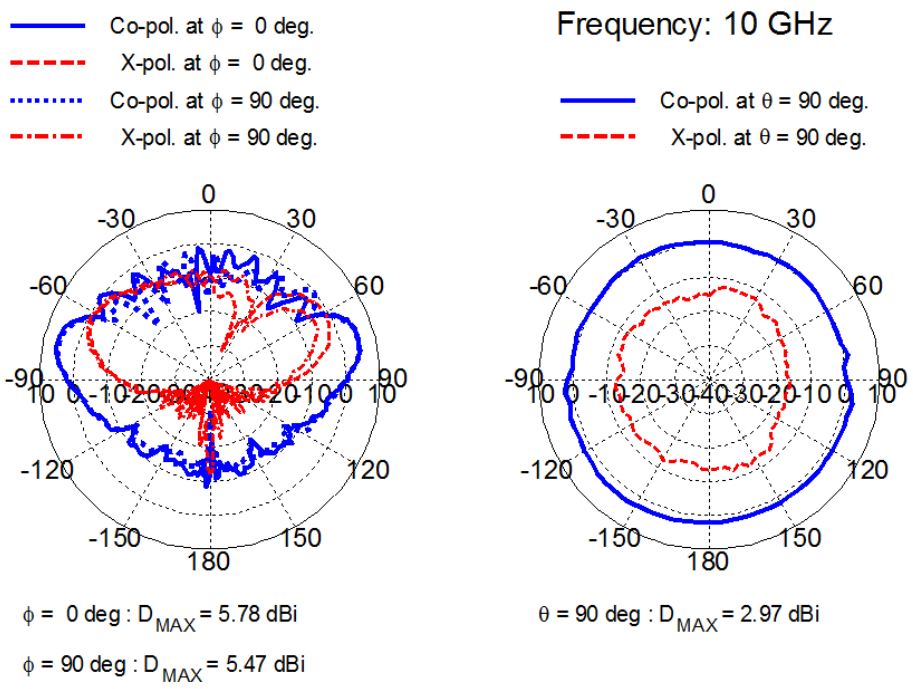


(a)

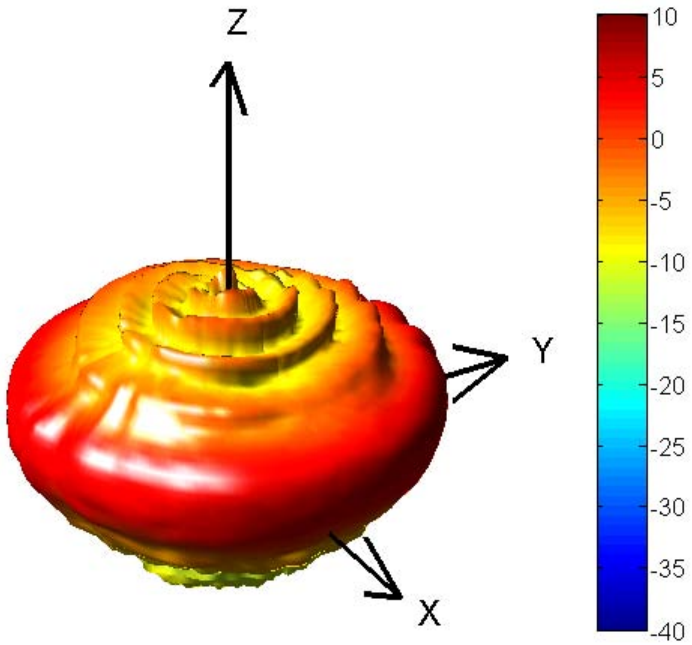


(b)

Figure A.8 Measured radiation patterns of the designed compact UWB antenna at 9 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

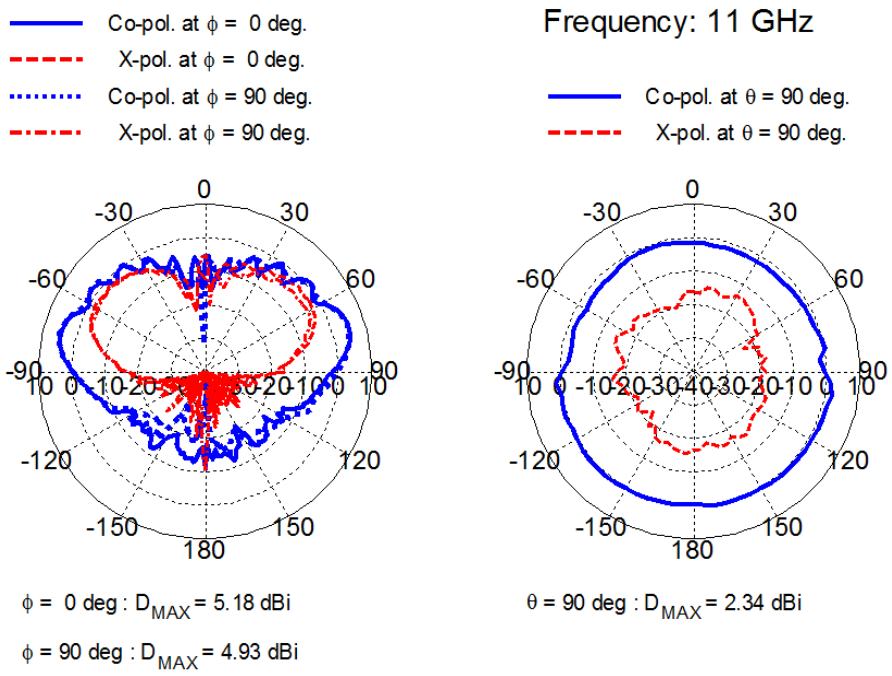


(a)

