

Silicon-based Magneto-optical Isolator and Circulator Fabricated by Direct Bonding Technology

Y. Shoji, K. Mitsuya, and T. Mizumoto
Tokyo Institute of Technology, Japan

Abstract— We have investigated magneto-optical isolator and circulator on silicon photonics circuits using a direct bonding of silicon and a magneto-optic garnet. The devices are based on a Mach-Zehnder interferometer (MZI) composed of Si waveguides. Bonded magnet-optic garnet as an upper cladding layer induces nonreciprocal phase shift in a push-pull manner by applying magnetic fields in the Voigt configuration. Depending on the number of input and output ports, an MZI becomes an isolator or a circulator. In our recent development, an optical isolator with 30 dB isolation and a 4-port circulator are demonstrated.

1. INTRODUCTION

Nonreciprocal devices are indispensable for highly-functional optical communication systems. An optical isolator protects laser diodes from unwanted reflections to keep the stable operation. An optical circulator is used to divide optical signals in bi-directional transmissions or to form an add-drop multiplexer with fiber Bragg gratings. Commercially available isolators and circulators based on the Faraday rotation are formed by bulk optics. Waveguide type optical isolators and circulators are examined on magneto-optic garnets [1–5], however, they have not yet been used in practical systems due to a lack of compatibility with other components. This is because the magneto-optic garnet must be epitaxially grown on a garnet substrate so as to be single crystal providing a large magneto-optic effect with relatively low optical absorption at a telecom wavelength.

Our approach to overcome this issue is to use a direct bonding of the single-crystalline garnet onto another optical waveguide platform. A nonreciprocal optical device with semiconductor platform partially bonded magneto-optic garnet can be integrated with other devices. Using the nonreciprocal phase shift which is the first order magneto-optic effect in the Voigt configurations is more suitable for waveguide devices than using the Faraday rotation which requires strict phase matching between TE and TM modes. Our device is based on a Mach-Zehnder interferometer (MZI) with nonreciprocal and reciprocal phase shifters. The nonreciprocal phase difference is provided in a push-pull manner by applying magnet field in anti-parallel directions. The reciprocal phase difference is provided by path length difference of optical waveguides.

We have demonstrated a silicon-based optical isolator with a high isolation over 30 dB [8]. Also, silicon-based 4-port circulator was demonstrated by using a 2×2 MZI configuration [9]. In this paper, we describe the recent progress of magneto-optical isolator and circulator based on silicon waveguides.

2. DEVICE STRUCTURE

Figure 1 shows schematic of silicon-based MZI optical isolator and circulator. The center ports of 3 dB coupler are input and output ports. A magneto-optic garnet Ce-substituted yttrium iron garnet (Ce:YIG) is bonded on the MZI as an upper cladding layer. A nonreciprocal phase shift is induced by the magneto-optic effect in Ce:YIG by applying magnetic fields to the arms in anti-parallel directions. Also, a reciprocal phase shift is induced by optical path difference between the MZI arms. The nonreciprocal phase shifter provides a $\pm\pi/2$ phase difference depending on the propagation direction. This is cancelled for the forward direction with a reciprocal phase difference of $\pi/2$, but is added for the backward direction. The forward light exhibits in-phase interference, while the backward light exhibits anti-phase interference. The light exhibiting in-phase interference coupled to the center port. The light exhibiting anti-phase interference is radiated to the side ports. Then, the MZI functions as an optical isolator.

In the same way, when the 3 dB directional coupler is composed of 2×2 ports, in- and anti-phase interferences select the transmission ports of the MZI. That is, the light transmits from port 1 to port 2 and from port 3 to port 4 for in-phase interference, on the other hand, from port 2 to port 3 and from port 4 to port 1 for anti-phase interference, respectively. Then, the MZI functions as an optical circulator.

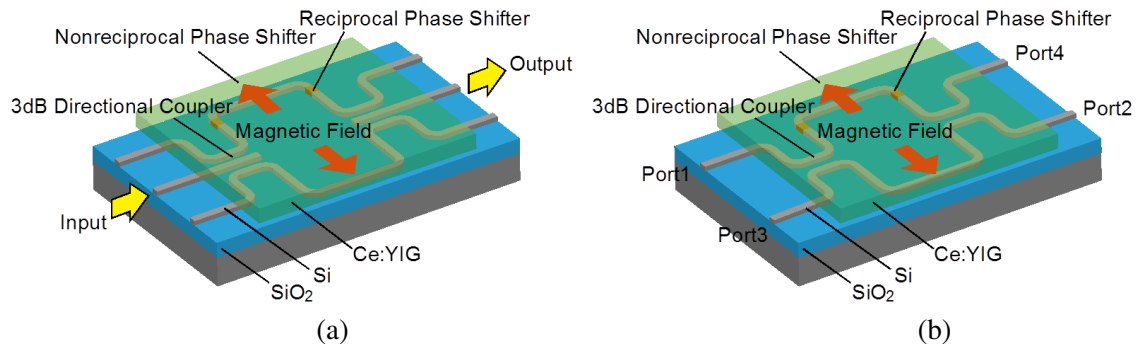


Figure 1: Schematics of (a) MZI optical isolator and (b) MZI optical circulator with silicon waveguides.

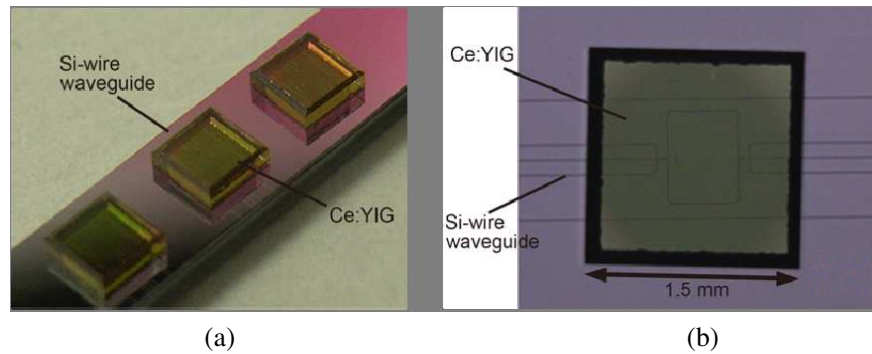


Figure 2: Photos of fabricated silicon waveguide optical isolators.

The nonreciprocal phase shift is a directionally dependent propagation constant and can be simulated by solving optical mode field and using perturbation theory [5, 6]. This shift is brought about by the directionally dependent dielectric permittivity of the magneto-optic garnet, where the off-diagonal elements change its sign due to the first-order magneto-optic effect. The length of the nonreciprocal phase shifter is designed to be $400\ \mu\text{m}$ for a silicon waveguide dimension of $450 \times 220\ \text{nm}^2$ and the Faraday rotation coefficient of $-4500\ \text{deg/cm}$ of Ce:YIG at a wavelength of $1550\ \text{nm}$.

Silicon waveguides were fabricated by electron-beam lithography and reactive ion etching technique. A single crystalline Ce:YIG layer was grown on a gadolinium garnet substrate by RF sputtering deposition technique. In this work, a $1.5 \times 1.5\ \text{mm}^2$ Ce:YIG die was bonded on the silicon waveguide using a plasma assisted direct bonding technique [7]. The surface of a patterned silicon layer and Ce:YIG die were exposed to N_2 plasma for 10 sec. Then, they were brought into contact and pressed at $6\ \text{MPa}$ at 200°C in a vacuum chamber. Figure 2 shows the photos of fabricated silicon waveguide optical isolators. Although the bonded region should be only on the nonreciprocal phase shifter, it is difficult to manipulate such a small die to place on the precise position. In this case, all the MZI was covered by the Ce:YIG die.

3. CHARACTERIZATION

The fabricated isolator was characterized by measuring the transmission spectra in a $1550\ \text{nm}$ wavelength range as shown in Figure 3 [8]. The resonant spectra are due to the optical path difference of the reciprocal phase shifter, which was designed much longer than that for minimum length to be $\pi/2$ so that the wavelength shift can be easily observed. By using a compact permanent magnet with three reversed poles, an external magnetic field was applied to the MZI arms in anti-parallel directions. The dashed line indicates the transmittance without applying a magnetic field. When the magnetic field was applied, the transmission spectra exhibited wavelength shifts with different signs depending on the light propagation direction as shown by the red and blue solid lines. A maximum isolation ratio of $30\ \text{dB}$, which was defined by the difference in transmittance between two propagation directions, was obtained at a wavelength of $1548\ \text{nm}$. This is the highest isolation among silicon waveguide optical isolators ever reported.

Figure 4 shows the measured transmittance spectra of silicon waveguide circulator [9]. The

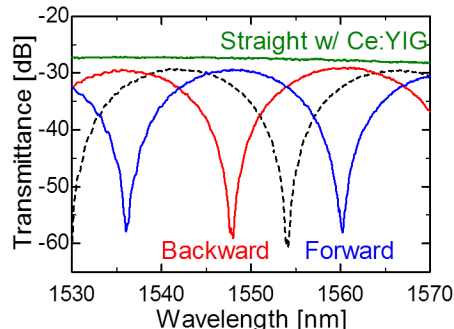


Figure 3: Measured transmittance of silicon waveguide isolator.

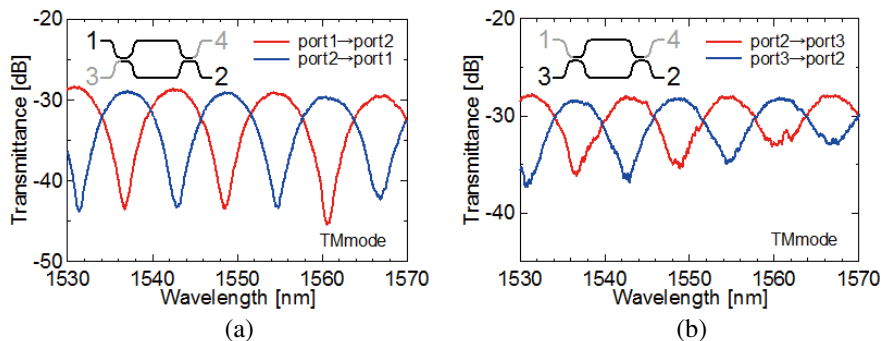


Figure 4: Measured transmittance spectra of a circulator with port pairs of (a) port 1-2 and (b) port 2-3.

Table 1: Measured transmittance for input/output port combinations at a wavelength of 1531 nm.

Input	Transmittance measured at output [dB]			
	Port 1	Port 2	Port 3	Port 4
Port 1	-	-28.4	-	-34.5
Port 2	-43.7	-	-28.0	-
Port 3	-	-37.3	-	-29.0
Port 4	-27.8	-	-42.1	-

transmittance measured at 1531 nm is summarized in Table 1 for four input/output combinations. Port 2, 3, 4, and 1 are conducting ports for the input from port 1, 2, 3, and 4, respectively, at this wavelength. On the other hand, port 4, 1, 2, and 3 are isolating ports for the input from port 1, 2, 3, and 4, respectively. Isolations defined as the difference in transmittance between two ports are better for cross port pairs 1-2 and 3-4 than bar port pairs 2-3 and 1-4. This is because an exact 3 dB dividing characteristic was not provided in the directional couplers. Since the deviations in two couplers are almost identical in actual fabrication, power imbalance due to the first coupler is nullified by the second coupler for the cross port. On the other hand, power imbalances are added for the bar port. Therefore, better isolations are obtained for the cross port pairs. A maximum isolation of 15.3 dB was obtained for the pair of port 1-2.

Let us discuss about the loss of these devices. The measured transmittance included a coupling loss of 8 dB/facet between fiber and silicon waveguide. The propagation loss of an air clad silicon waveguide is estimated to be 2.7 dB/cm at a wavelength of 1550 nm. The measured insertion loss of a straight waveguide adjacent to the device, shown by green line in Figure 3, is estimated from optical simulations to include the junction losses at the boundary between air and Ce:YIG cladding regions of 7.6 dB and the optical absorption of Ce:YIG of 4.1 dB. The junction loss can be reduced to be less than 2 dB by covering the air cladding region with SiO₂. The optical absorption of Ce:YIG can be reduced to be about 1 dB by annealing the Ce:YIG after the growth. The difference between the transmittance of straight waveguide and that of maximum peak of the MZI was 3 dB, which is due to excess loss in directional couplers by a fabrication error.

4. CONCLUSIONS

We have fabricated and demonstrated silicon-based optical isolator and circulator using direct bonding technology. A magneto-optic garnet Ce:YIG is directly bonded on a silicon waveguide without any adhesive. A magneto-optic effect induces nonreciprocal phase shift in an MZI, which enables nonreciprocal propagations. So far high isolation of 30 dB is the highest record among waveguide optical isolators ever reported. Also, a 4 port circulator with silicon waveguides has been demonstrated for the first time.

REFERENCES

1. Ando, K., T. Okoshi, and N. Koshizuka, “Waveguide magneto-optical isolator fabricated by laser annealing,” *Appl. Phys. Lett.*, Vol. 53, No. 1, 4–6, 1988.
2. Shintaku, T., “Integrated optical isolator based on efficient nonreciprocal radiation mode conversion,” *Appl. Phys. Lett.*, Vol. 73, No. 14, 1946–1948, 1998.
3. Sugimoto, N., T. Shintaku, A. Tate, M. Shimokozono, E. Kubota, M. Ishii, and Y. Inoue, “Waveguide polarization-independent optical circulator,” *IEEE Photon. Technol. Lett.*, Vol. 11, No. 3, 355–357, 1999.
4. Fujita, J., M. Levy, R. M. Osgood, Jr., L. Wilkens, and H. Dötsch, “Waveguide optical isolator based on Mach-Zehnder interferometer,” *Appl. Phys. Lett.*, Vol. 76, No. 16, 2158–2160, 2000.
5. Dötsch, H., N. Bahlmann, O. Zhuromskyy, M. Hammer, L. Wilkens, R. Gerhardt, P. Hertel, and A. F. Popkov, “Applications of magneto-optical waveguides in integrated optics: Review,” *J. Opt. Soc. Am. B*, Vol. 22, No. 1, 240–253, 2005.
6. Okamura, Y., H. Inuzuka, T. Kikuchi, and S. Yamamoto, “Nonreciprocal propagation in magneto-optic YIG rib waveguides,” *J. Lightwave Technol.*, Vol. 4, No. 7, 711–714, 1986.
7. Mizumoto, T., Y. Shoji, and R. Takei, “Direct wafer bonding and its application to waveguide optical isolators,” *Materials 2012*, Vol. 5, No. 5, 985–1004, 2012.
8. Shirato, Y., Y. Shoji, and T. Mizumoto, “High isolation in silicon waveguide optical isolator employing nonreciprocal phase shift,” *OFC/NFOEC 2013*, Anaheim, USA, Mar. 2013, OTu2C.5.
9. Mitsuya, K., Y. Shoji, and T. Mizumoto, “Demonstration of a silicon waveguide optical circulator,” *IEEE Photon. Technol. Lett.*, Vol. 25, No. 8, 721–723, 2013.

Quantum Electrodynamics in the Presence of Moving Bodies

S. A. R. Horsley

University of Exeter, United Kingdom

Abstract— The vacuum state of the electromagnetic field close to a dielectric body is not the same in all reference frames. Here we show that this has some surprising consequences. For instance, if we attempt to quantize electromagnetism interacting with a uniformly moving medium (properly including effects of dispersion and dissipation), then the resulting Hamiltonian lacks a lower bound. Consequently, a detector placed close to a moving surface can be excited by the vacuum! It is shown that when the dielectric is allowed to change velocity then the Hamiltonian again becomes bounded from below, thus demonstrating that the source of the energy exciting the detector is the kinetic energy of the dielectric.

1. INTRODUCTION

It is shown in Landau and Lifshitz' first book on statistical physics, that the state of maximum entropy for any system of moving macroscopic bodies is when those bodies are at relative rest [1]. However, it is not shown how the bodies get into this situation. At finite temperature, it seems that radiation pressure can do the job: each body emits an isotropic field in its rest frame which appears anisotropic when the body moves. This anisotropy causes an asymmetry in the radiation pressure between the bodies, which acts to bring the bodies to relative rest (for details see [2, 3]). However, the limiting behaviour of this interaction as the temperature of is reduced to zero is not so immediately clear, and a variety of results have been obtained using different approaches (e.g., [4–6]). To obtain a clear understanding of this limit requires an exact canonical treatment of quantum electrodynamics in the presence of moving dielectric bodies. Here we provide the details of such a treatment, illustrating some surprising features.

2. QUANTIZATION

It seems that a correct quantum mechanical description of a system must follow from an action principle. We look for an action that reproduces the classical macroscopic Maxwell equations for electromagnetism interacting with a moving dielectric, and properly includes the phenomena of dispersion and dissipation (the Kramers-Krönig relations). Huttner and Barnett [7] and Philbin [8] previously showed how to do this for stationary objects. In [9, 10] this was extended to objects in motion, where the Lagrangian density was found to be of the form,

$$\mathcal{L} = \mathcal{L}_F + \mathcal{L}_{INT} + \mathcal{L}_R \quad (1)$$

The first term is the Lagrangian density associated with the free electromagnetic field, $\mathcal{L}_F = \epsilon_0 [E^2 - c^2 B^2]/2$, and the final term,

$$\mathcal{L}_R = \frac{1}{2} \int_0^\infty d\omega \left[\left(\frac{\partial \mathbf{X}_\omega(\mathbf{x}, t)}{\partial t} + V \frac{\partial \mathbf{X}_\omega(\mathbf{x}, t)}{\partial x} \right)^2 - \omega^2 \mathbf{X}_\omega(\mathbf{x}, t)^2 \right] \quad (2)$$

represents a field (a reservoir) of simple harmonic oscillators, \mathbf{X}_ω that mimic the linear response of the moving dielectric to the field (velocity $\mathbf{V} = V\hat{\mathbf{x}}$). The interaction between the two is characterized by an interaction Lagrangian,

$$\mathcal{L}_{INT} = (\mathbf{E} + V\hat{\mathbf{x}} \times \mathbf{B}) \cdot \int_0^\infty \sqrt{\frac{2\omega \text{Im} [\epsilon(\mathbf{x}_\parallel, \omega)]}{\pi}} \mathbf{X}_\omega d\omega \quad (3)$$

where for simplicity of presentation it has here been assumed that the dielectric is without magnetic response ($\mu = 1$), and that the medium is only inhomogeneous in the plane normal to the direction of motion, with coordinates in this plane given by the two component vector, \mathbf{x}_\parallel . The classical equations of motion that follow from (1) are the macroscopic Maxwell equations for a dielectric slab in uniform motion along \hat{x} so long as the permittivity, ϵ — defined in the rest frame — satisfies the Kramers-Krönig relations. The Hamiltonian operator constructed from (1) is quadratic in

the fields and can thus be diagonalized into a continuum of normal modes. In [10], the required transformation was found, and the diagonalized Hamiltonian is,

$$\hat{H} = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \int d^2\mathbf{x}_{||} \hbar\omega_+ \hat{\mathbf{C}}_{\omega}^{\dagger}(k, \mathbf{x}_{||}) \cdot \hat{\mathbf{C}}_{\omega}(k, \mathbf{x}_{||}) \quad (4)$$

where $\omega_+ = \omega + Vk$. The $\hat{\mathbf{C}}_{\omega}$ operators appearing in (4) obey bosonic commutation relations and are combinations of both the electromagnetic field and reservoir operators. This is a very strange Hamiltonian, for the normal mode frequencies range from $-\infty$ to ∞ , rather than beginning at zero. However (4) seems to be correct, for when $V = 0$ it reduces to the result of [7, 8], and it is therefore clear that the diagonalization is simply telling us that the normal modes are Doppler shifted when the dielectric moves, which could have been guessed without following the full quantization procedure. In [10, 11] it is shown that if the velocity of the dielectric is included as a dynamical variable, then the Hamiltonian becomes bounded from below. One can therefore interpret these negative frequency normal modes as an accounting device for the fact that once it interacts with another system which is in relative motion, some external input is required to keep a medium moving at a constant velocity.