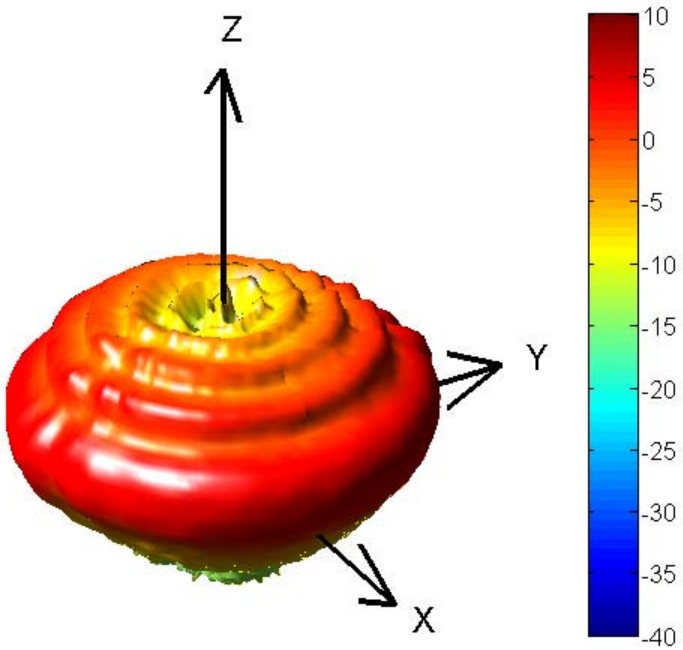


(b)

Figure A.9 Measured radiation patterns of the designed compact UWB antenna at 10 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

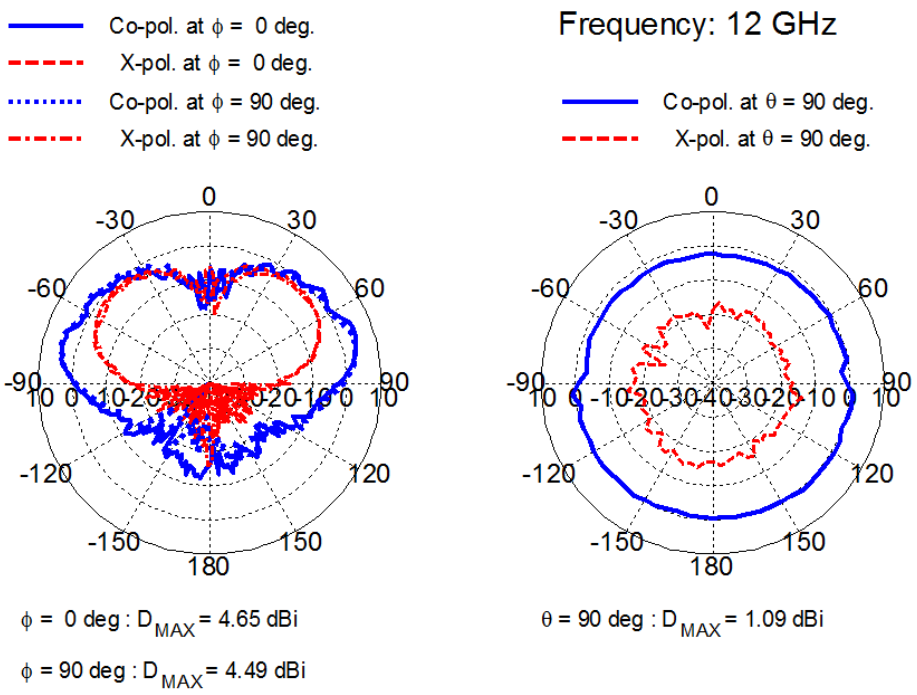


(a)

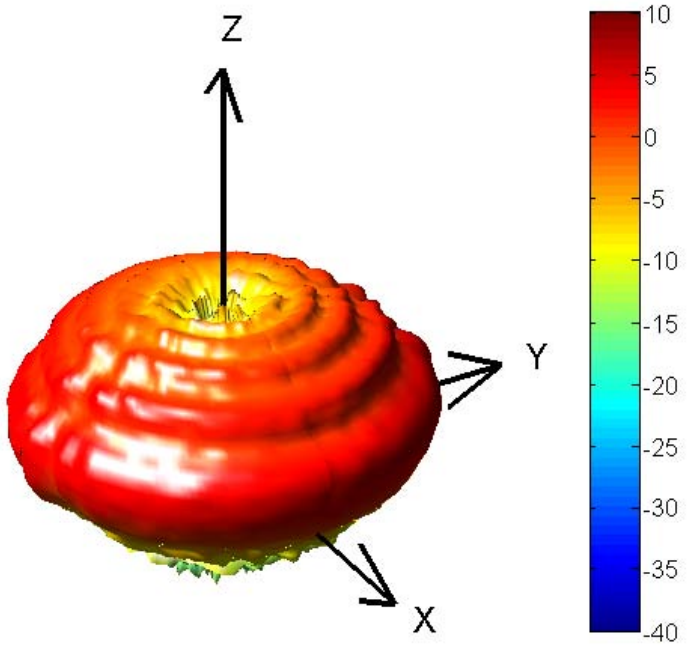


(b)

Figure A.10 Measured radiation patterns of the designed compact UWB antenna at 11 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

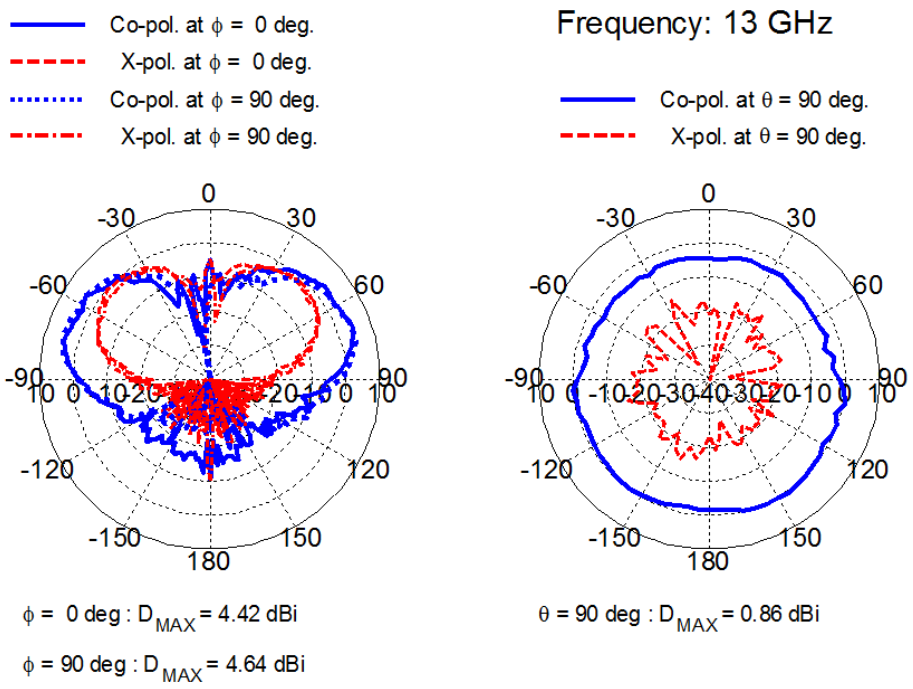


(a)

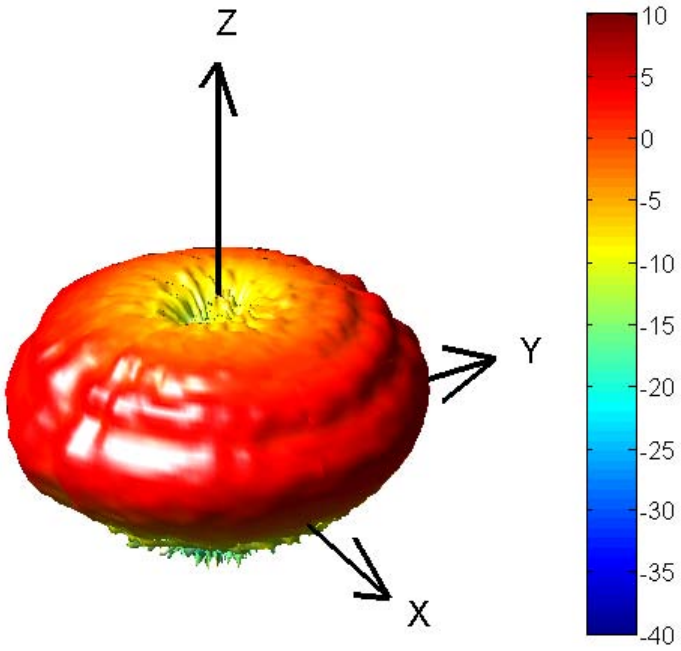


(b)

Figure A.11 Measured radiation patterns of the designed compact UWB antenna at 12 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