3. A DETECTOR CLOSE TO A MOVING SURFACE

If the vacuum is defined as the zero particle state, $\hat{\mathbf{C}}_{\omega}|0\rangle = 0$, then (4) tells us that it is not the lowest energy state of the system of field plus dielectric. Therefore energy can be extracted from the vacuum electromagnetic field outside of a moving dielectric when, for example, a two level system at rest and in its ground state is brought close to the surface (see Figure 1). In [10] it is found that the transition rate of this ‘detector’, $\Gamma_{0 \rightarrow 1}$ is given by the expression,

$$\Gamma_{0 \rightarrow 1} = \frac{2\omega^2}{\hbar} \int_{\omega/V}^{\infty} \mathbf{d}_0 \cdot \text{Im}[\mathbf{G}(-k, \mathbf{x}_{||}, \mathbf{x}_{||}, -\omega)] \cdot \mathbf{d}_0 \frac{dk}{2\pi} \quad (5)$$

where ω is the transition frequency of the detector, $\mathbf{x}_{||0}$ is the position of the detector in the plane normal to the velocity, \mathbf{d}_0 characterises the strength of the interaction between field and detector, and \mathbf{G} is the electromagnetic Green function. It is shown in [10] that explicitly evaluating (5) in terms of reflection coefficients gives expressions which are analogous to those of [5], which were used to calculate the rate of dissipation of energy between two dielectric plates in relative motion. Given that $V \ll c$, (5) shows that the excitation of the detector is caused by a sum over waves which exponentially decay from the surface into free space (e.g., plasmons, as considered in [5, 12]). Indeed unless the frequency of the detector is low (e.g., GHz), then the transition rate will be extremely small for reasonable distances from the dielectric, due to the fact that one is summing over extremely tightly bound surface waves. Nevertheless it may be measurable if one can operate at low enough temperatures and find a detector that operates at such low frequencies (perhaps flux qubits would be suitable for this purpose [13]).

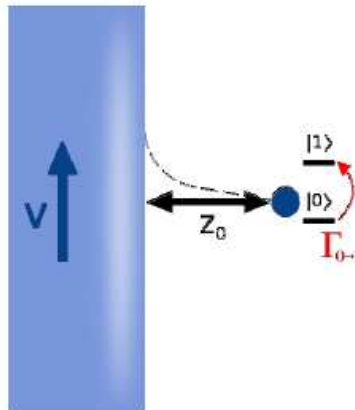


Figure 1: A dielectric slab moves up the x -axis at $\mathbf{V} = V\hat{\mathbf{x}}$. If a detector is positioned a distance z_0 away from the surface, with both field and detector initially in the ground state.

4. CONCLUSION

An outline has been given for the procedure necessary to quantize the electromagnetic field in the presence of moving bodies. It was found that if one fixes the velocity of a dielectric body by some external means, then the system of field plus dielectric must be described using a Hamiltonian that is not bounded from below. A detector placed outside such a moving surface can be excited by the vacuum field at a rate given by a sum over all waves bound to the surface of the medium.

ACKNOWLEDGMENT

The author thanks the EPSRC for financial support.

REFERENCES

1. Landau, L. D. and E. M. Lifshitz, *Statistical Physics — Part 1*, Butterworth-Heinemann, 2005.
2. Einstein, A., “On the quantum theory of radiation,” *Phys. Zs.*, Vol. 18, 121, 1917.
3. Mkrtchian, V., V. Adrian Parsegian, R. Podgornik, and W. M. Saslow, “Universal thermal radiation drag on neutral objects,” *Phys. Rev. Lett.*, Vol. 91, 220801, 2003.
4. Teodorovich, E. V., “On the contribution of macroscopic van der Waals interactions to frictional force,” *Proc. R. Soc. Lond. A*, Vol. 362, 71, 1978.
5. Pendry, J. B., “Shearing the vacuum — Quantum friction,” *J. Phys.: Cond. Mat.*, Vol. 9, 10301, 1997.
6. Philbin, T. G. and U. Leonhardt, “No quantum friction between uniformly moving plates,” *New J. Phys.*, Vol. 11, 033035, 2009.
7. Huttner, B. and S. M. Barnett, “Quantization of the electromagnetic field in dielectrics,” *Phys. Rev. A*, Vol. 46, 4306, 1992.
8. Philbin, T. G., “Canonical quantization of macroscopic electromagnetism,” *New J. Phys.*, Vol. 12, 123008, 2010.
9. Horsley, S. A. R., “Consistency of certain constitutive relations with quantum electromagnetism,” *Phys. Rev. A*, Vol. 84, 063822, 2011.
10. Horsley, S. A. R., “Canonical quantization of the electromagnetic field interacting with a moving dielectric medium,” *Phys. Rev. A*, Vol. 86, 023830, 2012.
11. Horsley, S. A. R., “Motion of macroscopic bodies in the electromagnetic field,” *Quantum Physics*, 1–15, 2013.
12. Scheel, S. and S. Y. Buhmann, “Casimir-Polder forces on moving atoms,” *Quantum Physics*, Vol. 80, 1–11, 042902, 2009.
13. Nakamura, Y., Y. Pashkin, and J. Tsai, “Quantum oscillations in two coupled charge qubits,” *Nature*, Vol. 398, 786, 1999.

Curve-fitting Formulas for Fast Determination of Frequency Band-notched Response of UWB Antennas

Ayman Ayd R. Saad¹, Mohamed Mamdouh M. Ali², and Elsayed Esam M. Khaled²

¹Kosseir Radio, Telecom Egypt, Kosseir 84712, Egypt

²Electrical Engineering Department, Assiut University, Assiut, Egypt

Abstract— This paper presents approximated formulas for fast determination of frequency band-notched response of ultra-wideband (UWB) antennas loaded with nearly quarter-/half-wavelength notch resonators. The closed-form formulas are derived using curve-fitting technique. They describe the influences of the physical length of these notch resonators on the corresponding notch frequencies in the UWB frequency band of 3.1–10.6 GHz. The calculated results obtained by using these new formulas show good correlation with the reported electromagnetic (EM) simulation results elsewhere.

1. INTRODUCTION

Since the Federal Communication Commission (FCC) of the US assigned 3.1–10.6 GHz frequency band of UWB systems in February 2002 [1], UWB technology has been attracting considerable interests in both the academic and commercial domains due to the potentially high data rate (more than 110 Mbits/s) for short range, low power consumptions and easy connections to different devices such as wireless USB, PCs, high-definition TVs, etc.. Consequently, UWB antennas have received more and more attention, as the only non-digital part of UWB system.

The use of the UWB antennas have more challenges to meet, namely, stable radiation pattern, gain, and group delay, all over the band. Covering this ultra wide bandwidth arouses a coexistence interference problem with narrowband technologies sharing with UWB some of the frequency bands such as WiMAX operating in the band 3.3–3.7 GHz and Wireless LAN operating in the band 5–6 GHz. To overcome this problem and avoiding interference, UWB antennas use filters to suppress these dispensable bands. An alternative approach to notch-out specific frequencies is to design UWB antenna with frequency band-notched characteristics.

With the development of UWB antenna having frequency band-notched features, literature is congested with different designs in various topologies. However, the most popular method to obtain the notched band is to insert resonator such as a slot or a parasitic strip. The length of that resonator may appear to be a nearly quarter-/half-wavelength of resonance frequency [2].

The aim of this paper is to present approximated closed-form expressions that help researchers and UWB antenna designers to fast determine the frequency band-notched response of the nearly quarter-/half-wavelength resonators, introduced in the UWB antennas. The formulas are obtained by means of curve-fitting technique which gives the best fit equation with the available results [3]. The formulas describe a relationship between the total length of a resonator and the corresponding notched frequency. To verify the proposed approach, the results obtained from the approximated formulas are compared with the simulation results obtained previously in the literature EM simulation for band-notched UWB antennas. The validity of the proposed technique is verified and high accuracy results are obtained.

2. ANALYSIS AND FORMULATION

In order to find the best fit expression, various sets of simulation traces data are required. In Fig. 1, the geometry of the proposed band-notched UWB antenna utilized to extract the required band-notch response data is shown. The details of the prototype antenna design can be found in [4]. For this study, in order to obtain the band-notched property, various nearly quarter-/half-wavelength resonators etched off the radiating patch as well as the microstrip feeding line are used. The radiating patch is integrated with either one of nearly half-wavelength slots corresponding to 5.5 GHz frequency notch such as a single complementary split-ring resonator (CSRR), as shown in Fig. 1(b) or a reversed U-shaped slot, as shown in Fig. 1(c). Also, a quarter-wavelength slot can be used corresponding to 3.5 GHz frequency notch such as a flipped L-shaped spur-line slot, as shown in Fig. 1(d). The microstrip feeding line can also be loaded with either a reversed U-shaped slot, as shown in Fig. 1(e) or a flipped L-shape spur-line slot, as shown in Fig. 1(f), which are corresponding

to nearly half- or quarter-wavelength resonators at 5.5 GHz notched frequency, respectively. The physical lengths of these resonators are used to optimize the band-rejected performance and the antenna's operating bandwidth.

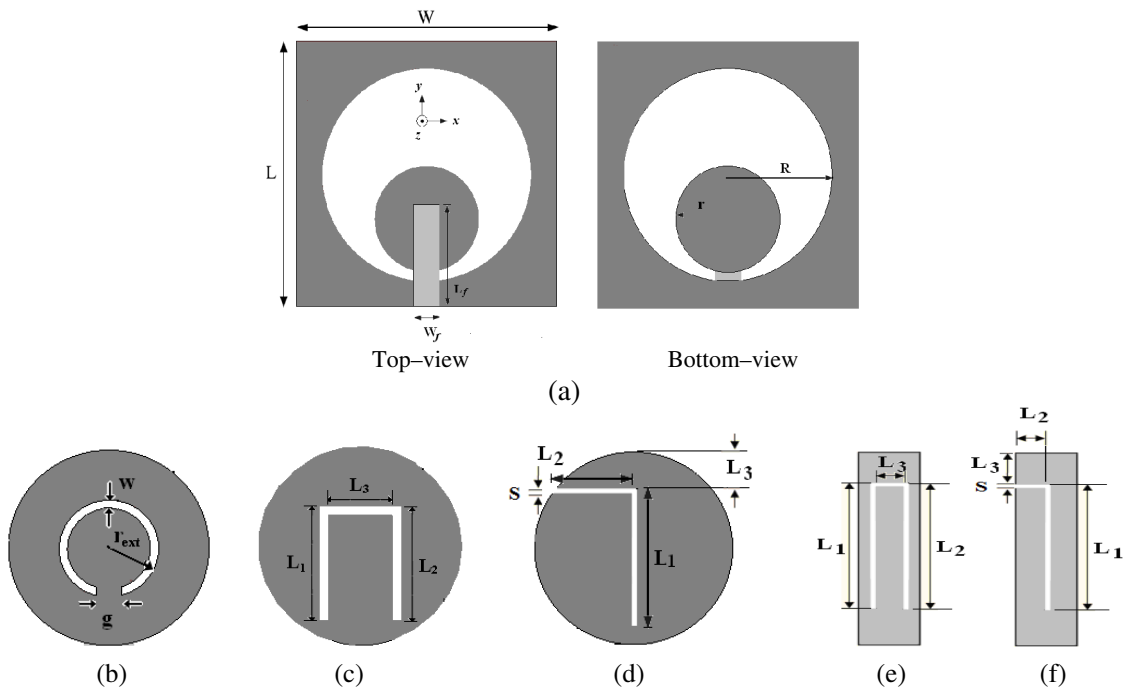


Figure 1: (a) Geometry of proposed antenna used in this study [4] ($L = 30$ mm, $W = 30$ mm, $R = 12$ mm, $r = 6$ mm, $L_f = 11.5$ mm, $W_f = 3$ mm). (b) A single CSRR etched off the radiating patch. (c) A reversed U-shaped slot etched off the radiating patch. (d) A flipped L-shape spur-line slot etched off the radiating patch. (e) A reversed U-shaped slot etched off the microstrip feeding line. (f) A flipped L-shape spur-line slot etched off the microstrip feeding line.

The external radius of the single CSRR element etched off the radiating patch is $r_{ext} = 3.0$ mm, its gap is $g = 1.5$ mm. The ring is of uniform width $w = 0.5$ mm. The three sides of the reversed U-shaped slot are L_1 , L_2 and L_3 . These lengths are optimized to control the band-notch performance. The reversed U-shaped slot etched off the radiating patch is of uniform width 0.5 mm and of dimensions of $L_1 = L_2 = 7$ mm, and $L_3 = 3.5$ mm. A gap of width 0.25 mm is considered between the reversed U-slot etched off the radiating patch and the left, upper, and right edges of the microstrip feeding line which is located on the other side of the substrate. On the other hand, the dimensions of the reversed U-shaped slot concentrically etched off the microstrip feeding line is $L_1 = L_2 = 7.5$ mm, and $L_3 = 1.5$ mm of uniform width 0.25 mm. The alternative technique to achieve a well defined rejection band uses a spur-line filter, which is realized by etching off a single flipped L-shape slot in the radiating patch as well as the microstrip feeding line. The good feature of a spur-line filter is that its physical structure is completely contained within the boundaries of the embedding structure. The flipped L-shape spur-line slot has four parameters which can be used to optimize the band-notch performance, which are, the length L_1 , the height L_2 , the location from the top edge of the embedding structure, L_3 and the width S of the slot.

The four parameters of the flipped L-shape spur-line slot etched off the radiating patch is $L_1 = 9.1$ mm, $L_2 = 4.6$ mm, $L_3 = 2$ mm, and $S = 1$ mm. On the other hand, the four parameters of the flipped L-shape spur-line slot etched off the microstrip feeding line is $L_1 = 5.75$ mm, $L_2 = L_3 = 1.75$ mm, and $S = 0.5$ mm. It is noticed that, the slot length of the spur-line, L_1 is approximately a quarter of the wavelength at the desired stop frequency band, measured in the microstrip line material [5].

The proposed antenna with the notched parameters under consideration was simulated using CST Microwave Studio software program [6]. All simulation results of parametric studies carried out on the proposed antenna with either one of the half-or quarter-wavelength of the proposed resonators are presented to demonstrate the influences of notched structures on the corresponding frequency response of the band-notch. Also, these parametric studies are used to derive the curve-

fitting formulas for the frequency band-notched response

Figures 2 and 3 address the effect of the total length of each resonator on the frequency response of the corresponding band-notch. It is found that the notched frequency is very sensitive to the total length of the resonator, L_t [4]. As L_t decreases the corresponding band notch increases. By adjusting the slot length, the notched frequency band within the antenna's operating bandwidth can be easily controlled.

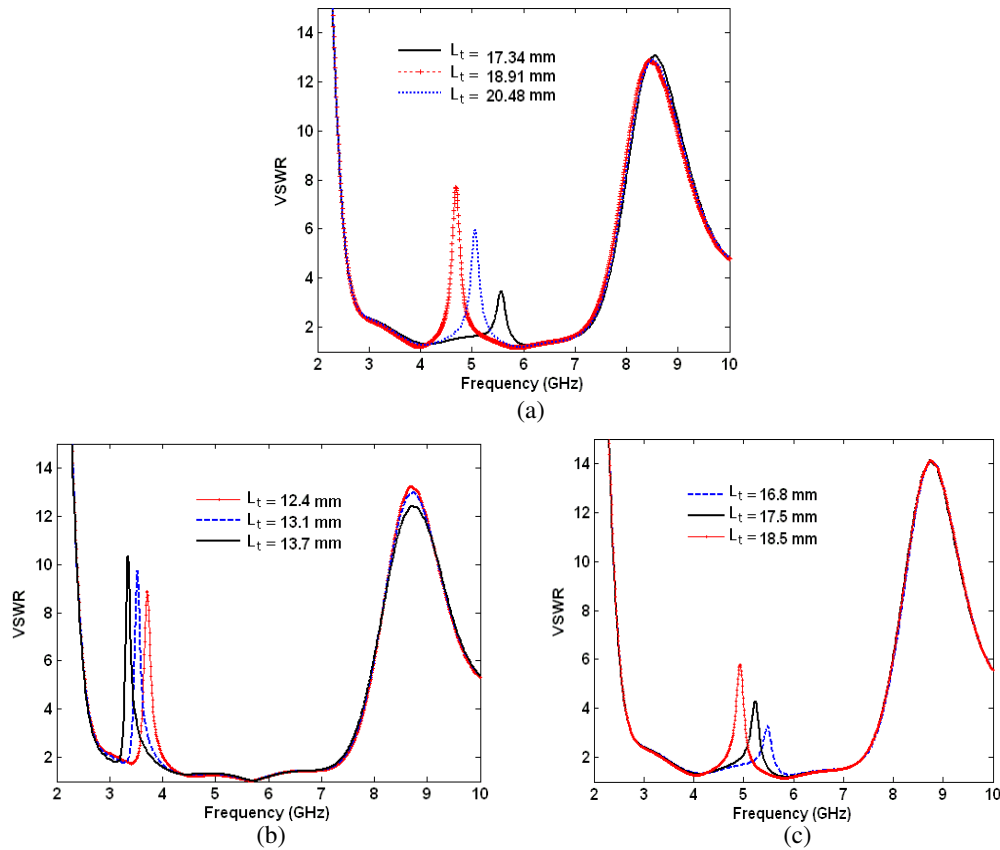


Figure 2: Effect of the total length of the slot resonators etched off the radiating patch on the band-notch response of proposed antenna. (a) A single CSRR element. (b) A flipped L-shape spur-line slot. (c) A reversed U-shaped slot.

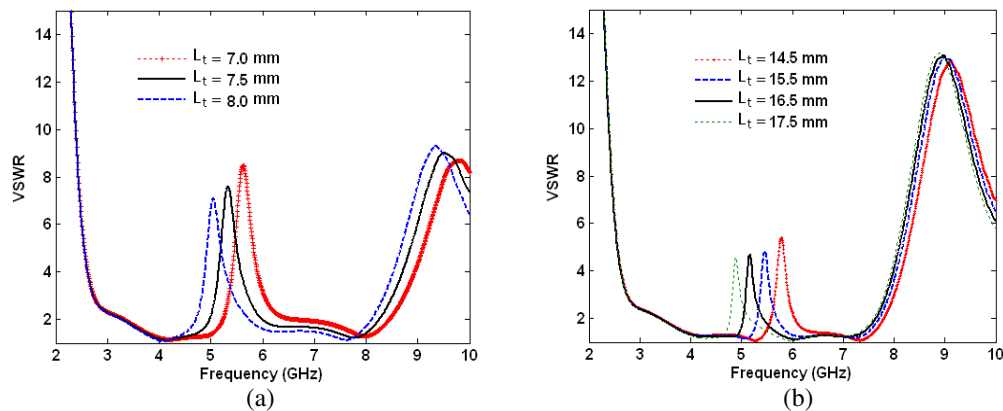


Figure 3: Effect of the total length of the slot resonators etched off the microstrip feed line on the band-notch response of proposed antenna. (a) A flipped L-shape spur-line slot. (b) A reversed U-shaped slot.

Based on these various sets of simulation traces in Figs. 2 and 3 and estimating lower and upper limits of each resonator response in the UWB frequency band, closed-form formulas for the total

length of the resonator, L_t as a function of frequency can be developed. This described process is carried out through the use of the principle of the curve-fitting polynomial expressions.

As a first order of approximation, the dimensions of the resonator can be chosen according to the following formula

$$L_t = \frac{\lambda_g}{n} \quad (1)$$

where L_t is the effective length of the resonator, and $n = 2$ corresponding to a half-wavelength and $n = 4$ corresponding to a quarter-wavelength. The wavelength can be approximately calculated by the formulas as follows [7]

$$\lambda_g = \frac{\lambda_0}{\sqrt{\varepsilon_{eff}}} \quad (2)$$

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2} \left[1 + 12 \frac{h}{w} \right]^{-1/2} \quad (3)$$

where λ_g and λ are the wavelength in the medium and in the free space, respectively ε_{eff} is the effective relative dielectric constant. Therefore the notch frequency, f_{notch} is given by

$$f_{notch} \approx \frac{c}{n * L_t \sqrt{\varepsilon_{eff}}} \quad (4)$$

Using the curve-fitting technique, the relationship between the notched frequency, f_{notch} and the total length of the nearly quarter-/half-wavelength resonators according to parametric studies can be approximated by second order polynomial as follows

$$L_t(\lambda/4) \approx 0.26 (f_{notch})^2 - 4.9 (f_{notch}) + 27 \quad 3 \leq f_{notch} \leq 11 \quad (5)$$

$$L_t(\lambda/2) \approx 0.51 (f_{notch})^2 - 9.7 (f_{notch}) + 55 \quad 3 \leq f_{notch} \leq 11 \quad (6)$$

where L_t in mm, and f_{notch} in GHz.