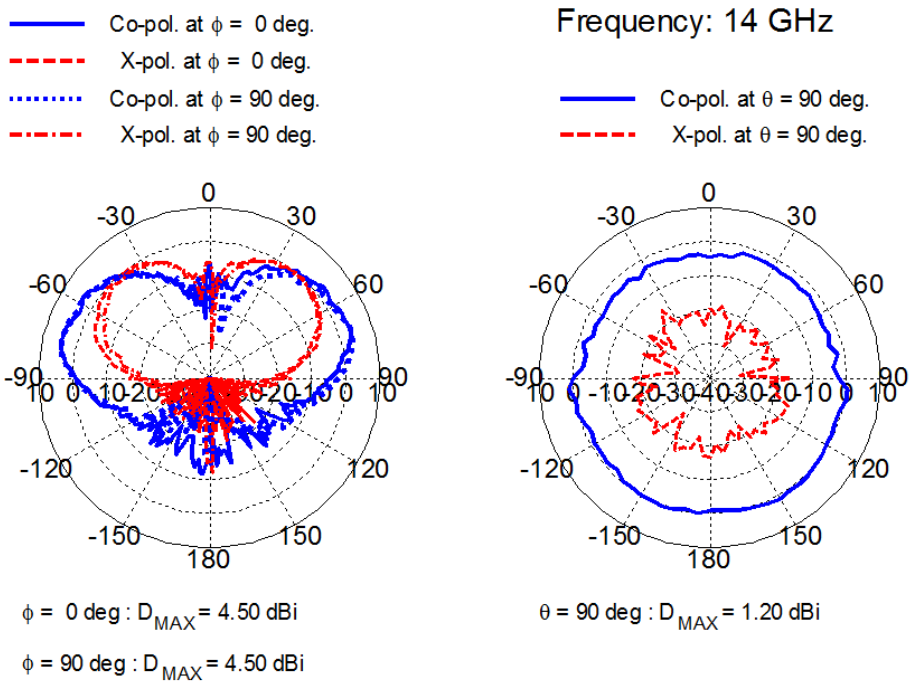


(a)

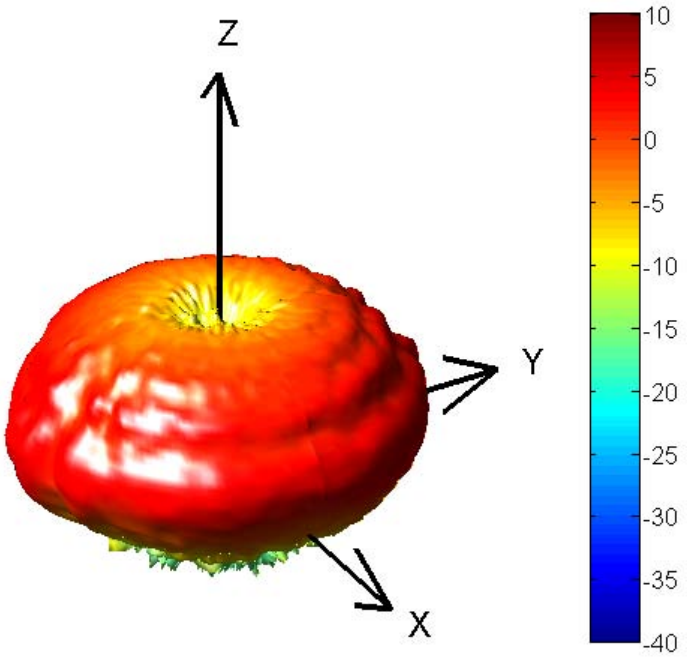


(b)

Figure A.12 Measured radiation patterns of the designed compact UWB antenna at 13 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

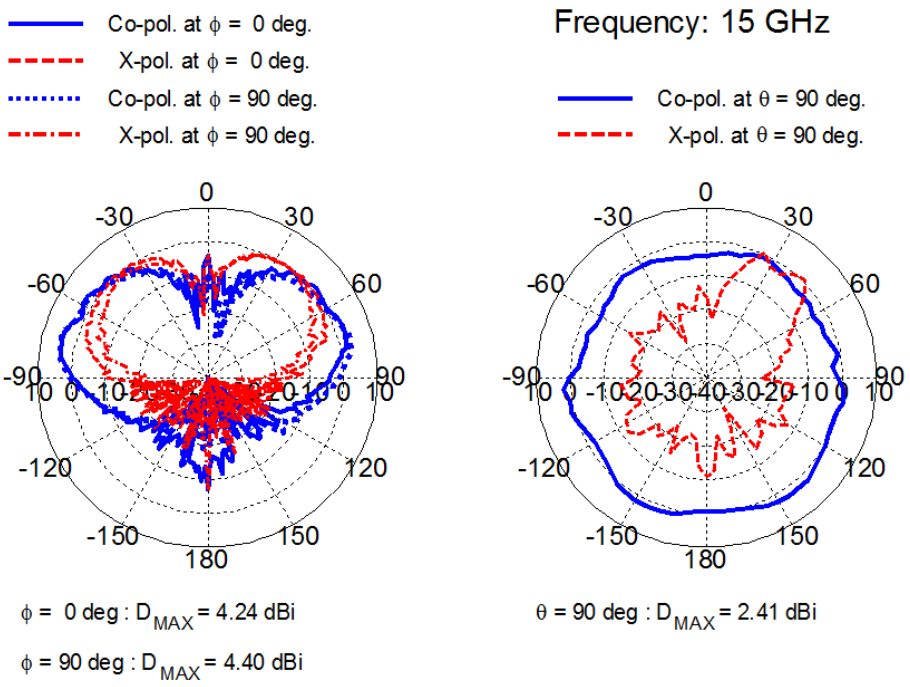


(a)

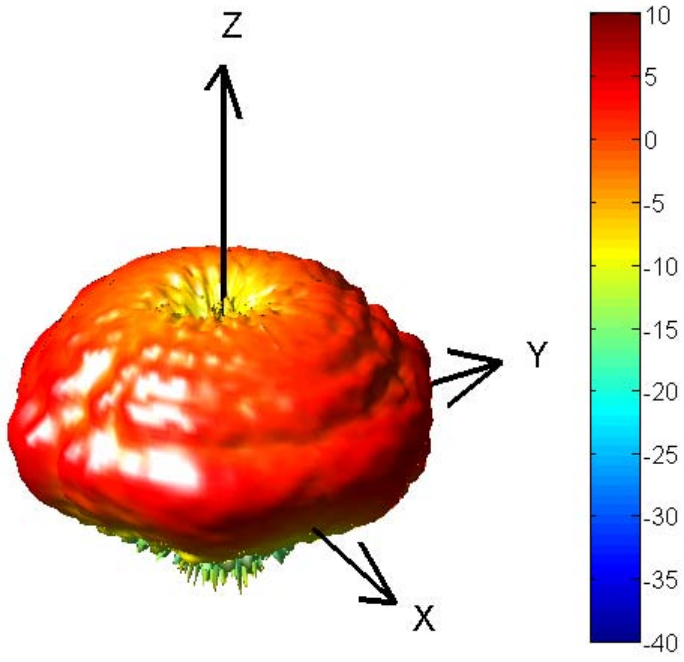


(b)

Figure A.13 Measured radiation patterns of the designed compact UWB antenna at 14 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

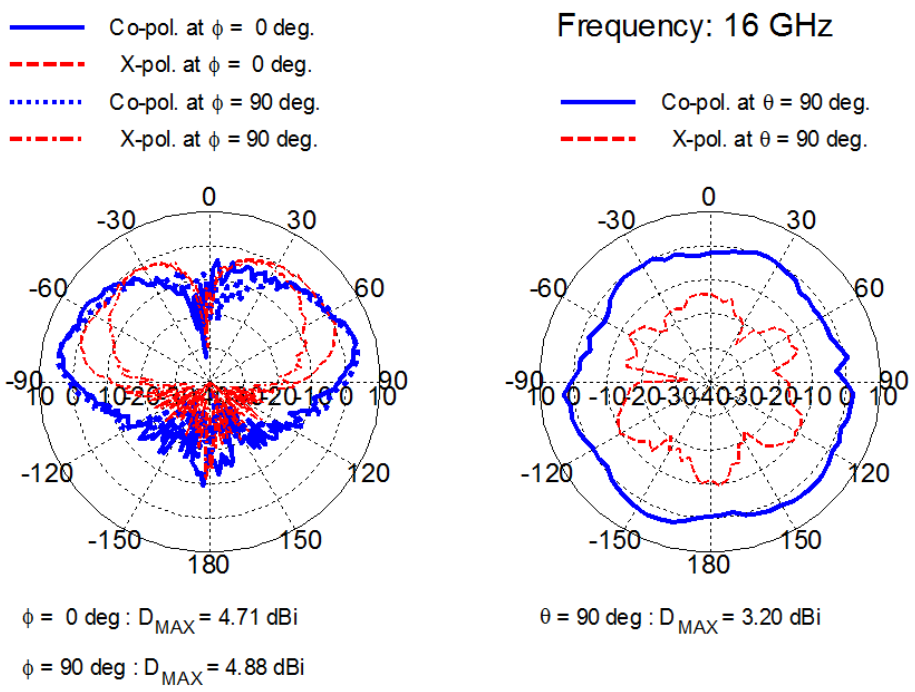


(a)

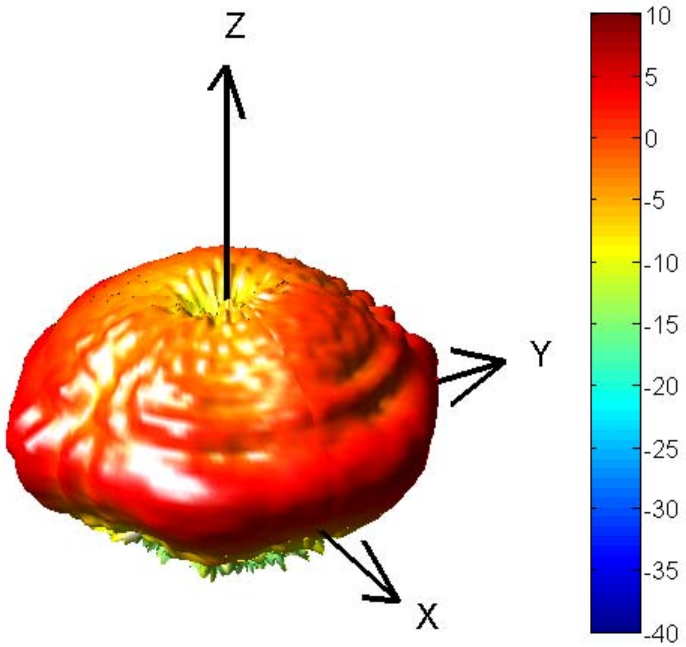


(b)