The proposed closed-form formulas to calculate the notched frequency are validated by comparing the calculated results with the simulated results obtained from the EM simulator. Figs. 4(a) and 4(b) show the comparisons between the calculated results obtained from Eqs. (5) and (6), and the simulated results of the notched frequency response for the nearly quarter-/half-wavelength notch resonators presented here and in literature [5, 8]. The proposed formulas are valid in limited space according to lower and upper limits of the resonator response in the UWB frequency band.

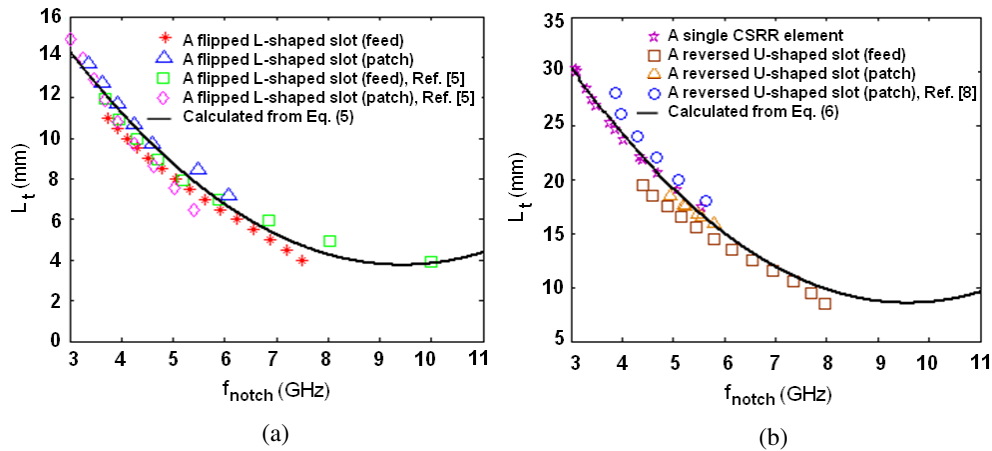


Figure 4: Comparison between the calculated and simulated values of the notched frequency for various total lengths, L_t of notched resonators presented here and in literature. (a) Nearly quarter-wavelength resonators. (b) Nearly half-wavelength resonators.

It can be observed that, the deviation between the slopes of the curve-fitting formulas and the slopes of the simulated results is very small and acceptable. With smoothly tunings the exact length of the notched resonator can be found.

3. CONCLUSION

In this paper, new and simple approximated closed-form formulas for computing the frequency band-notched response of UWB antennas loaded with nearly quarter-/half-wavelength notch resonators are presented. The formulas are obtained by means of the curve-fitting technique. They are useful for fast describing the influences of the total length of the notched resonators on the corresponding notched frequencies in the UWB frequency range through the use of a second order polynomial. Using this process, a fast and precise determination of frequency band-notched response of UWB antenna can be achieved without the need to perform the time-consuming tuning parametric studies on these notched structures. Calculated results accomplished with discussion are presented. The theoretical results obtained by using these new formulas are with very good agreement with the simulated results reported in the literature.

REFERENCES

1. Federal Communications Commission, "First report and order," Revision of Part 15 of the Commission's Rules Regarding Ultra-wideband Transmission Systems, February 2002.
2. Kim, C., *Ultra-wideband Antenna, Microwave and Millimeter Wave Technologies: Modern UWB Antennas and Equipment*, I. Minin, Ed., In-Tech, Croatia, 2010.
3. Bali, N. P. and M. Goyal, *A Textbook of Engineering Mathematics (Sem-III)*, Laxmi Publications Pvt Limited, 2006.
4. Khaled, E. E. M., A. A. R. Saad, and D. A. Salem, "A proximity-fed annular slot antenna with different a band-notched manipulations for ultra-wide band applications," *Progress In Electromagnetics Research B*, Vol. 37, 289–306, 2012.
5. Tilanthe, P., P. C. Sharma, and T. K. Bandopadhyay, "A monopole microstrip antenna with enhanced dual band rejection for UWB applications," *Progress In Electromagnetics Research B*, Vol. 38, 315–331, 2012.
6. CST Microwave Studio, Computer Simulation Technology, Framingham, MA, 2010.
7. Pozar, D. M., *Microwave Engineering*, 4th Edition, John Wiley & Sons, Inc., 2012.
8. Dissanayake, T. and K. P. Esselle, "Prediction of the notch frequency of slot loaded printed UWB antennas," *IEEE Trans. Antennas Propag.*, Vol. 55, No. 11, 3320–3325, November 2007.

A Novel Frequency-tunable Antenna with a Wide Tuning Range

Ningli Zhu and Wei Zhang

Institute of Microelectronics, Peking University, Beijing 100871, China

Abstract— In this paper a varactor-loaded antenna with a wide tuning range and good directivity is designed. The variable capacitance diode loading method is taking in detail, which enables the impedance matching from 1.9 GHz to 2.56 GHz for $VSWR \leq 2$. The HFSS and ADS software simulate this antenna independently, and the results show that it realizes a 500 MHz tunable range, from 1.44 GHz to 2.06 GHz in HFSS, and 1.944 GHz to 2.568 GHz in ADS, moreover, there exists a perfect matching resistor. At the same time, the shape of the H radiation pattern is nearly omnidirectional at those chosen frequency points. This paper uses the vector network analyzer to test the S parameter of the fabricated antenna. It shows that this antenna realizes 370 MHz tunable range, from 1.52 GHz to 1.89 GHz. This proposed design has a simple structure and a compact dimension of $24.2 \text{ mm} \times 32 \text{ mm} \times 1.5 \text{ mm}$.

1. INTRODUCTION

During the past decades and even more in the future, the demand for tunable antennas are gorgeous in communication applications and many military uses. There has been lots of researches about tunable antenna. One of the traditional way to load an antenna is using the PIN diodes [1, 2] which can decide the transmission part by its on or off diode state. Similarly, the newest RF MEMS switches [3] find many places in antennas as they have a great potential in terms of large capacitance ratio, low power consumption and better linearity. However, both of the two ways above cannot achieve continuous frequency tune range. What's more, the MEMS devices are relatively costly and not as accessible as common commercial tuning devices. On the contrary, antenna loaded by varactor diodes [4–6] shows ability to easily achieve continuous tuning and considerable size reduction. Moreover, varactor diodes are the devices accessible and economical.

In this paper, a frequency-tunable CPW Fed monopole antenna have a good wideband loading with a varactor is proposed [8, 9]. The radiation characteristics for this proposed antenna element such as return loss, bandwidth, voltage standing wave ratio (VSWR), E -plane and H -plane radiation patterns are obtained using HFSS simulation. This proposed design has a simple structure and a compact dimension of $24.2 \text{ mm} \times 32 \text{ mm} \times 1.5 \text{ mm}$. Vector network analyzer shows that the fabricated antenna realizes 370 MHz tunable range, from 1.52 GHz to 1.89 GHz.

2. CONFIGURATION AND WORKING PRINCIPLE

2.1. Configuration

The proposed antenna is composed of two parts, one is the radiation part, and another is the coplanar waveguide feedline. The connection of the two parts is a J slot [8].

To investigate the frequency reconfigurability of the antenna, a varactor diode is selected to cover the range of the required capacitance. The selected varactor diode SMV1405 has been manufactured by Skyworks Industries, and its equivalent circuit model is shown in the inset of Fig. 1(b). The capacitance of the intrinsic device can be approximately tuned from 0.63 pF to 2.67 pF when the bias voltage is varied from 30 V to 0 V, respectively. The value of the parasitics given by the manufacturer are $L_s = 0.7 \text{ nH}$, $C_p = 0.29 \text{ nH}$ and $R_s = 0.80 \Omega$. Fig. 1(c) shows the equivalent circuit model for the proposed antenna.

2.2. Working Principle

As can be seen from surface electric field in Fig. 2, the E -field across the CPW slots is asymmetric. This phenomenon indicates that the CPW odd mode as well as even mode is significantly influences the capability of this antenna. From this surface current picture, the horizontal line at the top of the antenna is not involved in the radiation, due to the odd symmetry of the distribution of the current. Accordingly, this narrow slit is equivalent to a loaded inductance. At the same time, the J-slot can be equivalent to the field current of the space radiation, which can be explained by the surface electric field. We can see the strongest radiation part and the electric field gradually attenuated along the J-slot.

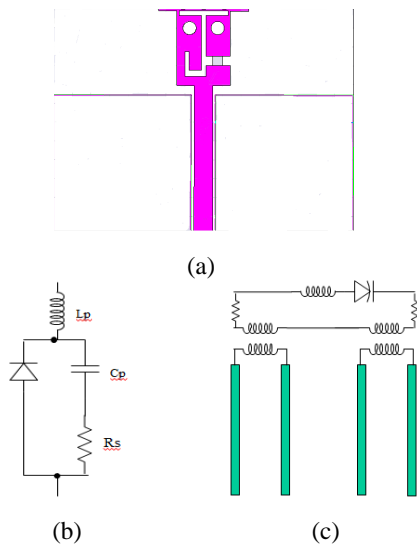


Figure 1: (a) Configuration of the proposed antenna. (b) Varactor circuit model. (c) Equivalent circuit model for the proposed antenna.

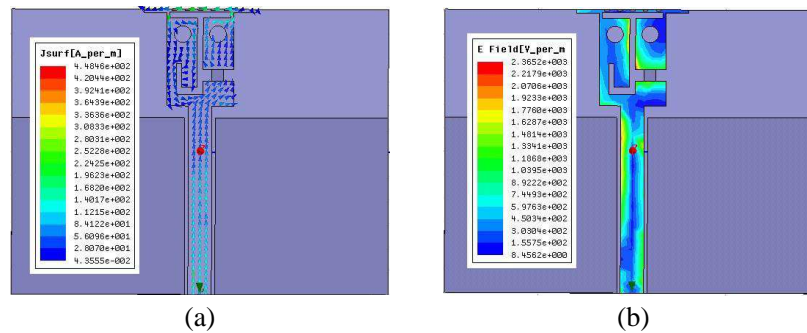


Figure 2: Surface current and electric field simulation in HFSS. (a) Surface current. (b) Surface electric field.

3. SIMULATION RESULTS

3.1. HFSS Simulation Results

From the HFSS simulation chart in Fig. 3, this antenna realizes a 600 MHz tuning range from 1.44 GHz to 2.06 GHz as the varactor diode transferred from 0.63 pF to 2.67 pF, and there is about 50 MHz bandwidth at -10 dB.

Figure 4 shows that there is an ideal value 50 Ohm matching at 2.08 GHz.

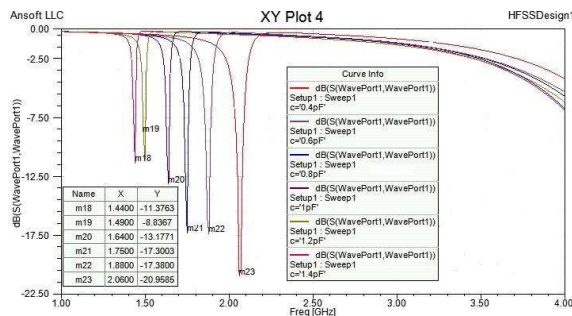


Figure 3: Simulated return loss of the proposed antenna.

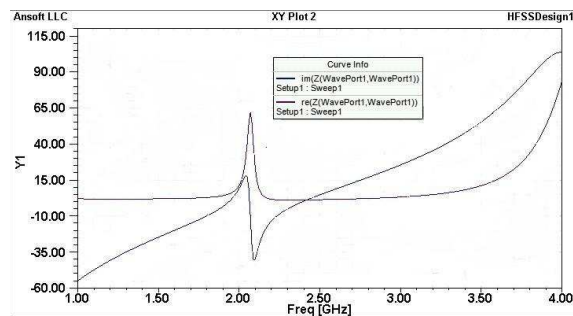


Figure 4: Resistor matching diagram.

From Fig. 5 to Fig. 7, we selected three frequency points, 2.06 GHz, 1.75 GHz and 1.44 GHz, to analyze its radiation pattern.

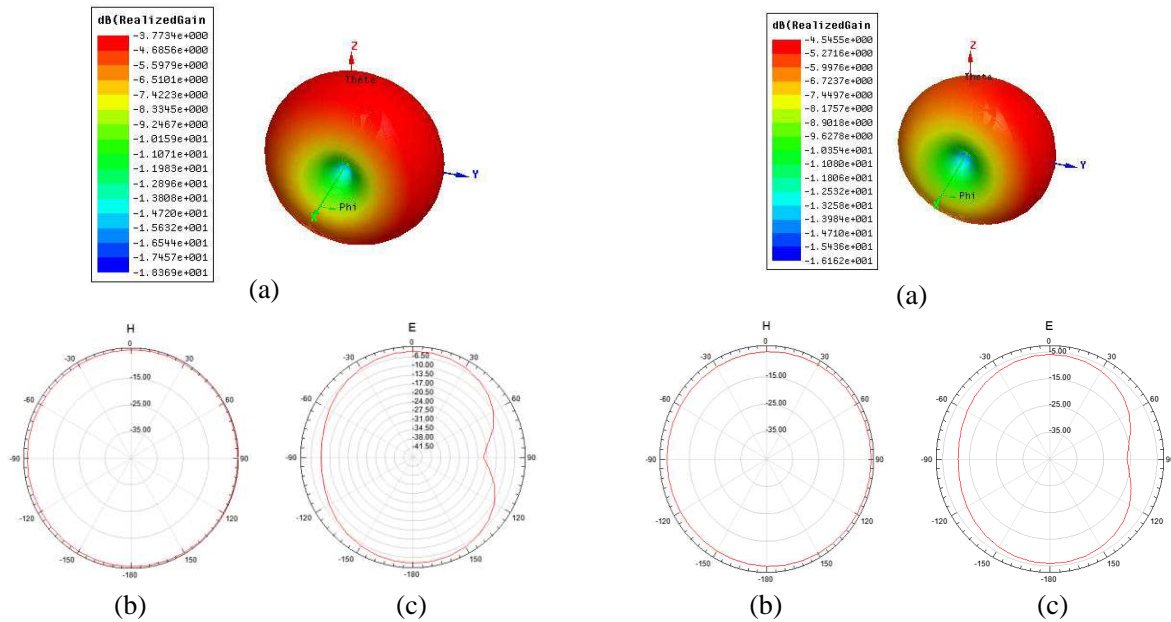


Figure 5: (a) 3D pattern of the operating frequency at 2.06 GHz. (b) Simulated radiation patterns in the H plane of the operating frequency at 2.06 GHz. (c) Simulated radiation patterns in the E plane of the operating frequency at 2.06 GHz.

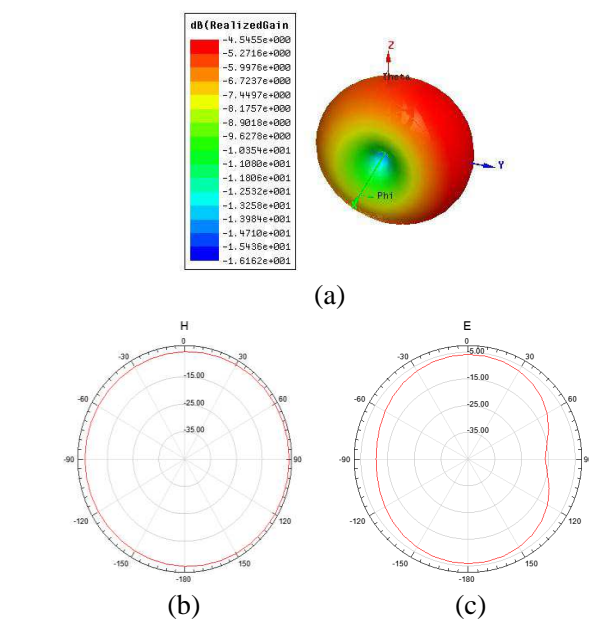


Figure 6: (a) 3D pattern of the operating frequency at 1.75 GHz. (b) Simulated radiation patterns in the H plane of the operating frequency at 1.75 GHz. (c) Simulated radiation patterns in the E plane of the operating frequency at 1.75 GHz.

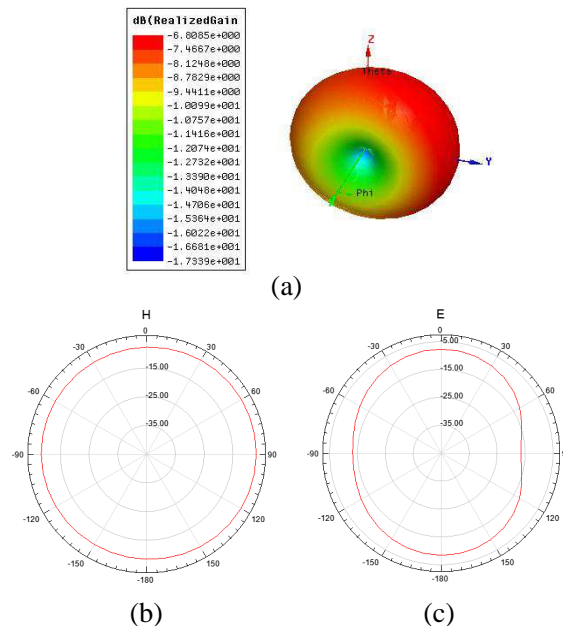


Figure 7: (a) 3D pattern of the operating frequency at 1.44 GHz. (b) Simulated radiation patterns in the H plane of the operating frequency at 1.44 GHz. (c) Simulated radiation patterns in the E plane of the operating frequency at 1.44 GHz.

From the above pictures, we can clearly see that the pattern in the H -plane is nearly a omnidirectional radiation, and the pattern in the E -plane is nearly an 8 figure shape.

3.2. ADS Simulation Results

Because the equivalent circuit parameters of the varactor diode will vary with frequency change, so we have to establish it in the ADS equivalent circuit simulation for a more accurate result.

In Fig. 8, We set a circuit model in the ADS software. The left one is 50 Ohm transmission

line, next by it is a three-port varactor which using the HFSS software to import. In the middle is the antenna. The lines above the antenna leads the impedance feeder which is loaded with the DC power supply. The S -parameters regulation can control the sweep range.

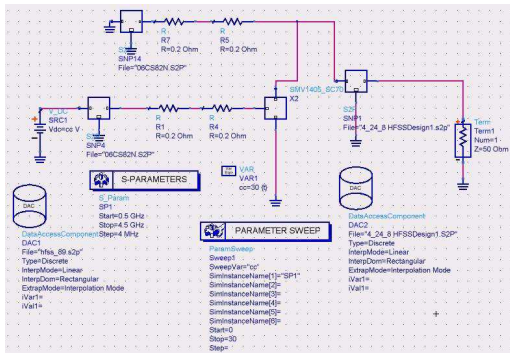


Figure 8: The ADS simulation with the varactor insertion loss.

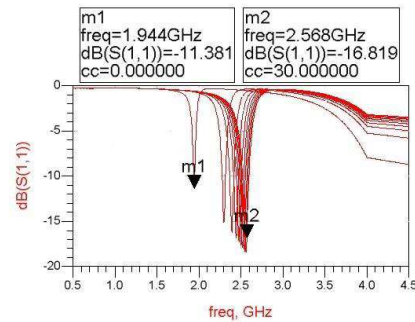


Figure 9: ADS simulation of return loss S_{11} .

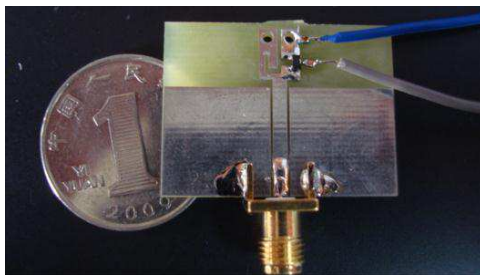


Figure 10: Image of the fabricated antenna.

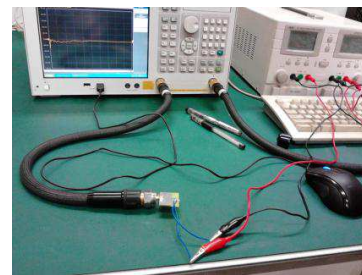


Figure 11: Testing scene.