Figure A.14 Measured radiation patterns of the designed compact UWB antenna at 15 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

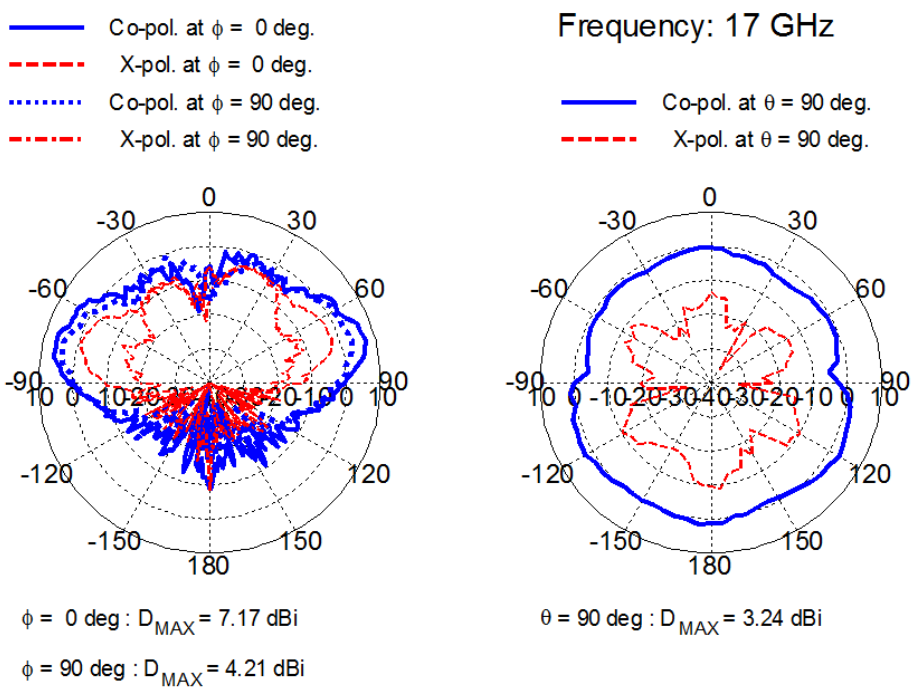


(a)

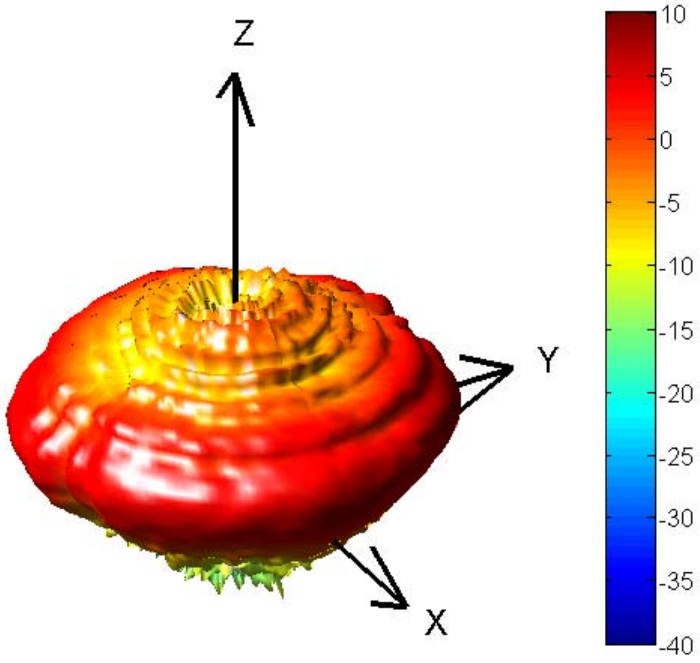


(b)

Figure A.15 Measured radiation patterns of the designed compact UWB antenna at 16 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

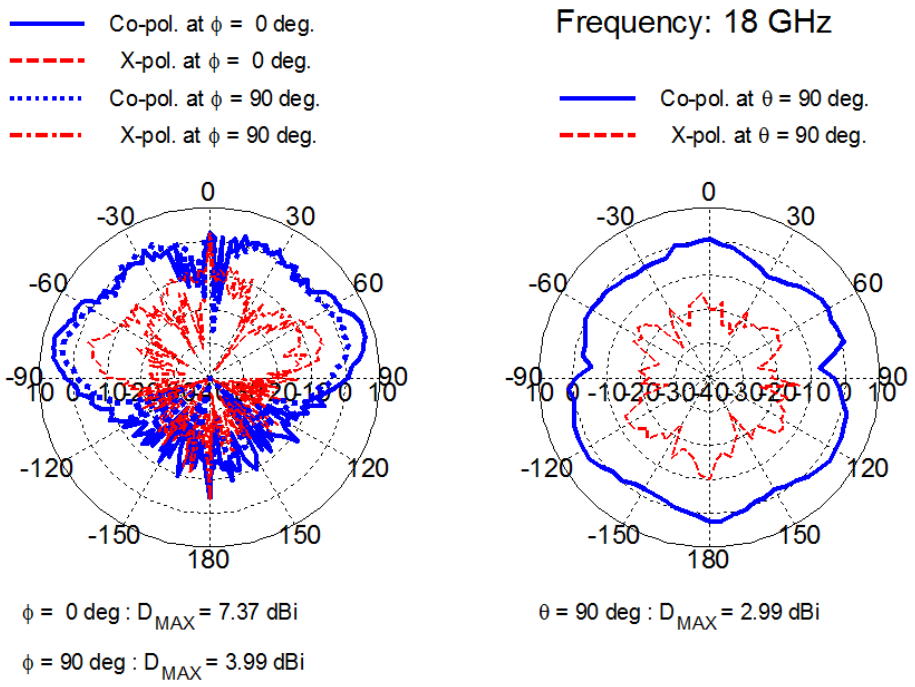


(a)

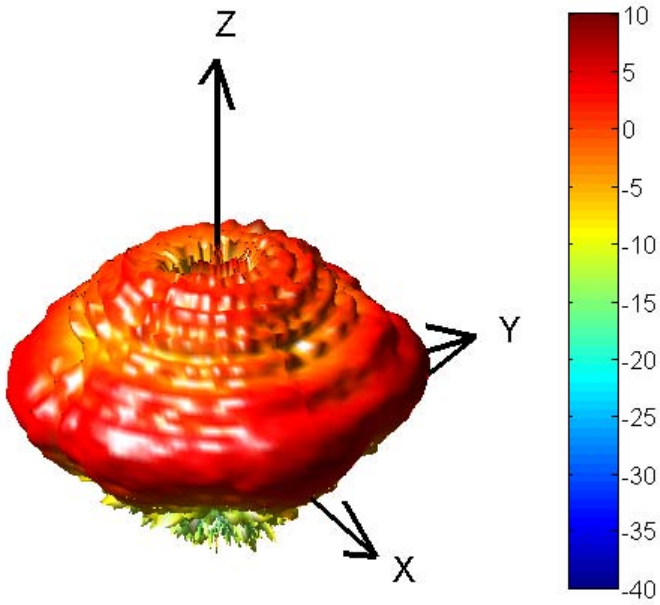


(b)

Figure A.16 Measured radiation patterns of the designed compact UWB antenna at 17 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

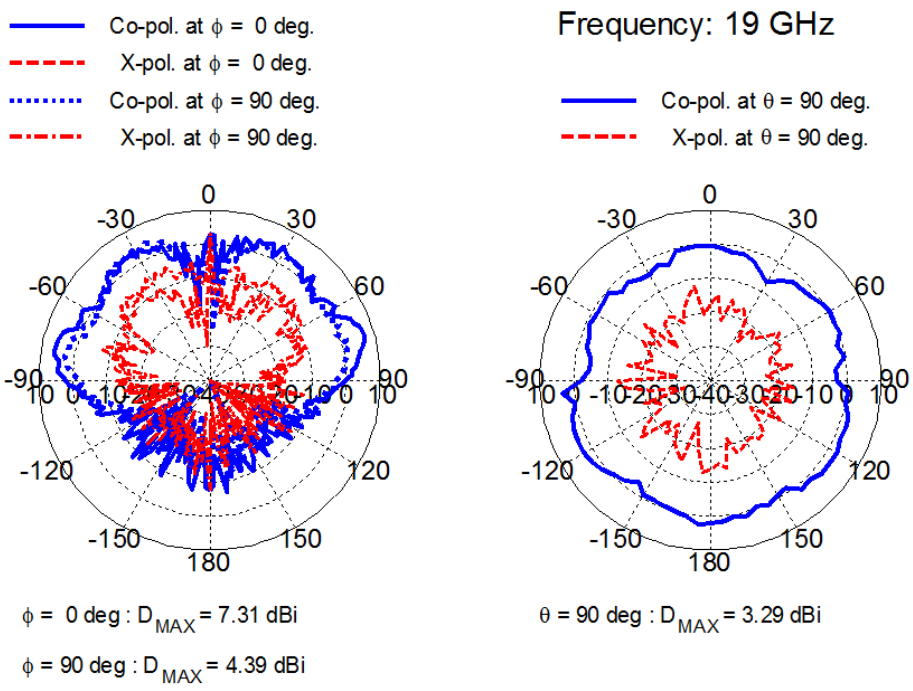


(a)