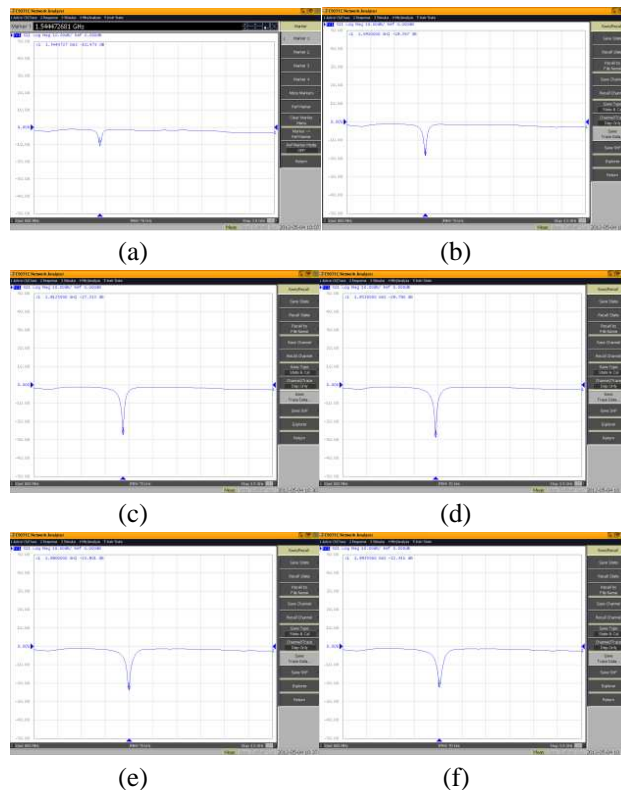


Figure 12: Original testing images of S_{11} by network analyzer. (a) Testing voltage at 4 V. (b) Testing voltage at 8.5 V. (c) Testing voltage at 15.5 V. (d) Testing voltage at 19 V. (e) Testing voltage at 22 V. (f) Testing voltage at 26 V.

From the simulation of ADS in Fig. 9, it shows that the antenna realizes a tunable range from 1.944 GHz to 2.568 GHz with a varactor diode insertion loss. Whats more, the antennas bandwidth in -10 dB is about 50 MHz and it realizes 600 MHz tunable range. Its performance is very good, so this design deserve to be processed.

4. FABRICATION RESULTS

The fabricated antenna shown in Fig. 10 uses a 1.5-mm-thick FR4 substrate with a relative permittivity (ϵ_r) of 4.4, loss tangent (δ) of 0.02, the size of it is $24.2 \text{ mm} \times 32 \text{ mm} \times 1.5 \text{ mm}$. However, due to the limitation of the processing technology which is accurate only to 1.5 mm, cannot accurate to 0.3 mm. Therefore the manufacturers set the width of the uppermost line of the antenna into 1.5 mm instead of 0.3 mm as designed. At the same time, there is no existence of the three slots which should be in the middle part of the line to in accordance with the design requirements. Consequently, this will lead to a larger inductive effect than before and the overall operating frequency tuning range will be lower than before, so it will affected the working frequency to some degree. The testing scene is shown in Fig. 11.

The S -parameter of each operating voltage is tested by network analyzer, and the tested results show that this proposed antenna realizes about 370 MHz tuning range continuously from 1.52 GHz to 1.89 GHz under the voltage from 4 V to 26 V. What's more, the return loss S_{11} can reach -20 dB from 1.7 GHz and the bandwidth at -10 dB is nearly 50 MHz.

For the limitation of space, this paper chooses some original images as example show in Fig. 12 from (a) to (f). And all the csv data of original images in the network analyzer is processed by MTLAB, the tunable results are shown in Fig. 13.

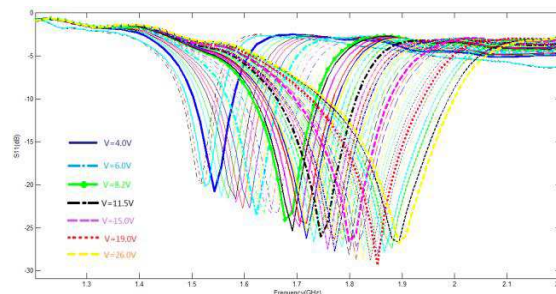


Figure 13: Measured return loss of the proposed antenna by the Vector network.

5. CONCLUSIONS

The frequency-tunable CPW Fed monopole antenna, designed in this paper, has the advantage of small size, wide tunable range, good directivity, easy to process, which can be applied to near field wireless transmission system such as WLAN, Buletooth, UWB, ZigBee and RFID.

ACKNOWLEDGMENT

The author would like to thank friends and teachers both in Peking University and South West Jiao Tong University for their continuous support not only in finance but also in mentality. Especially, the author would thank Professor Quanyuan Feng, who gives his most concern and care for all of his students.

REFERENCES

1. Sheta, A.-F. and S. F. Mahmoud, "A widely tunable compact patch antenna," *IEEE Antennas and Wireless Propagation Letters*, Vol. 7, 40–42, 2008.
2. Komulainen, M., M. Berg, H. Jantunen, E. T. Salonen, and C. Free, "Afrequency tuning method for a planar inverted-F antenna," *IEEE Trans. Antennas Propag.*, Vol. 56, 944–950, Apr. 2008.
3. Ko, C.-H., K. M. J. Ho, and G. M. Rebeiz, "An electronically-scanned 1.8–2.1 GHz base-station antenna using packaged high-reliability RF MEMS phase shifters," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 61, No. 2, 979–985, 2013.

4. Adhikari, S., A. Ghiotto, and K. Wu, “Simultaneous electric and magnetic two-dimensionally tuned parameter-agile SIW devices,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 61, No. 1, 423–435, Jan. 2013.
5. Zhu, S., D. G. Holtby, K. L. Ford, A. Tennant, and R. J. Langley, “Compact low frequency varactor loaded tunable SRR antenna,” *IEEE Transactions on Microwave Theory and Techniques*, 2013.
6. Huang, L. and P. Russer, “Electrically tunable antenna design procedure for mobile applications,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 56, No. 12, 2789–2797, Dec. 2008.
7. Mirzaei, H. and G. V. Eleftheriades, “A compact frequency-reconfigurable metamaterial inspired antenna,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 10, 1154–1157, 2011.
8. Zhu, J. and G. V. Eleftheriades, “A simple approach for reducing mutual coupling in two closely-spaced metamaterial-inspired monopole antennas,” *IEEE Antenna and Wireless Propagation Letters*, Vol. 9, 379–383. 2010.
9. Zhu, J. and G. V. Eleftheriades, “Dual-band metamaterial-inspired small monopole antenna for WiFi applications,” *IET Electronics Letters*, Vol. 45, No. 22, 1104–1106, Oct. 2009.

Design of Linear Biconical Antenna Array for the Generation of Trapezoidal Pattern

A. Sudhakar¹, C. Subba Rao², and A. V. Yashwanth¹

¹Department of ECE, RVR & JC College of Engineering
Chowdavaram, Guntur, Andhra Pradesh 522019, India

²Department of ECE, PVP Siddhartha Institute of Technology
Vijayawada, Andhra Pradesh 520007, India

Abstract— The radiation patterns of individual antennas are either omnidirectional or directional in shape. The patterns of a linear array are either broad side or end fire in shape and the radiation patterns are omnidirectional in collinear arrays. These shapes are common and are widely used in many broad casting applications with a single element. Shape of the radiation pattern can be altered by using reflectors of various geometries. But these antennas with reflectors are heavy and are not suitable for the present day applications where miniaturization is the prime requirement of any gadget. Antenna arrays are highly suitable for effectively controlling the radiation pattern properties such as position of nulls, number of side lobes and their amplitudes using parameters of the array such as input excitation of the elements and the separation between them. Geometrical shape of the array also plays an important role in determining the radiation properties but complex configurations are not preferred due to high cost. Complex shaped beams can be easily generated by using simple linear antenna array with excitation control. Excitation may be input amplitude or phase of the individual elements of the array. Amplitude control method is not usually preferred as the scanning of the beam is not possible and usually losses in the feed configuration are more. More over the amplitude taper between center element and edge elements becomes pronounced as the number of elements in the array increases. Control of excitation phase of individual elements to generate complex beam shapes is more optimum as the input amplitudes of the elements are maintained constant and is suitable for scanning the beam at high speeds. The radiation intensity from a linear array may be expressed as a Fourier Transform which consists of complex exponential term. Many times such transforms cannot be solved by standard methods. In such cases numerical or approximation methods are used to obtain desired result. In this paper a biconical linear antenna array is proposed to generate trapezoidal shaped radiation pattern using input phase control method. The stationary phase concept is used to evaluate required phase coefficients necessary for the elements of the proposed array to generate the Trapezoidal shaped radiation pattern which is highly useful in ground mapping and surveillance applications.

1. INTRODUCTION

Many wireless communication systems in VHF and UHF range require wide area coverage and uses wire antennas because of their low cost. The directivity of single element antenna is very poor and side lobes are more. Point to point communication systems require pencil beams whose beam width is around 15 degrees but design of these antennas are highly involved. Generation of sector and cosecant patterns has been reported by many authors in the past. H. J. Zhou et al. proposed [1] planar array for the generation of flat-topped shaped beam, but the proposed array is very large and expensive. J. A. R. Azevedo has proposed synthesis of shaped patterns with non uniform sample phases [2], but the resultant pattern has more side lobes. An implicit constraint is used on the excitation current of elements of the linear array and the overall pattern function is minimized by placing constraints on the excitation currents [3]. Analysis of Biconical antenna for its wide band characteristics is reported in detail in [4–6]. Linear Biconical antenna characteristics are analyzed and reported in [7]. In this paper the phase control method is used to generate trapezoidal shaped beam.

2. LINEAR BICONICAL ANTENNA ARRAY

Biconical antennas exhibit wide band radiation characteristics. Radiation characteristics of biconical antenna with 90 degrees cone angle is reported in [6]. The electric and magnetic fields are related by

$$\nabla \times E = -j\omega\mu H \quad (1)$$

The electric field is oriented along the axis of the cones and the magnetic field in the horizontal plane.

$$\frac{1}{r} \frac{\partial}{\partial r} (rE_\theta) = -j\omega\mu H_\phi \quad (2)$$

$$\nabla \times H = j\omega\varepsilon E \quad (3)$$

$$\frac{1}{r \sin \theta} \frac{\partial}{\partial r} (r \sin \theta H_\phi) = -j\omega\varepsilon E_\theta \quad (4)$$

$$\frac{1}{r} \frac{\partial}{\partial r} (rH_\varphi) = -j\omega\varepsilon E_\theta \quad (5)$$

$$\frac{\partial^2 (rH_\varphi)}{\partial r^2} = -\omega^2 \mu \varepsilon (rH_\varphi) = -K^2 (rH_\varphi) \quad (6)$$

$$H_\phi = \frac{H_0}{\sin \theta} \frac{e^{-jkr}}{r} \quad (7)$$

$$E_\theta = \eta H_\varphi = \eta \frac{H_0}{\sin \theta} \frac{e^{-jkr}}{r} \quad (8)$$

A linear biconical antenna array consists of required number of biconical antennas are arranged along a line and the array excitation can be used effectively to control the radiation properties of the pattern of the array, such as introducing nulls in the desired direction, to introduce desired side lobe topology or to generate desired shaped radiation pattern.

For the linear array of discrete radiators under consideration, the radiation pattern is given by

$$E(u) = \sum_{n=1}^N b(x) e^{\frac{j2\pi L}{\lambda} [ux + \varphi(x)]} \quad (9)$$

where $x = \frac{2n-1-N}{N}$.

The stationary phase concept method is used for evolving the phase coefficients necessary for generation of Trapezoidal pattern. $\varphi(x)$ is the phasor corresponding to $E(u)$. Here $b(x)$ is the excitation amplitudes of the elements of the array and is maintained constant in phase control method.

$$\frac{\partial}{\partial x} [ux + \varphi(x)] = 0 \quad (10)$$

$$u = -\frac{\partial \varphi}{\partial x} \quad (11)$$

$$\frac{L}{\lambda} \int_{u_1}^{u_2} |E(u)|^2 du = \int_{x_1}^{x_2} |(b(x))|^2 dx \quad (12)$$

$$E(u) = (1 - u); \quad -u_0 \leq u \leq u_0 \quad = 0 \text{ otherwise} \quad (13)$$

3. SYNTHESIS OF LINEAR BICONICAL ANTENNA ARRAY TO GENERATE TRAPEZOIDAL PATTERN

A biconical dipole of height of half wave length and cone angle of 90 degrees is designed, modeled, simulated and its radiation characteristics are analyzed for its wide band characteristics in [6]. A linear array of biconical antennas of equal separation is formed and analyzed in [7].

$$\lambda_g = \lambda/\sqrt{2.1} = 34.5 \text{ mm} \quad (14)$$

$$\lambda_g = 360 = 34.5, \text{ hence for } 1 \text{ degree} = 0.096 \text{ mm} \quad (15)$$

A linear biconical antenna arrays consisting of eight and fifty elements are simulated in WIPL-D. Separation between the elements of the array is half wave length. The cone angle of each cone is 90 degrees and is designed for 6 GHz frequency. Height of each cone is 12.5 mm corresponding to quarter wavelength. Each element is excited using phase coefficients necessary for the generation of Trapezoidal pattern. Also the eight element biconical linear antenna array is fabricated with

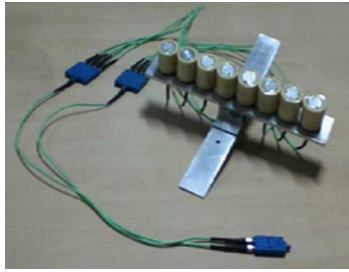


Figure 1: Fabricated 8-element biconical array.



Figure 2: Cables used in the feed for the measurement.

Table 1: Position and phase level of each element in the array.

Element Number	Element Location (mm)	Phase angle (Deg.)	Phase angle (normalized)	Cable length (mm)
1	-87.5	394.07	0	0
2	-62.5	339.54	54.53	5.235
3	-37.5	299.42	94.65	9.1
4	-12.5	269.218	124.85	11.99
5	12.5	246.89	147.18	14.13
6	37.5	231.19	162.88	15.64
7	62.5	221.29	172.78	16.59
8	87.5	216.57	177.5	17.04

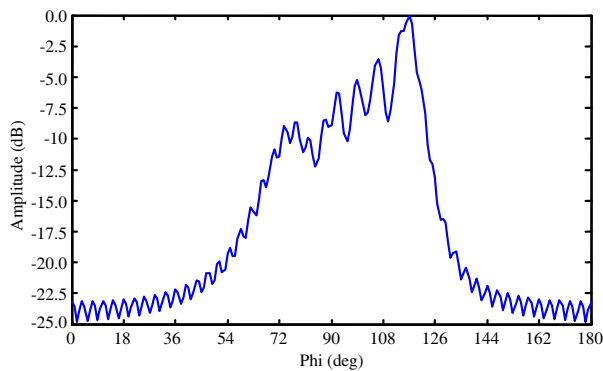


Figure 3: Simulated trapezoidal pattern of 8-element array.

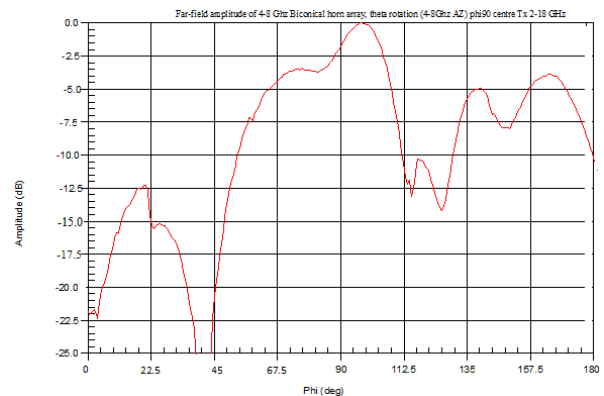


Figure 4: Measured trapezoidal pattern.

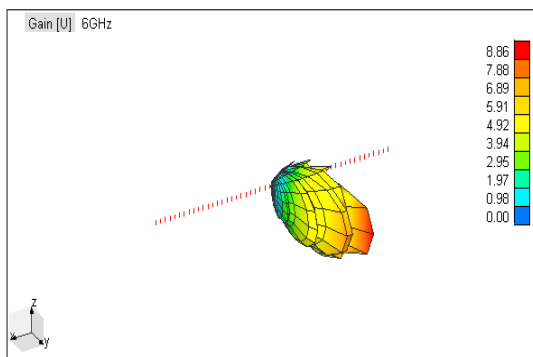


Figure 5: Simulated trapezoidal pattern of 50-element array (3D).

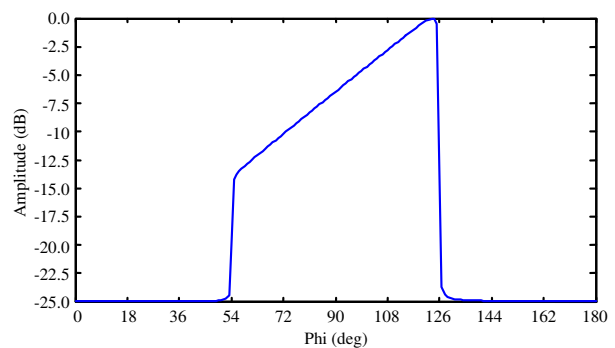


Figure 6: Radiation pattern in horizontal plane.

the same dimensions for operation at 6 GHz frequency as shown in Figure 1. The feed of the array consists of one 1:2 power divider and two 1:4 power dividers. The out puts of the power

dividers are connected to the cables shown in Figure 2, that are designed and cut to the length proportional to the phase coefficients of the corresponding element and every element in the array is fed at a different phase as listed in Table 1, necessary for the generation of the desired Trapezoidal pattern. All the elements in the array are excited at same amplitude. Figure 3 shows the simulated trapezoidal pattern from eight element array and Figure 4 shows the respective measured pattern. Whereas Figure 5 shows the trapezoidal pattern (3D) simulated from biconical array of 50 elements and Figure 6 shows the respective pattern in horizontal plane.

4. RESULTS AND CONCLUSION

Linear biconical antenna arrays consisting of eight and fifty elements are modeled and simulated to generate Trapezoidal Radiation patterns. Simulated pattern with small array consists of more side lobes outside the main beam region and the main beam has ripples and is oscillatory. The transition between shaped and unshaped regions is also not sharp. But the pattern generated from large array consisting of more number of elements has no side lobes outside the main beam and there are no ripples in the main beam region. The transition of the beam is very sharp and shape of the main beam is very close to the desired pattern. An eight element biconical linear array is fabricated and its radiation characteristics are measured by exciting with phase coefficients generated using phase control method, keeping amplitudes of the elements constant. Comparing the simulated and measured patterns it shows that the pattern in the shaped region is in close agreement with that of the simulated pattern in both cases in shape and level. But the number of side lobes and their amplitudes in the measured pattern are not in exact agreement with that of simulated pattern. This is due to losses in the cables used in the feed for the excitation of the array elements. Many techniques have been reported for the generation of shaped beams but did not address the problem of reduction of side lobes and their amplitudes. In the proposed method the results show that side lobes can be completely eliminated by using large arrays consisting of more number of elements. But increasing the number of elements in the array also increases the complexity of the feed and cost. Hence an optimum selection is to be made based on the application. Low precision and low cost applications where approximate pattern with reasonable amount of side lobe levels are suitable can use small arrays. But high precision radar applications where perfect shape with minimum side lobes is required the large arrays are the best option. As the number of elements in the array becomes more the beam in the shaped region approaches the desired pattern as with the fifty element array. As the number of elements in the array increases the side lobes and ripples of the main beam gets minimized and the pattern will be very close to the desired pattern.

Excitation phase coefficients can be generated for arrays consisting of even up to four hundred elements.

ACKNOWLEDGMENT

Authors thank the University Grants Commission, Govt. of India, New Delhi and the Management of R.V.R & J.C College of Engineering, Guntur for their financial support for this work.

REFERENCES

1. Zhou, H.-J., B.-H. Sun, J.-F. Li, and Q.-J. Liu, "Efficient optimization and realization of a shaped-beam planar array for very large array application," *Progress In Electromagnetics Research*, Vol. 89, 1–10, 2009.
2. Azevedo, J. A. R., "Shaped beam pattern synthesis with non-uniform sample phases," *Progress In Electromagnetics Research B*, Vol. 5, 77–90, 2008.
3. Buckley, M. J., "Synthesis of shaped beam antenna patterns using implicitly constrained current elements," *IEEE Transactions on Antennas and Propagation*, Vol. 44, No. 2, 192–197, Feb. 1996.
4. Balanis, C. A., *Antenna Theory: Analysis and Design*, 3rd Edition, John Wiley & Sons, 2008.
5. Stutzman, W. L. and G. A. Thiele, *Antenna Theory and Design*, John Wiley & Sons, 1981.
6. Subba Rao, C. and A. Sudhakar, "Design and modeling of biconical dipole for broad band applications," *International Conference on Advances in Computing, Control, and Telecommunication Technologies (ACT-2009)*, 248–251, Trivandrum, 2009, DOI: 10.1109/ACT.2009.69.
7. Subba Rao, C. and A. Sudhakar, "High gain ultra wide band biconical antenna array," *Association for the Advancement of Modeling & Simulation Techniques in Enterprises (AMSE)*, Vol. 83, No. 1, 68–78, France, Feb. 2010.

Layer Homogenization Method for Modeling Additive Color Effect in Natural Photonic Polycrystals — The Case of *Entimus imperialis* Weevil

Sébastien Mouchet, Jean-Pol Vigneron, Jean-François Colomer, and Olivier Deparis

Department of Physics, University of Namur, 61 rue de Bruxelles, B-5000 Namur, Belgium

Abstract— An original method for computing the reflectance of photonic crystal whose unit cell is composed of form-birefringence anisotropic elements (e.g., cylinders) is proposed. This method, based on the layer homogenization of the photonic crystal, is particularly useful when the reflectance is calculated for different crystal orientations. The far-field color due to an additive color effect in a natural photonic polycrystal found on *Entimus imperialis* weevil was investigated.

1. INTRODUCTION

Modeling the optical properties of photonic polycrystals is crucial for the investigation of natural photonic structures as well as for the design and fabrication of bioinspired devices [1, 2]. The computation of reflectance (or transmittance) of photonic crystals (PCs) by the Rigorous Coupled Wave Analysis (RCWA) method [1, 3] is sometimes cumbersome to perform due to the presence of form-birefringence anisotropic elements in the crystal unit cell (e.g., cylinders or parallelepipeds), particularly when it has to be calculated for various crystal orientations. The Layer Homogenization (LH) method proposed here solves this issue specifically for crystals with periodicities such that there is only specular reflection and no other diffraction orders [4]. For a single crystal orientation, the dielectric function is homogenized within the layers in which the crystal is sliced. A standard thin film solver is then used for the calculation of the reflectance. The disorder in the crystal orientations of the photonic polycrystal is taken into account by averaging reflectance spectra computed for various crystal orientations and incidence angles.

We demonstrate the usefulness of this method by applying it to the study of a natural photonic polycrystal. Polycrystals found in living organisms are known to give rise to multiscale visual effects such as additive colors. A nice example is found in cavities on the cuticle of *Entimus imperialis* [2, 4–8]. This photonic polycrystal is made of one single PC structure with different orientations [5], which is relevant for developing bioinspired applications. For instance, the *E. imperialis* structure has been recently replicated in order to serve as biomimetic devices such as gas, temperature or pH sensors [9].

2. PHOTONIC CRYSTAL WITH FORM ISOTROPIC AND ANISOTROPIC ELEMENTS

One of the most common method used to calculate the reflectance spectrum of 3D PC films is the RCWA method [1, 3] which is based on Fourier expansions of the dielectric function and of the electric and magnetic fields. In this method, a PC film is treated as a stratified medium with lateral periodicities. The layers of the PC film are perpendicular to the crystal orientation \vec{l} and taken homogeneous along that orientation. Elements in the crystal unit cell (defined by translation vectors \vec{a}_x , \vec{a}_y and \vec{a}_z) are therefore discretized into sets of non overlapping islands of different materials, embedded in a host and having simple shapes such as cylinders or parallelepipeds with axes parallel to \vec{l} . For instance, a spherical element is approximated by a set of cylinders (Fig. 1). Calculation of the reflectance according to different crystal orientations is easily performed due to the form isotropy of the sphere: the discretization is the same whatever the crystal orientation is [1].

However, when form anisotropic elements are found in the unit cell, calculation of the reflectance according to various crystal orientations is more complicated to perform (Fig. 2): The discretization is not the same for the different orientations and may involve islands with non simple shapes [4]. In this case, the LH method is needed to solve the problem.

3. LAYER HOMOGENIZATION (LH) METHOD AND POLYCRYSTAL MODELING ASPECTS

Before modeling the additive color effect in the polycrystal, the various crystal orientations were treated separately. In order to compute the reflectance spectrum of a PC with form anisotropic

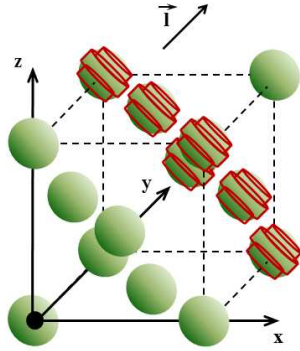


Figure 1: In the framework of RCWA, any element in the unit cell is discretized by simple shapes such as cylinders or parallelepipeds. A sphere, in this case, is approximated by a set of cylinders whose axes are parallel to the crystal orientation \vec{l} under study. Whatever the crystal orientation is, the discretization of a sphere is the same due to its form isotropy.

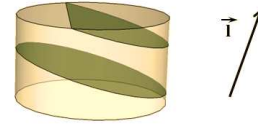


Figure 2: A cylinder cannot easily be discretized by a set of simple shapes when studied under any crystal orientation \vec{l} that is neither perpendicular nor parallel to the axis of the cylinder. The computation of reflectance or transmittance is therefore more complicated to perform.

elements in the crystal unit cell whatever the orientation is, we developed the LH method [4] in which the crystal was sliced into a set of layers for each orientation under study. The dielectric function was homogenized within each layer by a simple spatial averaging (Fig. 3). Thanks to this method, the resolution of Maxwell's equations is highly simplified and the reflectance of a single crystal orientation can be computed using a standard thin film solver. However, the averaging of the dielectric function can be performed only if the refractive index contrast within each layer is weak. Such a weak contrast is found in photonic structures occurring in insects since they are made of chitin ($n_{chitin} = 1.56$ [10]) and air ($n_{air} = 1$).

Since the PC is approximated by a stratified medium with homogenized layers (Fig. 3(b)), a translational invariance appears along any direction parallel to the layer interface. That implies that only specular reflection occurs in that layer-homogenized medium. In order to apply our method, the lateral period of the PC has to be small in comparison with the incident wavelength. This feature is rather common in natural PCs, such as *Entimus imperialis* investigated hereafter.

The reflectance spectrum due to an additive color effect can be modeled by averaging reflectance spectra computed for various incidence angles and crystal orientations. This procedure is often used for computing optical properties of multiscale structures [11, 12].

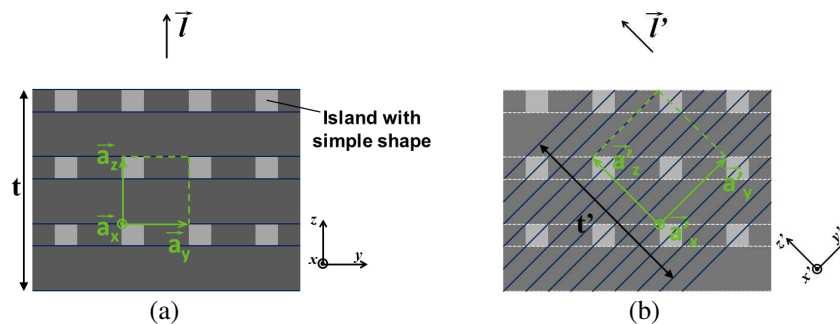


Figure 3: Layer Homogenization (LH) method. (a) A photonic crystal is first defined for a given crystal orientation \vec{l} as a stratified medium with 2D lattices of islands having simple shapes (cylinders or parallelepipeds). The unit cell is described by \vec{a}_x , \vec{a}_y and \vec{a}_z vectors. \vec{a}_z is parallel to the crystal orientation \vec{l} . (b) The photonic crystal is redescribed by a laterally homogeneous stratified medium with an orientation \vec{l}' that is not parallel to \vec{a}_z . A new crystal unit cell is defined by other translation vectors \vec{a}'_x , \vec{a}'_y and \vec{a}'_z (\vec{a}'_z being parallel to \vec{l}').

4. ADDITIVE COLOR EFFECTS IN *ENTIMUS IMPERIALIS*

The polycrystal found on the cuticle of *E. imperialis* is a noticeable example of additive photonic color [2, 5] (Fig. 4(a)). A photonic polycrystal with different crystal domains is found in the scales which cover the cuticle. In each domain, the same photonic structure is found but with a different crystal orientation. Any PC domain with a given orientation gives rise to a single bright iridescent color (from blue to orange) (Fig. 4(b)). The unit cell of the PC can be modeled by perforated chitin layers on which a lattice of protrusions is found (Figs. 4(c), (d)) [2, 5]. Analysis of this morphology by a method presented in [12] showed that this structure corresponded to a FCC crystal with a lattice parameter equal to 396 nm [5]. We demonstrated in previous studies that this structure solely reflected the incident light in the specular direction [4, 5]. In spite of the very directional and bright color produced by each photonic domain, the color perceived in the far-field by human eye is a non-iridescent matt green color (Fig. 4(a)). As we will see hereafter, this matt color is due to the disorder in the crystal orientations.

The LH method was used to calculate the reflectance according to the (111) and (100) crystal orientations at different incidence angles (Fig. 5). For the (111) orientation at normal incidence, the specular reflectance peak was found at 577 nm (i.e., yellow) and an harmonic was found at 297 nm. For the sake of comparison, the reflectance spectra in this crystal orientation were simulated by RCWA as well. We note that for this orientation, the elements composing the crystal unit cell could be easily described by islands with simple shapes. The RCWA spectra are in good agreement with those computed by the LH method. In the case of the (100) orientation, the main reflectance peak was located at 530 nm (i.e., green). For that orientation, the computation could be performed only by the LH method because the unit cell is made of form anisotropic elements. When the incidence angle θ is increased, the spectrum is blue-shifted for both crystal orientations. For instance, a violet color is obtained at an incidence angle of 45° for both orientations.

The diversity of colors due to the polycrystalline structure is better highlighted on a chromaticity diagram. Chromaticity coordinates were computed for the two investigated crystal orientations, at incidence angles ranging from 0° to 75° by step of 15° . They were found to be distributed across the diagram (Fig. 6(a)): Different domains gave rise to yellow, green and blue colorations, as we observed in single scales (Fig. 4(b)).

In order to take into account the disorder in the crystal orientations and to simulate its far-field coloration, the reflectance spectra calculated at various incidence angles and for various crystal orientations were averaged. The chromaticity coordinates calculated from the average of the spectra were found in the faint green central area of the diagram (Fig. 6(a)). They represent the color perceived by the naked eye, which appeared here to be a faint green color. A previous study [5] provided us with the reflectance spectra measured for different crystal domains by microspectropho-

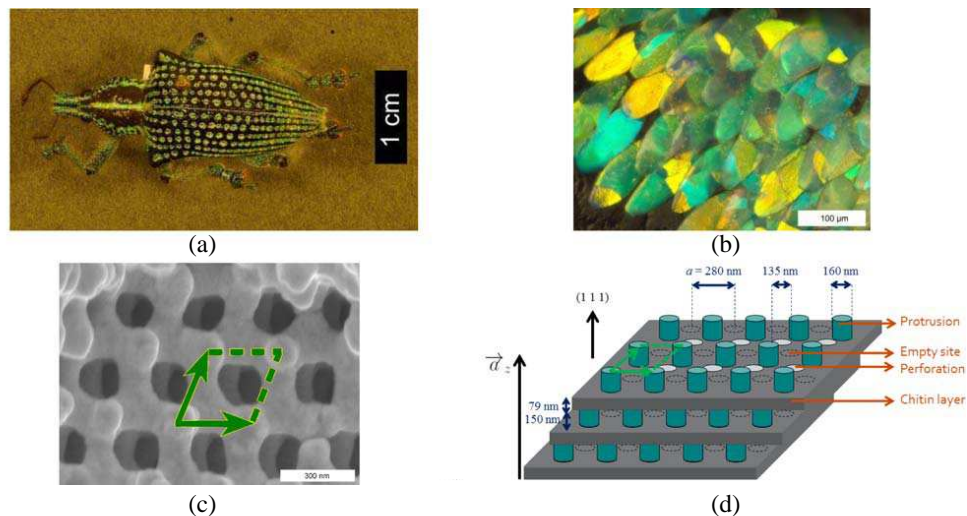


Figure 4: (a) *Entimus imperialis* is a remarkable example of additive color effect. (b) A photonic polycrystal made of different colorful domains covers its cuticle. (c) The photonic structure found in each domain was evidenced by SEM. (d) This structure was modeled by 2D lattices of protrusions covering chitin layers with cylindrical perforations. This model had a FCC symmetry.

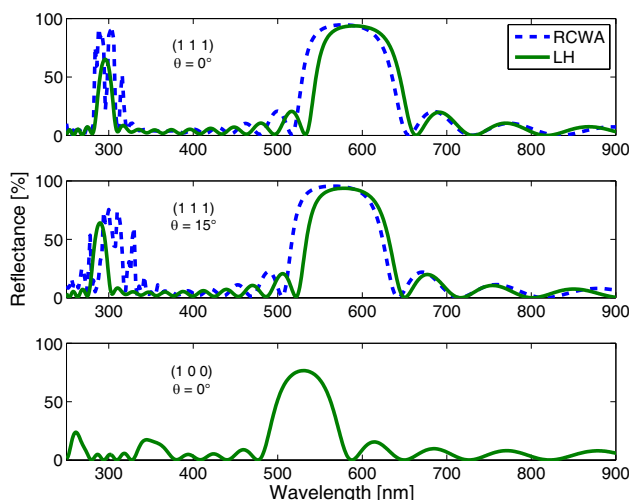


Figure 5: Reflectance spectra in the (111) and (100) orientations calculated using two computational methods at different incidence angles θ . Curves labelled RCWA were computed by RCWA method and curves labelled LH were simulated by the Layer Homogenization method. The reflectance spectrum in the (100) orientation could only be calculated by the LH method because of the form anisotropy of the elements in the unit cell (i.e., cylinders).

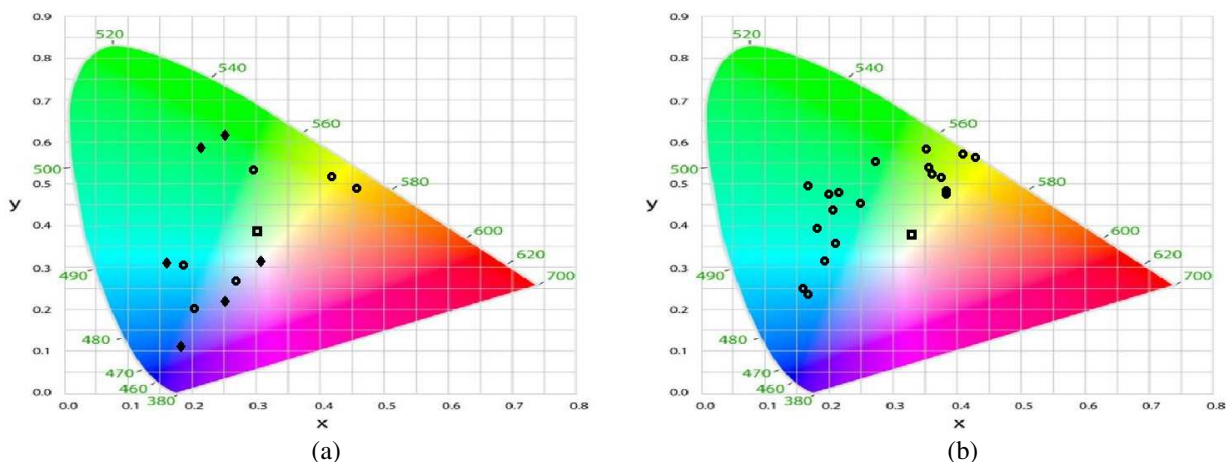


Figure 6: (a) Chromaticity coordinates computed at different incidence angles θ (from 0° to 75° by step of 15°) in the cases of the (111) orientation (\circ) and (100) orientation (\blacklozenge). There is a great diversity of colorations produced by the crystal domains. The average of the reflectance spectra (\square) leads to coordinates in the faint green central area. (b) The chromaticity coordinates calculated from microspectrophotometry measurements (\circ) [5] are distributed across the whole diagram. The coordinates from far-field reflectance measurements (\square) [5] also fall in the faint green area.

tometry and in the far-field by spectrophotometry (Fig. 6(b)). The diversity of colors produced by the crystal domains was reproduced numerically by the spectra simulated for different incidence angles and two specific crystal orientations. The faint green color measured in the far-field [5] is well reproduced by the average of the simulated spectra due to an additive color effect.

5. CONCLUSION

The additive colors effect occurring in polycrystals made of one single photonic structure was investigated. The calculation of reflectance (or transmittance) spectra according to various crystal orientations often turns out to be problematic when form anisotropic elements are present in the unit cell. The method we developed relied on layer homogenization and solved this problem by computing the reflectance thanks to a standard thin film computation code. The method assumed that the photonic crystal reflected the incident light only in the specular direction (at least in the

visible range). The additive color effect of the photonic polycrystal found on the cuticle of *Entimus imperialis* was predicted by averaging reflectance spectra computed for different incidence angles and crystal orientations.

ACKNOWLEDGMENT

The authors thank Cédric Vandenberg for fruitful and helpful discussions. S.M. and J.-F.C. were supported by the Belgian National Fund for Scientific Research (F.R.S.-FNRS) as Research Fellow and Research Associate, respectively. The authors acknowledge the use of resources of the Electron Microscopy Facilities of the University of Namur, Belgium and the Interuniversity Scientific Computing Facility (iSCF) located at the University of Namur, which is supported by the F.R.S.-FNRS under convention No. 2.4617.07.

REFERENCES

1. Vigneron, J.-P. and V. Lousse, "Variation of a photonic crystal color with the Miller indices of the exposed surface," *Proceedings of SPIE*, Vol. 6128, 61281G, San Jose, CA, United States, January 2006.
2. Deparis, O. and J.-P. Vigneron, "Modeling the photonic response of biological nanostructures using the concept of stratified medium: The case of a natural three-dimensional photonic crystal," *Mater. Sci. Eng. B — Adv.*, Vol. 169, No. 1–3, 12–15, 2010.
3. Moharam, M. G. and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *J. Opt. Soc. Am.*, Vol. 71, No. 7, 811–818, 1981.
4. Mouchet, S., J.-F. Colomer, C. Vandenberg, O. Deparis, and J.-P. Vigneron, "Method for modeling additive color effect in photonic polycrystals with form anisotropic elements: The case of *Entimus imperialis* weevil," *Opt. Express*, Vol. 21, No. 11, 13228–13240, 2013.
5. Mouchet, S., J.-P. Vigneron, J.-F. Colomer, C. Vandenberg, and O. Deparis, "Additive photonic colors in the Brazilian diamond weevil, *Entimus imperialis*," *Proceedings of SPIE*, Vol. 8480, 848003, San Diego, CA, United States, August 2012.
6. Wilts, B. D., K. Michielsen, J. Kuipers, H. De Raedt, and D. G. Stavenga, "Brilliant camouflage: photonic crystals in the diamond weevil, *Entimus imperialis*," *P. Roy. Soc. B — Biol. Sci.*, Vol. 279, No. 1738, 2524–2530, 2012.
7. Wilts, B. D., K. Michielsen, H. De Raedt, and D. G. Stavenga, "Hemispherical Brillouin zone imaging of a diamond-type biological photonic crystal," *J. Roy. Soc. Interface*, Vol. 9, No. 72, 1609–1614, 2012.
8. Wu, X., A. Erbe, D. Raabe, and H.-O. Fabritius, "Extreme optical properties tuned through phase substitution in a structurally optimized biological photonic polycrystal," *Adv. Funct. Mater.*, 1–6, 2013.
9. Van Opdenbosch, D., M. Johannes, X. Wu, H.-O. Fabritius, and C. Zollfrank, "Fabrication of three-dimensional photonic crystals with tunable photonic properties by biotemplating," *Photonics Nanostruct.*, Vol. 10, No. 4, 516–522, 2012.
10. Sollas, I. B. J., "On the identification of chitin by its physical constants," *P. Roy. Soc. B — Biol. Sci.*, Vol. 79, No. 354, 474–481, 1907.
11. Rassart, M., J.-F. Colomer, T. Tabarrant, and J.-P. Vigneron, "Diffractive hygrochromic effect in the cuticle of the hercules beetle *Dynastes hercules*," *New J. Phys.*, Vol. 10, No. 033014, 1–14, 2008.
12. Welch, V. L., V. Lousse, O. Deparis, A. R. Parker, and J.-P. Vigneron, "Orange reflection from a three-dimensional photonic crystal in the scales of the weevil *Pachyrrhynchus congestus pavonius* (Curculionidae)," *Phys. Rev. E*, Vol. 75, No. 4, 041919-1–9, 2007.