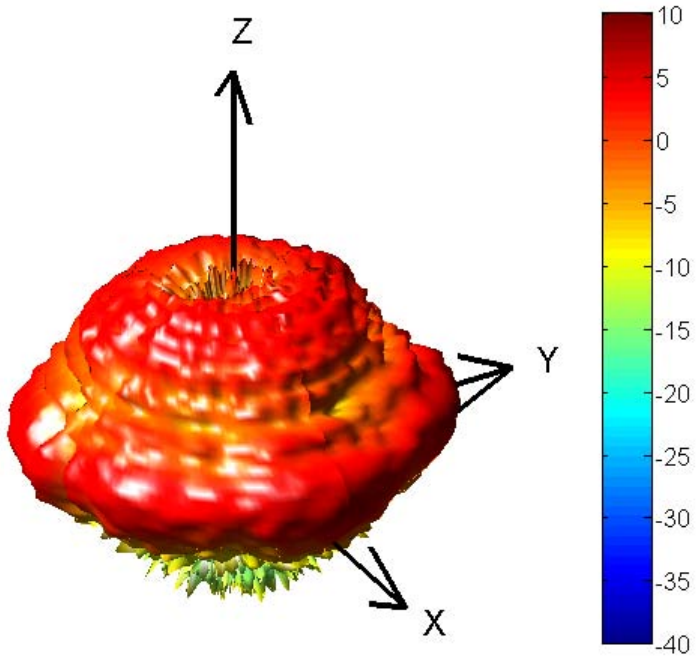


(b)

Figure A.17 Measured radiation patterns of the designed compact UWB antenna at 18 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).

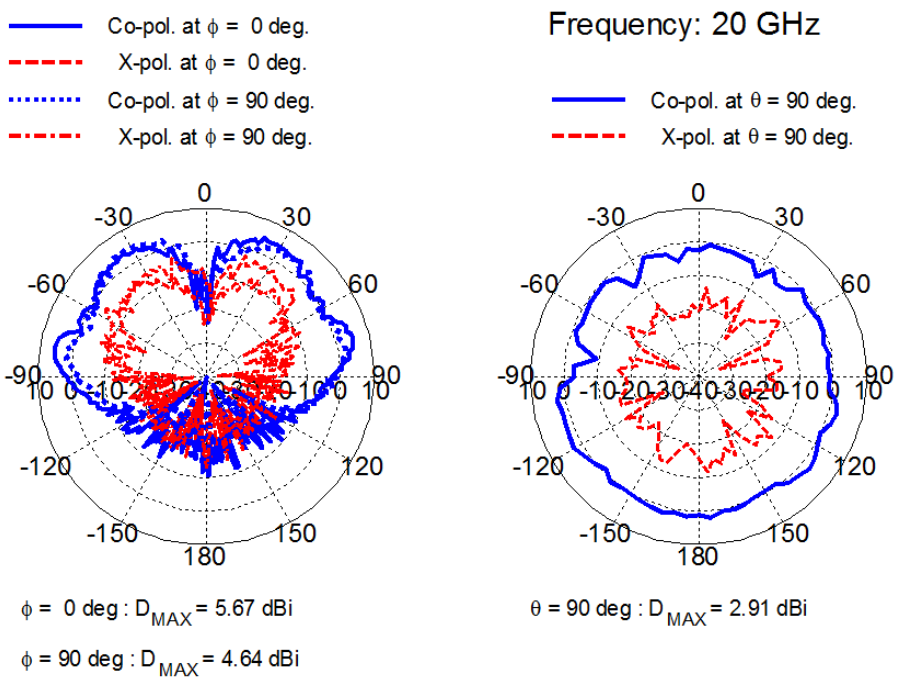


(a)

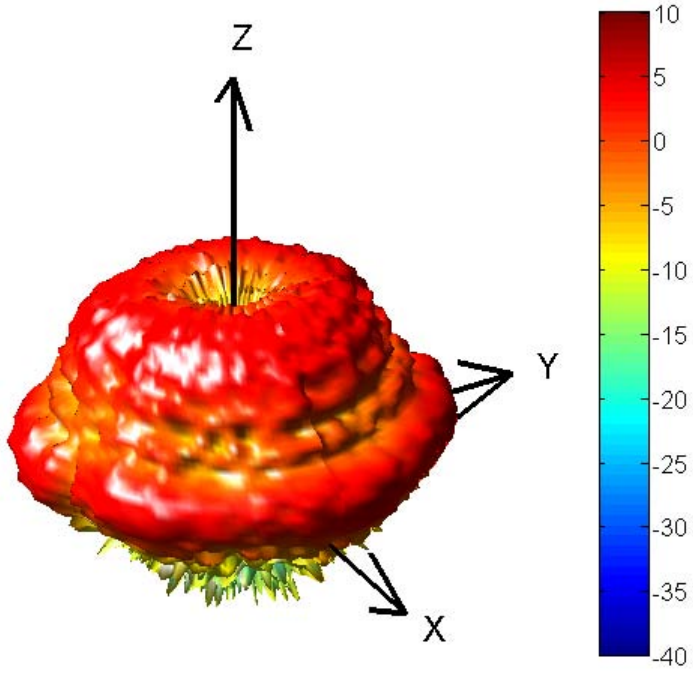


(b)

Figure A.18 Measured radiation patterns of the designed compact UWB antenna at 19 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).



(a)



(b)

Figure A.19 Measured radiation patterns of the designed compact UWB antenna at 20 GHz – (a) Principal plane cuts and (b) 3-dimensional pattern (total field).