Numerical Analysis of Grating and Prism Coupled Surface Plasmon Resonance

Sorin Miclos¹, Dan Savastru¹, Ion Lancranjan¹, Roxana Savastru¹, and Constantin Opran²

¹National Institute of R&D for Optoelectronics — INOE 2000
409 Atomistilor St., Magurele, Ilfov RO-077125, Romania

²Politehnica University of Bucharest, 313 Splaiul Independentei Ave., Bucharest, Romania

Abstract— The main purpose of the paper consists of presenting simulation results obtained in analyzing the surface plasmons excitation effect, presently commonly referred as surface plasmon resonance (SPR) and its main possible application as a highly accurate and sensitive sensor. Both electromagnetic wave coupling configurations to the medium, grating and Kretschmann prism configurations are investigated. In the grating coupling configuration, the performed numerical simulations are pointing mainly to nanowire and/or nanostructured gratings. The numerically simulation model developed in the case of the grating coupling configuration consider Cu, Au, Ag or other metal nanowire having a diameter of 40 nm to 800 nm, formed on dielectric substrates with a refractive index between 1.4 and 2.4, such as SF11, BK7 glasses, GaP and TiO₂-o (rutile). The same range of dielectric materials is considered for the prism coupling configuration. The thin layers formed on the face of prism coupler is considered as being formed by up to 15 Cu, Au, Ag or other metal layers. Several shapes of the groves formed onto the high refractive index substrate are investigated, such as square/rectangular (including the rounded corner case), circular or elliptical. As the low-index background material air, water and sea-water were considered. In order to get more insight into the plane electromagnetic wave propagation in such waveguides a simulation procedure is developed on the basis of the Finite Element and Finite Difference Time Domain methods and a finite element Helmholtz solver. For practical applications, it is of interest to evaluate how much electromagnetic radiation can be coupled into the medium. Therefore, the coupling coefficient of electromagnetic radiation from and into the medium using the grating and Kretschmann prism coupling configurations has been examined.

1. INTRODUCTION

Since the first evidence report of surface plasmons excitation effect, presently commonly referred as surface plasmon resonance (SPR), its main possible application is as a highly accurate and sensitive sensor [1–5]. Optical sensors based on SPR belong to the group of refractometric sensing devices including the resonant mirror sensor [1, 2], the grating coupler sensor [3–5], the integrated optical Mach-Zehnder interferometer [6, 7], the integrated Young interferometer [8, 9] and the white light interferometer [10, 11], which measure changes in the refractive index occurring in the field of an electromagnetic wave supported by the optical structure of the sensor [12–14]. Related to these possible sensor applications, numerical simulation, modeling of surface plasmon resonance (SPR) is of increased interest [8, 10, 15]. This paper presents numerical simulation results obtained in investigating SPR sensors based on grating couplers, pointing mainly to nanowire and/or nanostructured grating.

When performing an analysis of surface plasmon resonance effect pointing to its application as high sensitivity sensors, several items have to be clearly understood:

- Basically, plasmon resonances in metallic nanostructures are collective oscillations of the conduction band electrons. Plasmon resonances are excited when radiation of certain wavelengths is incident on these nanostructures. Among the metallic nanostructures can be enumerated thin films, nanopillars fabricated on planar surfaces, and nanoparticles.
- The main effect which appears when excitation of surface plasmons is performed consists into an increase in the electromagnetic field intensity in the vicinity of metallic thin films and nanostructures. This increase of electromagnetic field intensity leads to surface plasmon resonance [1–4], localized surface plasmon resonance [5–8] and surface-enhanced Raman scattering.
- Plasmon resonances of noble metal nanostructures (mainly gold and silver) are conventionally used for developing sensors based on the fact that these nanostructures resonantly scatter or absorb light in the visible and near-infrared spectra.

- Surface plasmon resonance sensors are conventionally based on detecting changes of refractive indices — both in the bulk media around the metallic films or in the vicinity of the continuous metallic films having nano-scale thickness, conventionally between 20 and 100 nm — using both angular and wavelength interrogation methods.
- Surface plasmon resonance sensor can also rely on detection of differential reflectance — before and after the localized changes of refractive indices on the surface of the plasmonic films.
- A basic problem to be solved for a proper development of surface plasmon resonance sensors consists into the technique to be used for coupling the incident radiation into surface plasmons.

These techniques employ either prism coupling (Kretschmann or Otto configurations), grating-based coupling [1, 2] or waveguide coupling (which includes optical fiber waveguides [1, 2]).

Essentially, using a Kretschmann prism configuration (as in Fig. 1), the electromagnetic radiation is coupled onto a grating formed by depositing metal nanowires or digging nano wells on a plane metallic substrate. The surface plasmon resonance sensors operate by detecting the changes of the oscillation of the surface plasmons, changes induced by modifying the surface mechanical and/or chemical characteristics [1–7]. The net effect is a modulation of surface characteristics, modulation which could be monitored by the angular distributions of reflection and transmission, wavelength, intensity, phase, and polarization changes. Four grating structures were investigated: (a) metal nanowire grating formed on dielectric substrate, (b) cylindrical groove nanostructured grating formed on a dielectric substrate, (c) rectangular groove nanostructured grating and (d) wedge groove nanostructured grating.

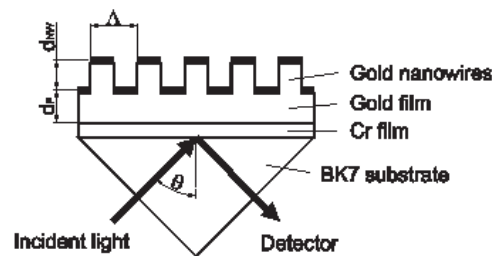


Figure 1: Schematic of a surface plasmon resonance structure.

Such a structure is investigated in order to detect sharp variations of grating reflection index with incidence angle. The model is developed using FEM for calculating the electromagnetic field distribution [7, 8, 12, 16]. It computes the transmission and reflection coefficients for the refraction, specular reflection, and first order diffraction. The TE wave has the electric field component towards z direction, outgoing from the xy -plane. For the TM wave, the incident electric field is specified in the xy -plane and perpendicular to the direction of propagation. Because the periodicity boundaries are parallel to the y -axis, only the x -component is required. During the analysis, the geometric characteristic dimension of the grating structure, i.e., of the metallic wire, of the cylindrical groove, of the rectangular groove or of the wedge groove is in nano domain. The investigated metallic wire has a diameter of 40 to 400 nm. The Floquet “cell”, defined by Floquet boundary conditions has a 200-800 nm length.

2. RESULTS

The simulation results are similar to a large extent for the four investigated grating structures. As example, we present results obtained in the nanowire grating case.

In Fig. 2 is given as example the simulation of EM wave propagation through a Floquet cell with $W = 400$ nm and $H = 3200$ nm. The silver nanowire has a diameter of 80 nm, the propagation medium is sea water, dielectric substrate is GaP and the studied wavelength is 1064 nm.

In Fig. 3, the simulated variations of grating reflectivity versus EM wave angle of incidence for several models are presented. There are 9 models: three wavelengths (633, 1064 and 1550 nm) and three substrates (BK7, SF11 optical glasses and GaP), the grid pitch (400 nm) and nanowire diameter (80 nm) remaining unchanged. Each model has a name including wavelength, propagating medium, grating pitch, wire diameter and substrate. The example in Fig. 3 gathers all models for sea water as propagation medium.

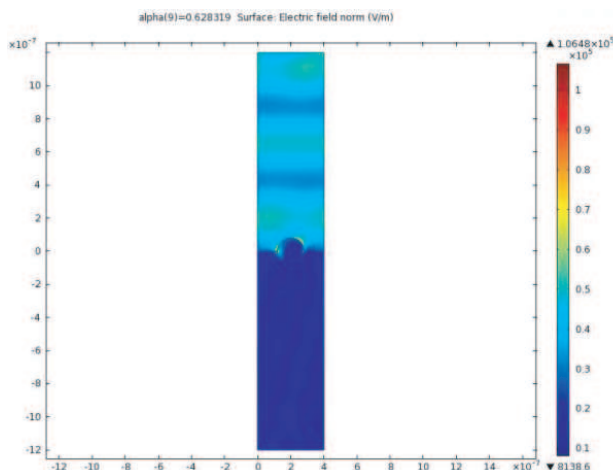


Figure 2: Electric Field Intensity distribution simulated for 1064SEA400D80GaP model: propagation medium — sea water, dielectric substrate — GaP, wavelength — 1064 nm.

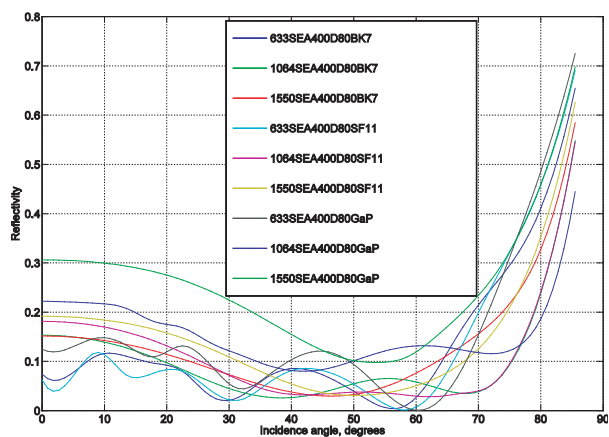


Figure 3: Reflectivity vs incidence angle. Wave propagation medium: sea water.

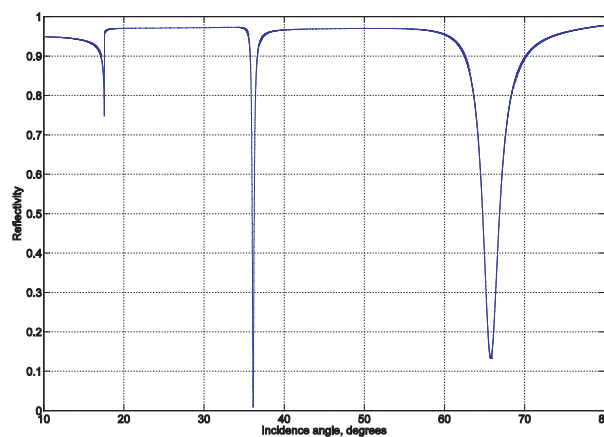


Figure 4: Reflectivity vs incidence angle. Wave propagation medium: air. Kretschmann configuration.

Regarding the presented simulation curves several observations can be made, namely that the reflectivity variation versus EM incidence angle are smooth continuous curves with a minimum located in the EM incidence angle range 45° – 60° , with no “breaking points” indicating possible nonlinear effects; more simulation curves have extremum points indicating null first derivative and sharper variations can be observed; the curve 633SEA400D80GaP which corresponds to a grating structure consisting in the use of a GaP substrate and 633 nm (more exactly 632.8 nm—the He-Ne laser emission wavelength) has more points of extremum in comparison with its “air” counterpart. This curve has sharper variations than in the case of air EM propagation medium. Such a variation of nanowire grating reflectivity could be exploited for a possible sensor application. Such a variation of nanowire grating reflectivity could be exploited for a possible sensor application.

In Fig. 4 is shown the reflectivity curve for a Kretschmann setup, the prism is made of GaP, it is a 50 nm silver layer and a 250 nm GLS layer. The propagation medium is air. Plasmonic resonance occurs at about 66° , while a narrower and better resonance occurs at about 36° [17].

3. CONCLUSIONS

The results obtained in numerical simulation of a nano scale grating system, structure are presented pointing to possible application as a sensor. The performed simulations are used for investigating the variations of the nano scale grating system parameters measurable on macroscopic scale, such as reflectivity functional dependencies on various grating system parameters, namely the grating constant, the wire diameter, refractive index of the EM wave propagation medium, refractive index of the dielectric substrate.

Both grating and prism coupling types prove to be useful in building sensors. The resonance in the prism case appears to be better, but this layout comprises two thin layers (silver and GLS) while the grating layout comprises just one layer. A next step we plan to do is to extend the grating layout to more layers, thus improving its sensitivity.

ACKNOWLEDGMENT

These results were obtained in the frame of Romanian National Authority for Scientific Research "PARTNERSHIP IN PRIORITY DOMAINS" Programme Contract nr. 25/2012 "MEMOPLAS".

REFERENCES

1. Kretschmann, E., "Determination of optical constants of metals by excitation of surface plasmons," *Z. Phys.*, Vol. 241, 313–334, 1971.
2. Homola, J., S. S. Yee, and G. Gauglitz, "Surface plasmon resonance sensors: Review," *Sens. Actuators B Chem.*, Vol. 54, No. 1–2, 3–15, 1999.
3. Jorgenson, R. C. and S. S. Yee, "A fiber optic chemical sensor based on surface plasmon resonance," *Sens. Actuators B Chem.*, Vol. 12, No. 3, 213–220, 1993.
4. Homola, J., *Surface Plasmon Resonance Based Sensors*, Springer, Berlin, 2006.
5. Ctyroky, J., J. Homola, and M. Skalský, "Tuning of spectral operation range of a waveguide surface plasmon resonance sensor," *Electron. Lett.*, Vol. 33, No. 14, 1246–1248, 1997.
6. Slavík, R., J. Homola, J. Čyroký, and E. Brynda, "Novel spectral fiber optic sensor based on surface plasmon resonance," *Sens. Actuators B Chem.*, Vol. 74, No. 1–3, 106–111, 2001.
7. Schuck, P., "Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules," *Annu. Rev. Biophys. Biomol. Struct.*, Vol. 26, No. 1, 541–566, 1997.
8. Malmqvist, M., "Surface plasmon resonance for detection and measurement of antibody-antigen affinity and kinetics," *Curr. Opin. Immunol.*, Vol. 5, No. 2, 282–286, 1993.
9. Bohren, C. F. and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley, New York, 1983.
10. Haes, A. J. and R. P. Van Duyne, "A nanoscale optical biosensor: Sensitivity and selectivity of an approach based on the localized surface plasmon resonance spectroscopy of triangular silver nanoparticles," *J. Am. Chem. Soc.*, Vol. 124, No. 35, 10596–10604, 2002.
11. Mock, J. J., D. R. Smith, and S. Schultz, "Local refractive index dependence of plasmon resonance spectra from individual nanoparticles," *Nano Lett.*, Vol. 3, No. 4, 485–491, 2003.
12. Otto, A., I. Mrozek, H. Grabhorn, and W. Akemann, "Surface-enhanced Raman scattering," *J. Phys. Condens. Matter*, Vol. 4, No. 5, 1143–1212, 1992.
13. Kneipp, K., M. Moskovits, and H. Kneipp, *Surface-enhanced Raman Scattering: Physics and Applications*, Springer, Berlin, 2006.
14. Vo-Dinh, T., "Surface-enhanced Raman spectroscopy using metallic nanostructures," *Trends in Anal. Chem.*, Vol. 17, 557–582, 1998.
15. Chang, R. K. and T. E. Furtak, *Surface-enhanced Raman Scattering*, Plenum, New York, 1982.
16. Futamata, M., Y. Maruyama, and M. Ishikawa, "Local electric field and scattering cross section of Ag nanoparticles under surface plasmon resonance by finite difference time domain method," *J. Phys. Chem. B*, Vol. 107, No. 31, 7607–7617, 2003.
17. Popescu, A. A., R. Savastru, D. Savastru, and S. Miclos, "Application of vitreous As-S-Se chalcogenides as active layer in surface plasmon resonance configuration," *Dig. J. Nanomat. Biostr.*, Vol. 6, No. 3, 1245–1252, 2011.

Determination of the Optimal Value of the Radius of a Circular Cylindrical Post in a Rectangular Waveguide for Measurement of the Dielectric Permittivity

Roman Kushnin and Janis Semenjako

Riga Technical University, Riga, Latvia

Abstract— The goal of this paper is to determine the optimal value of the radius of a circular cylindrical post in a rectangular waveguide, that is, such a value of the radius of the post at which measurement uncertainty takes minimum value. To that end, we employ the well-known Monte Carlo method. Probability distributions for input quantities of the Monte Carlo model are assigned according to the principle of maximum entropy.

1. INTRODUCTION

This paper is inspired by a recently published paper [1], in which a method for determination of the optimal radius of a dielectric post in a cylindrical metallic resonator for measurement of the dielectric permittivity has been proposed. This task requires a lot of computation time, since its based upon the Monte Carlo method that, in turn, requires a very large number of trials in order for estimation of the measurement uncertainty to be reliable and each of these trials requires solution of the inverse scattering problem, that in our case is being solved by iteratively minimizing the distance between simulated and measurement data. In order to accomplish this task in a reasonable amount of time, we need a very fast approach for solving the direct scattering problem, since the computation time required to solve the inverse scattering problem is proportional to that required by an algorithm for solving the corresponding direct scattering problem. A simpler way to evaluate measurement uncertainty is to follow GUM framework. Unfortunately, validity of the GUM framework is restricted only to models that lend themselves to an adequate linear approximation.

The problem of scattering of the dominant mode in a rectangular waveguide by a cylindrical post was first treated in Notes on lectures by Julian Schwinger [2] by using the variational method. Numerical results obtained by application of this method also may be found in [3]. Unfortunately, it has been found that results obtained by the variational method are highly inaccurate for posts of large electric radius, particularly near resonance frequencies. An attempt to improve the accuracy of the variational method was made in [4] by using the so-called second order approximation, that is, by using one more term in both series expansions instead of using just one term. Also many other approaches have been proposed by various authors over the last several decades [5–9].

A first approach based on division of geometry of a problem into subregions of simple regular geometry has been proposed by Nielsen [10]. Since the fields in regular regions are expressed in terms of a series of solutions of the homogenous Helmholtz equation, it allows one to solve the boundary problem on the interfaces between layers of the post analytically that, in turn, considerably reduce computational efforts as well as provides sufficiently high resolution of the field distribution. In order to find unknown expansion coefficients, infinite series are truncated and then by applying the point matching procedure the system of linear equations is set up and solved for unknown expansion coefficients. Unfortunately, this approach converges to the correct result only for post of small electric radius (the product of the radius of the post and wavenumber). It was found that this limitation can be overcome by replacing the rectangular interaction region with the circular one, where the center of the circular interaction region (the region where point-matching procedure applies) coincides with the axis of the post and its radius is equal to one-half the width of the broader wall of the waveguide [11]. It was also found that applying numerical integration on the surface of the interaction region instead of the point matching procedure yields faster convergence [12].

2. DISCUSSION

Analytical solution of the inverse scattering problems is possible only for very simple problem geometries. Even when the direct scattering problem may be solved analytically, it may not be possible to solve the corresponding inverse problem analytically without any approximations. Due to this fact, some numerical procedure has to be used. One of the most common and general ways

to solve this problem numerically is to convert the inverse scattering problem into an equivalent numerical optimization problem, where an objective function is typically chosen as the distance between calculated and measured values of S parameters.

$$Q(f) = \sqrt{\sum_{m=1}^2 \sum_{n=1}^2 (\log S_{nm}^e - \log S_{nm}^s)^2} \quad (1)$$

where $Q(f)$ — is the objective function; S_{nm}^e — measured values of scattering matrix entries; S_{nm}^s — values of the scattering parameters obtained by solving the corresponding direct scattering problem. It is obvious that the objective function takes minimum value when values of the coordinates correspond to the solution of the inverse scattering problem. There are many algorithms for minimization of objective functions, but in this study we use a simple pattern search method [13].

In case when measurements are made only at one particular frequency the solution of the problem is not unique, since the coefficients may take the same value at different values of the complex permittivity. This issue may be overcome by making measurements at two different frequencies. Nevertheless, this multi-frequency method cannot be applied in a case of highly dispersive materials, where constitutive parameters changes very rapidly with frequency. In other words, without at least approximate knowledge of this dependence, it is most likely that error in the results will be large. In such cases, we have to use other measurement methods that not only ensures uniqueness of the solution, but also allows one to make all measurements at some fixed frequency. One such method is to make measurements for two samples with different values of some dimension at a fixed frequency value. Another commonly used method is to use a movable short circuit at one of ports of the measurement cell. Unfortunately, in this case it is possible to measure only the reflection coefficient, the absolute value of which in a case of lossless sample is always equal to unity that means that in a lossless case the only quantity we can measure is the phase of the reflection coefficient.

3. THE DIRECT SCATTERING PROBLEM

Throughout this study it is assumed that the walls of the rectangular waveguide are perfectly conducting. Also we assume that only dominant mode TE_{10} may propagate in the waveguide and all other modes are evanescent and exhibit very rapid exponential decay with distance from the sample. Both these assumptions lead to simplification of the numerical analysis. In order to solve the problem we divide the waveguide into three separate regions as depicted in Figure 1. In region I and III scattered fields are represented in terms of waveguide modes.

$$E_y^I = \sum_{m=1}^{\infty} A_m \cos \frac{m\pi x}{a} e^{-jk_m z} \quad (2)$$

$$E_y^{III} = \sum_{m=1}^{\infty} B_m \cos \frac{m\pi x}{a} e^{jk_m z} \quad (3)$$

In region II fields and inside the cylindrical sample fields are represented in terms of cylindrical waves.

$$E_y^{II} = \sum_{n=0}^{\infty} (C_n J_n(k_o r) + D_n Y_n(k_o r)) \cos n\Theta, \quad (4)$$

$$E_y^p = \sum_{n=0}^{\infty} E_n J_n(k_p r) \cos n\Theta, \quad (5)$$

where k_m — is the waveguide wavenumber, k_o — wavenumber in the air, k_p — wavenumber in the material of the sample. Expressions for corresponding magnetic fields in these regions may be obtained by using the second Maxwell's equation.

Enforcing boundary conditions on the surface of the post as well as taking the advantage of the mutual orthogonality of cylindrical waves with respect to azimuthal coordinate Θ and eliminating the unknown constant E_n , the expression relating unknown constants D_n and C_n may be obtained as follows.

$$D_n = \alpha_n \cdot C_n \quad (6)$$

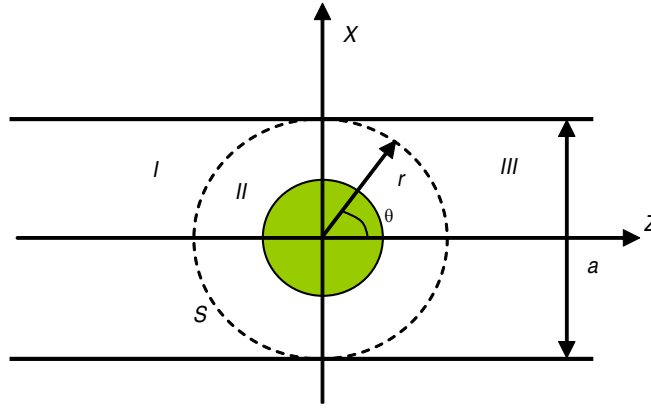


Figure 1: The rectangular waveguide containing the cylindrical dielectric sample.

where

$$\alpha_n = \frac{k_p J'_n(k_p r_p) J_n(k_o r_o) - k_o J_n(k_p r_p) J'_n(k_o r_p)}{k_o J_n(k_p r_p) Y'_n(k_o r_p) - k_p J'_n(k_p r_p) Y_n(k_o r_p)}$$

By enforcing boundary conditions on the surface of the interaction region S we obtain the set of four equations. As it can be seen, this set contains infinite series of cylindrical waves and waveguide modes. However, it is impossible to solve this system by any analytical method, since cylindrical waves and waveguide modes are not mutually orthogonal. It means that we are forced to resort to the truncation of infinite series of waveguide modes and cylindrical terms, thereby converting it into the system of linear algebraic equations. Application of the truncation procedure yields finite series of waveguide modes and cylindrical waves with only M and N first terms retained, respectively. Projecting the difference between field representations on both sides of the interaction region S upon the set of test functions $\cos(p \cdot \Theta)$ (where $p = 1, 2, \dots, M$) and carrying out a number of simple algebraic manipulations the final system of linear algebraic equations for two sets of unknown coefficients A_m and B_m is obtained.

4. NUMERICAL RESULTS

Among the factors influencing the measurement accuracy are the limited resolution, residual systematic error, connection mismatch, and geometrical imperfections of the sample, such as a small shift in the position of the sample and the accuracy of the measurement of the radius of the cylindrical sample. Since the influence of different sources of uncertainty on the uncertainty of the output quantity may be dependent on values of parameters of the model it may be possible that for some optimal combinations of values of these parameters the standard uncertainty $u(\varepsilon)$ of the output quantity will be smaller than for all other combinations. In this paper we consider only the dependence of the standard uncertainty $u(\varepsilon)$ of the output quantity upon only one of the model parameters, namely, the radius of the cylindrical sample. In order to find optimal values of the radius of the sample we estimate the standard uncertainty in measurement of the dielectric permittivity by using the Monte Carlo method. Probability distributions for input quantities of the model are assigned according to the maximum entropy principle. In our case we take into account uncertainties due to the frequency accuracy, imperfect measurement of the radius of the sample and scattering data and assume that all these quantities are distributed according to the normal distribution. Normally distributed random numbers are generated by using uniformly distributed random numbers generated by the pseudo-random number generator and applying Box-Muller transform. Values of the dielectric constant can be extracted from scattering data by means of a pattern search optimization algorithm [11], using the objective function (1).

Application of the Monte Carlo method for each value from some discrete set of values of the radius of the post allows one to find optimal (minimal) values of the standard uncertainty of the dielectric constant. In the present case, for all numerical simulations we have used the number of trials equal to 100000 and values of the first and second measurement frequencies equal to 2.4 GHz and 2.6 GHz, respectively. The results of Monte Carlo simulations for three different types of extraction of the dielectric constant from measured values of reflection data of the field reflected by the cylindrical dielectric post with $\varepsilon = 5$ inserted in the rectangular waveguide with width of the broader wall a equal to 100 mm are depicted in Figures 2–4. The results obtained by numerical

simulation of measurement of both the absolute value and phase of the reflection coefficient are depicted in Figure 2. In turn, the results obtained by numerical simulation of measurement of only the absolute value and phase of the reflection coefficient are depicted in Figure 3 and Figure 4, respectively. In Figure 5, the results obtained for the post with $\varepsilon = 10$ inserted in the waveguide with $a = 100$ mm by simulating measurement of both the absolute value and phase of the reflection coefficient, are depicted. As these graphical results show, there is a variation of the standard uncertainty of measurement of the dielectric constant with the value of the relative radius of the sample r/a . As seen in Figures 2 and 4, the value of the standard uncertainty for certain values of r/a is less than for other values.

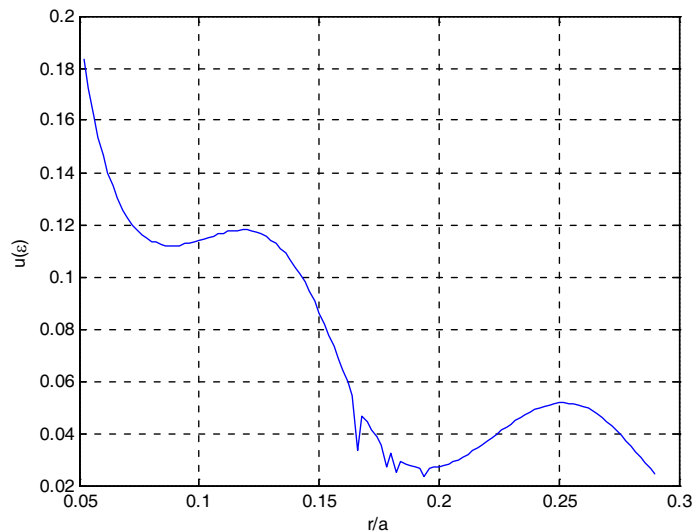


Figure 2: The standard uncertainty $u(\varepsilon)$ of the dielectric constant versus the relative radius of the post r/a , measuring both the absolute value and phase of the reflection coefficient with the following values of standard uncertainties of input quantities: $u(|R|) = 0.05|R|$, $u(\arg(R)) = 1.7$ degrees, $u(r) = 0.01$ mm, $u(f) = 2.4$ MHz.

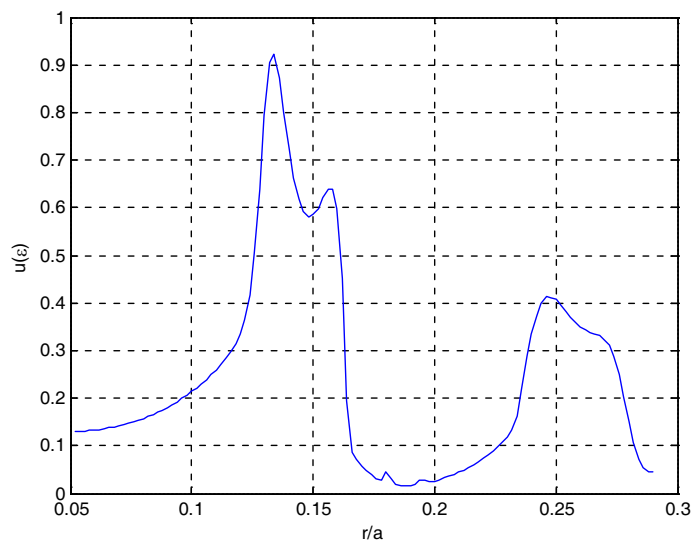


Figure 3: The standard uncertainty $u(\varepsilon)$ of the dielectric constant versus the relative radius of the post r/a , measuring only the absolute value $|R|$ of the reflection coefficient with the following values of standard uncertainties of input quantities: $u(|R|) = 0.05|R|$, $u(r) = 0.01$ mm, $u(f) = 2.4$ MHz.

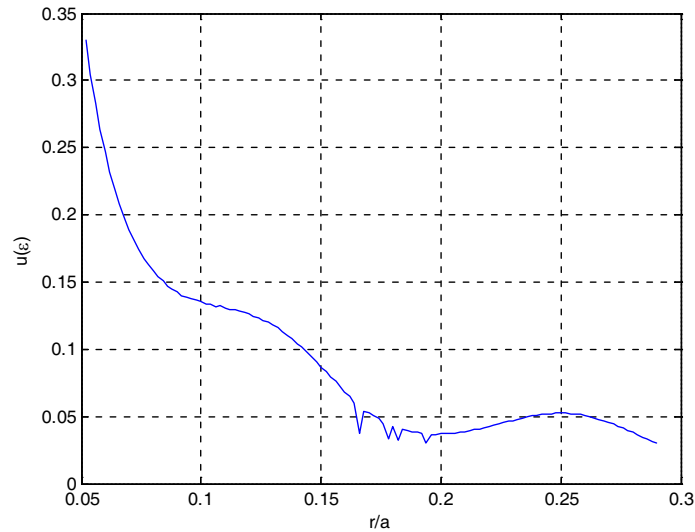


Figure 4: The standard uncertainty $u(\varepsilon)$ of the dielectric constant versus the relative radius of the post r/a , measuring only the phase of the reflection coefficient with the following values of standard uncertainties of input quantities: $u(\arg(R)) - 1.7$ degrees, $u(r) - 0.01$ mm, $u(f) - 2.4$ MHz.

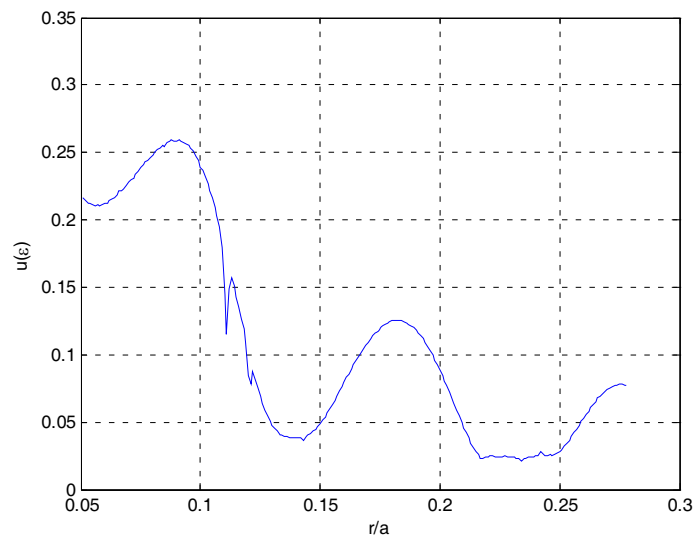


Figure 5: The standard uncertainty $u(\varepsilon)$ of the dielectric constant versus the relative radius of the post r/a , measuring both the absolute value and phase of the reflection coefficient with the following values of standard uncertainties of input quantities: $u(|R|) - 0.05|R|$, $u(\arg(R)) - 1.7$ degrees, $u(r) - 0.01$ mm, $u(f) - 2.4$ MHz.

5. CONCLUSIONS

We have chosen from a number of approaches that have been proposed by various authors the one that provides accurate results and at the same time provides sufficiently fast convergence. In order to find values of the radius of the cylindrical sample at which the value of measurement error has the smallest influence on the accuracy of determination of the dielectric constant, we have applied the Monte Carlo method to find values of the radius of the sample at which the standard uncertainty of the dielectric constant has its smallest value. Uncertainties due to the frequency accuracy, imperfect measurement of the radius of the sample and scattering data have been taken into account, assuming that all these quantities are distributed according to the normal distribution. Numerical results show that the standard uncertainty of measurement of the dielectric constant changes with the value of the relative radius of the sample r/a and for certain values of r/a , the standard uncertainty is smaller than for other values.

REFERENCES

1. Paez, E., M. A. Azpurua, C. Tremola, and R. C. Callarotti, "Uncertainty estimation in complex permittivity measurements by shielded dielectric resonator technique using the monte carlo method," *Progress In Electromagnetics Research B*, Vol. 41, 101–119, 2012.
2. Schwinger, J., Notes on lectures, *Discontinuities in Waveguides*, Gordon and Breach, New York, 1968.
3. Marcuvitz, N., *Waveguide Handbook*, London, 1986.
4. Araneta, J. C., M. E. Brodwin, and G. A. Kriegsmann, "High temperature characterization of dielectric rods," *IEEE Trans. Microwave Theory and Techniques*, Vol. 32, 1328–1334, 1984.
5. Leviatan, Y., P. G. Li, A. T. Adams, and J. Perini, "Single-post inductive obstacle in rectangular waveguide," *IEEE Trans. Microwave Theory and Techniques*, Vol. 31, 1983.
6. Leviatan, Y. and G. F. Sheaffer, "Analysis of inductive dielectric posts in rectangular waveguide," *IEEE Trans. Microwave Theory and Techniques*, Vol. 35, 48–59, 1987.
7. Leviatan, Y. and G. F. Sheaffer, "Composite inductive posts in waveguide — A multifilament analysis," *IEEE Trans. Microwave Theory and Techniques*, Vol. 36, 779–783, 1988.
8. Ise, K. and M. Koshihara, "Dielectric post resonances in a rectangular waveguide," *IEE Proceedings, Part H — Microwaves, Antennas and Propagation*, Vol. 137, 61–66, 1990.
9. Chung, I. G. H. and A. A. Hesham, "Multiple dielectric post in a rectangular waveguide," *IEEE Trans. Microwave Theory Technuques*, Vol. 34, 883–889, 1986.
10. Nielsen, E. D., "Scattering by a cylindrical post of complex permittivity in a waveguide," *IEEE Trans. Microwave Theory Technuques*, Vol. 17, 148–153, 1969.
11. Sahalos, J. N. and E. Vafiadis, "On the narrow-band microwave filter design using dielectric rod," *IEEE Trans. Microwave Theory Technuques*, Vol. 33, 1165–117, 1985.
12. Abdalnour, J. and L. Marchildon, "Scattering by a dielectric obstacle in a rectangular waveguide," *IEEE Trans. Microwave Theory Technuques*, Vol. 41, 1988–1994, 1993.
13. Yu, W. C., "Positive basis and a class of direct search techniques," *Scientia Sinica [Zhongguo Kexue]*, 53–68, 1979.

Image Appraisal of Full Waveform Inverted GPR Data

G. Meles¹, S. Greenhalgh², H. Maurer², and A. Green²

¹School of Geosciences, University of Edinburgh

The Kings Buildings, West Mains Road, Edinburgh EH9 3JN, UK

²Institute of Geophysics, ETH Zurich, Sonneggstrasse 5, CH 8092 Zurich, Switzerland

Abstract— We have recently developed a novel method for explicitly computing the permittivity ε and conductivity σ sensitivity functions (Jacobian matrix \mathbf{J}) of ground penetrating radar (GPR) data, based on a time domain adjoint approach which uses a finite difference modeling method. This not only opens up the possibility for performing Gauss-Newton type inversions, which offer distinct advantages over standard gradient-based approaches, but also permits a methodology for assessing the reliability of inverted GPR images from full waveform data. Image appraisal is performed through a joint analysis of the eigenvalue spectrum of the pseudo Hessian matrix $\mathbf{H} = \mathbf{J}^T \mathbf{J}$ and the formal model resolution matrix $\mathbf{R} = (\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^T \mathbf{J}$, where λ is a damping value to stabilize the matrix inversion. In seismic and geoelectric inversion, the damping factor is often chosen as the median value of \mathbf{H} , but the justification for this is not clear. The diagonal values of \mathbf{R} give the resolution of each cell in the model, with values varying from 0 (unresolved) to 1 (perfectly resolved). The off-diagonal elements convey the trade-offs (cross-coupling) between the different parameters, and the degree of image blurring. The eigenvalue distribution of the pseudo-Hessian matrix provides a measure of the information content of an experiment and shows the unresolved model space. The effect of model perturbation along a given eigenvector direction on the cost function is established in terms of the size of the corresponding eigenvalue. The relative eigenvalue range (RER), which is the width of the normalized spectrum (with entries assembled in descending eigenvalue order) at the level of the noise floor, is a measure of the resolved model space. Four and three-sided radar acquisition geometries (e.g., combination crosshole and borehole-to-surface) yield higher RER values than a one-sided (surface reflection) or two-sided (crosshole) experiment, indicating greater information content and smaller null spaces. Clearly, the better the angular and spatial coverage of the target, the more reliable the image. Resolution is not just a function of the recording geometry and the quantity measured but also the underlying model itself.

We show that cumulative sensitivity (i.e., the column sum of absolute values of the Jacobian) images can be used as a reasonable proxy for formal resolution. Cumulative sensitivity is far less expensive than obtaining the resolution matrix, which involves large matrix inversion and multiplication. We also show that only minor differences exist between the resolution images provided by normalized Jacobians for the full set of ε and σ parameters and the sub-Jacobians for the individual ε and σ values. Permittivity sensitivity values are typically 10^9 times larger than the conductivity values, because they involve a time derivative of the electric field ($i2\pi f$ term, where frequency f is ~ 100 MHz) as opposed the field itself. Without normalization of sensitivities or using sub-Jacobians, conductivity updating would be impossible.

To provide more insightful meaning to resolution, we have undertaken a singular value decomposition of the pseudo-Hessian matrix $\mathbf{W}\Omega\mathbf{D}^T = \mathbf{H}$ and then extract the eigenvectors of \mathbf{D} corresponding to the ‘ a ’ largest singular values (because H is self adjoint, these are identical to the eigenvalues) of Ω : $\mathbf{A}_{ij} = \mathbf{D}_{ij}$, $i = (1, 2, \dots, M)$; $j = (1, 2, \dots, a < M)$. We then form an alternative expression for the resolution matrix $\mathbf{R}^a = \mathbf{A}\mathbf{A}^T$ and take the diagonal elements of \mathbf{R}^a as representative of resolution in each cell. There is a close relationship between resolution provided by \mathbf{R} and \mathbf{R}^a . The effect of the damping factor λ in the formal model resolution formula for \mathbf{R} is basically equivalent to the role of a in the truncated SVD resolution \mathbf{R}^a . Small values of λ have a very similar effect as choosing small values of a . (i.e., just the most dominant eigenvalues). Since SVD resolution is clearly connected to the spectrum of the Hessian, and because the singular values of Ω and the eigenvalues of \mathbf{H} are identical, it provides insight and guidance on the effect of damping in computing \mathbf{R} . However, SVD analysis is extremely expensive from a computational point of view and not intended for routine applications.

In this contribution we illustrate the sensitivity patterns, eigenvalue spectra and resolution plots for a variety of heterogeneous models and recording setups.

1. INTRODUCTION

Most waveform inversions of ground penetrating radar (GPR) data [1] are based on gradient methods [5] that are less expensive computationally than Gauss-Newton or full Newton approaches [3, 4],

because they do not require the inversion of large matrices in the updating process. A critical aspect of any inversion procedure is the assessment of the reliability of the final image. Most often, mere convergence in the data space (i.e., the matching of observed and synthetic GPR traces) is the only criterion used to appraise the goodness of a final result. A better indication of the correctness of an inverted model could be obtained by means of a formal model resolution analysis [3]. This requires the computation of the sensitivity or Jacobian matrix. We recently developed an efficient and novel scheme for computing the permittivity and conductivity sensitivity functions explicitly based on a time-domain adjoint method [2]. The Fréchet derivatives, which form elements of the Jacobian matrix, are obtained by cross correlating forward propagated fields and backward propagated Green's functions from the receiver positions (adjoint sources). The procedure was implemented using a standard finite difference time domain modeling method.

The availability of the Jacobian not only enables formal resolution analysis to be undertaken but also Gauss-Newton style inversions to be performed, which generally converge faster than gradient approaches. In addition, it opens up the possibility to carry out an information content analysis of the radar data by singular value decomposition of the Jacobian (or what is equivalent, an eigenvalue analysis of the pseudo (approximate) Hessian matrix). The eigenvalue spectrum is a useful tool in experimental design because it conveys the extent of the resolved model space and permits comparison of the efficacy of different recording configurations and data sets.

2. EXPLICIT EXPRESSIONS FOR THE SENSITIVITY KERNELS

The GPR sensitivity functions for each model cell or block in the subsurface are a measure of how a perturbation in the electrical property (permittivity ε or conductivity σ) in that cell causes a perturbation in the measured electric field for a given radar transmitter-receiver configuration. The explicit formulae for the columns of the sensitivity functions for any receiver location \mathbf{x}_r and observation time t_r were derived in [2] and can be stated thus:

$$[\mathbf{J}_\varepsilon^S(x')]_{x_r, t_r}^i = \langle \delta(\mathbf{x} - \mathbf{x}') \partial_t \mathbf{E}^S, \mathbf{G}^T \delta^i(\mathbf{x} - \mathbf{x}_r, t - t_r) \rangle \quad (1)$$

$$[\mathbf{J}_\sigma^S(x')]_{x_r, t_r}^i = \langle \delta(\mathbf{x} - \mathbf{x}') \mathbf{E}^S, \mathbf{G}^T \delta^i(\mathbf{x} - \mathbf{x}_r, t - t_r) \rangle \quad (2)$$

where the bracket symbol $\langle \rangle$ indicates inner product. Here the superscript i indicates the individual vectorial components, the subscripts ε and σ denote either a permittivity or conductivity derivative, \mathbf{x}' is any domain point, \mathbf{E}^S is the electric field in the whole space-time domain, \mathbf{G} is the Green's operator, T indicates transposition, and the superscript 's' on E refers to a particular source. We now concentrate on the right side term $\mathbf{G}^T \delta^i(\mathbf{x} - \mathbf{x}_r, t - t_r)$ in Eqs. (1) and (2) where the adjoint source at the receiver has a single component (i.e., electric dipole direction) that corresponds to the same component of the sensitivity in which we are interested, namely $\partial E^i(x_r, t_r) / \partial \varepsilon(x')$ or $\partial E^i(x_r, t_r) / \partial \sigma(x')$. The inner products in (1) and (2) involve an integration over space and time. In these equations, the argument of the delta function, $\mathbf{x} - \mathbf{x}'$, indicates the point at which we are computing the sensitivities. The dummy variables over which the integration implicit in (1) and (2) takes place are indicated in the following by x^* and t^* .

$$[\mathbf{G}^T \delta(x - x_r, t - t_r)_{x^*, t^*}^i]^j = G^{ij}(x^*, t_r, x_r, t^*) \quad (3)$$

The inner product sums over the j different field components, which are forward and back propagated in (1) and (2). Due to reciprocity (3) can be written as

$$[\mathbf{G}^T \delta(x - x_r, t - t_r)_{x^*, t^*}^i]^j = G^{ji}(x^*, t_r, x_r, t^*) \quad (4)$$

These are the different components of the wavefield generated by the delta-function source oriented along a single direction i . From time invariance, we can write (3) as:

$$G^{ji}(x^*, t_r, x_r, t^*) = G^{ji}(x^*, t_r - t^*, 0) \quad (5)$$

The computation of (5) involves the solution of a standard forward problem, with the source placed at x_r and discharged at $t = 0$. Depending on the value of t_r , different wavefields $G^{ji}(x^*, t_r - t^*, x_r, 0)$ will be cross correlated with $\partial_t \mathbf{E}^s(x^*, t^*)$ and $\mathbf{E}^s(x^*, t^*)$. The presence of $\delta(\mathbf{x} - \mathbf{x}')$ in (1) and (2) reduces in practice the integration to that over time alone.

Practical issues of back propagating a Green's delta function in an FDTD scheme can be circumvented due to the orthogonality of the Fourier base (any frequency component not common to both vectors in an inner product has no effect on the result) so it is only necessary to back propagate a filtered delta function in accordance with the frequencies contained in the source spectrum [2].

The size of the Jacobian is $N \times M$, where N is the total number of measurements (number of sources \times number of receivers \times number of times along each trace) and M is the number of model parameters (number of cells $\times 2$, if both permittivity and conductivity are considered). The temporal derivative of the electric field appearing in (1), suggests that the GPR permittivity sensitivities will be much higher (by $2\pi f$, $f \approx 100$ MHz) than the corresponding conductivity sensitivities (which depend on the electric field itself), and so it is necessary to normalize the Jacobian or work with sub-Jacobians for each class of model parameter.

Figures 1(a), (b) show a heterogeneous model involving two cross-shaped anomalous bodies embedded in a homogeneous background of relative permittivity $\epsilon_r = 4$ and conductivity $\sigma = 3$ mS/m. Cross 1 is resistive ($\sigma = 0.1$ mS/m) with low $\epsilon_r = 1$, whereas cross 2 is conductive ($\sigma = 10$ mS/m) with high ϵ_r . The waveform in Fig. 1(c) represents the GPR trace for the crosshole-polarized recording configuration as shown, with transmitter and receiver both located at a depth of 5.3 m and 10 m apart. Apart from the direct transmitted arrival, there are additional forward scattered signals present. Point A corresponds to the first arrival, B to an arbitrary relative

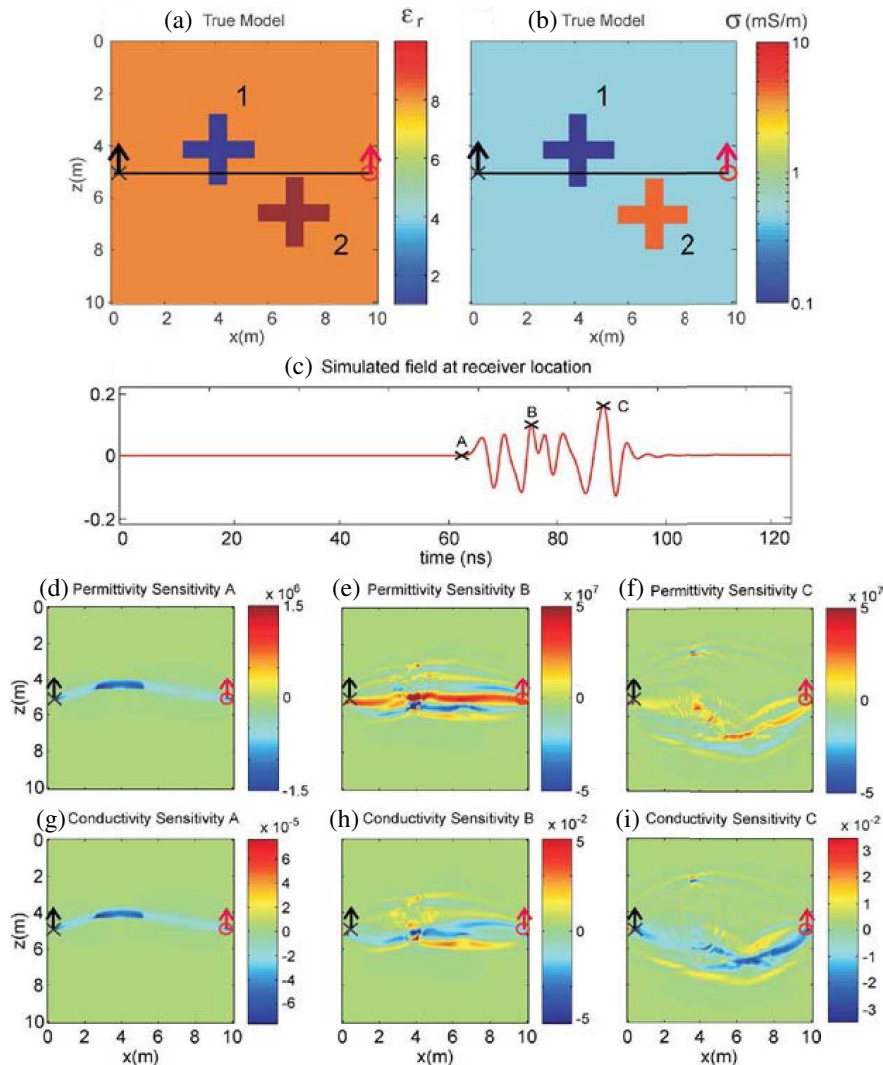


Figure 1: (a) Permittivity ϵ and (b) conductivity σ distributions for heterogeneous model. (c) Radar trace for Tx and Rx antennae positions as shown. (d)–(i) show the ϵ and σ sensitivity patterns for the three observation times A, B and C indicated in (c). Units of ϵ permittivity sensitivity are $V^2/A \cdot s$ and σ sensitivity are V^2/A .

amplitude maximum and C to the absolute maximum amplitude of the entire trace. The ε and σ sensitivities throughout the medium are shown in Figs. 1(d)–(f) and 1(g)–(i) for these three time samples A, B and C. There are significant amplitude differences between the pairs of corresponding sensitivity plots. The sensitivity patterns are significantly distorted from the quasi-elliptical shape expected in a homogenous medium. As time increases, the sensitivity patterns expand to occupy a larger region of the model.

3. INFORMATION CONTENT ANALYSIS

The availability of the Jacobian values for various receiver locations and observation times (Eqs. (1) and (2)) also allows the computation of the pseudo-Hessian matrix, given by $\mathbf{H}^A = \mathbf{J}^T \mathbf{J}$. The size of this matrix is $M \times M$. The inverse of this matrix is used in Gauss-Newton inversion as a more refined version of a step length operator common in gradient based inversion schemes. It specifies the extent of model updating in the gradient direction. The eigenvalue spectrum of the pseudo Hessian also reveals the information content of a data set, and can be used to appraise the imaging capability of different recording configurations. Fig. 2 shows in semi-log form the normalized eigenvalues, arranged in descending order, for the model of Fig. 1 but for four different recording configurations: surface, cross-hole, three-sided and four sided. The surface and crosshole experiments involve 5 transmitters and 5 receivers placed uniformly along the upper side of the model (surface case) or in the left and right sides of the model (crosshole), whereas for the 3-sided experiment 15 transmitters and 15 receivers are placed evenly along the upper and lateral sides of the model. For the 4-sided experiment, and additional 5 transmitters and 5 receivers are placed along the bottom side of the model. The three separate plots in Fig. 2 are for the complete model space (ε and σ together), and the separate model spaces (ε only, and σ only). The intersection of the eigenvalue spectra with the horizontal threshold line (relative eigenvalue = 10^{-7}) delineate the resolved model space and the mostly unresolved model space. This threshold corresponds to a noise floor below which no significance should be attached to the eigenvalues. The portion to the left of the intersection point is referred to as the relative eigenvalue range (RER) and it provides a simple measure of the goodness of a particular experimental design. The increasing size of the RER for the four different experiments (with the highest for the 4 sided experiment) and second highest for the 3-sided experiment give a clear indication of the increasing reliability of the inverted models for the different configurations. This can be linked to the improved angular and spatial coverage of the target.

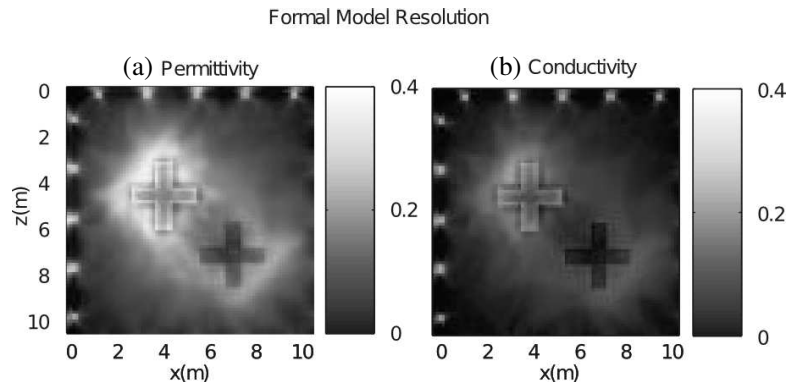


Figure 2: Formal model resolution for (a) permittivity and (b) conductivity for the four-sided experiment.

4. FORMAL RESOLUTION AND CUMULATIVE SENSITIVITY

The resolution of each cell within an inverted model can be expressed formally by the model resolution matrix \mathbf{R} defined as [3]:

$$\mathbf{R} = (\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^T \mathbf{J} \quad (6)$$

where λ is a damping factor (to ensure that the matrix can be inverted) and \mathbf{I} is the identity matrix. \mathbf{R} relates the estimated model parameters m^{est} with the true model parameters m^{true} through the relationship $m^{est} = \mathbf{R}m^{true}$. Of particular interest are the diagonal elements of \mathbf{R} . Values close

to 0 indicate poorly resolved parameters, whereas values close to 1 indicate well resolved model parameters. The rows of \mathbf{R} indicate the degree of smearing or point spread functions for each cell.

Figure 3 shows the formal model resolution images obtained for the 4-sided experiment on the double cross model (Fig. 1) and described above. The two plots correspond to the permittivity and conductivity sub-Jacobians, and illustrate better resolution for the resistive and low permittivity cross.

A surrogate or proxy for assessing formal resolution is the cumulative sensitivity function, obtained by summing over all transmitters, receivers and observation times the absolute values of the sensitivities, i.e., a column sum of the Jacobian, given by:

$$\mathbf{CS}(m) = \sum_s \sum_r \sum_\tau |\mathbf{J}^s(m)|_{x_r, t_r} \quad (7)$$

The quantity CS is rarely used in GPR or seismic waveform inversion but its appeal is that it does not require any matrix multiplication or inversion, and is therefore relatively inexpensive. Space limitations preclude showing comparison CS plots to Fig. 3. Although differences exist there is a strong similarity in shape, and high cumulative sensitivity values are concentrated in the same regions of the model where formal resolution is high.

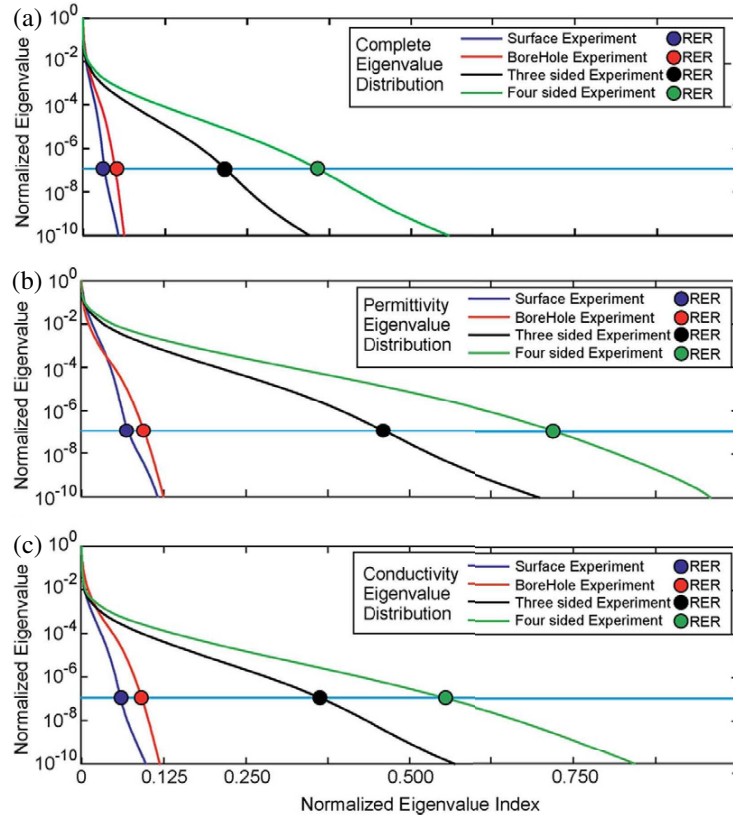


Figure 3: Eigenvalue spectra (a) ϵ and σ together, (b) ϵ separate, and (c) σ separate for the four different recording configurations.

5. SVD RESOLUTION

An alternative to the definition of formal resolution can be provided by means of a singular value decomposition (SVD) of the pseudo-Hessian matrix, given by

$$\mathbf{W}\mathbf{\Omega}\mathbf{D}^T = \mathbf{H}^A \quad (8)$$

and then extract the eigenvectors of \mathbf{D} corresponding to the ‘ a ’ largest singular values (or eigenvalues) of $\mathbf{\Omega}$:

$$A_{ij} = D_{ij}, \quad i = (1, 2, \dots, M), \quad j = (1, 2, \dots, a < M) \quad (9)$$

We then form the following alternative expression for the resolution matrix, referred to as SVD resolution:

$$\mathbf{R}^a = \mathbf{A}\mathbf{A}^T \quad (10)$$

and take the diagonal elements of \mathbf{R}^a as representative of resolution in each cell. Values close to zero indicate poorly resolved model parameters, whereas values close to one indicate well-resolved model parameters. Both the natural model space and that expressed by the columns of \mathbf{D} are orthonormal.

Each vector of the natural base, m_i , can be expressed as a linear combination of vectors in the orthonormal base provided by \mathbf{D} :

$$m_i = \sum_{k=1}^M c_k \mathbf{v}_k \quad (11)$$

where \mathbf{v}_k are the orthonormal columns of \mathbf{D} and c_k are the numerical coefficients of the i -th column of \mathbf{D}^{-1} . Since $\mathbf{D}^{-1} = \mathbf{D}^T$, it is straightforward to see that

$$R_{i,i}^a = \sum_{k=1}^a c_k^2 = \sum_{k=1}^a D_{ik}^2 \quad (12)$$

Because \mathbf{D} consists of orthonormal columns, $R_{i,i}^a$ is always smaller than 1.

More specifically, the resolution valued defined in XX is equivalent to the relative amplitude of the projection of the natural base element m_i on the space spanned by the eigenvectors corresponding to the a largest singular values (i.e., the best resolved ones). There is a very close relationship between the resolution provided by (6) and (10). The effect of the damping factor in (6) is basically equivalent to the role of a . Small values of λ correspond to large values of a , whereas large values of λ have a very similar effect as choosing small values of a (i.e., just the most dominant eigenvalues).

Since SVD resolution is clearly connected to the spectrum of the Hessian, and because the singular values of Ω and the eigenvalues of \mathbf{H}^A are identical, it provides an insightful meaning to resolution. However, since SVD is extremely expensive from a computational point of view, and not intended for routine practical applications, formal resolution analysis is most often presented.

Figure 4 gives a comparison between formal resolution (6) and SVD resolution (10) for the same heterogeneous model comprising the two anomalous crosses. The upper set of diagrams (a)–(e) depict formal model resolution images for different choices of the damping factor λ , in the range 0.01 to 100 times the median value of the diagonal value of \mathbf{H}^A . By comparison, the bottom set of diagrams (f)–(j) show SVD resolution for different sets of eigenvalues ranging from the 80 largest ones to the 3000 largest ones (from a total of 7744). The (a)–(e) and (f)–(j) diagrams of Fig. 4 are extremely similar, establishing a tight connection between the two definitions of resolution.

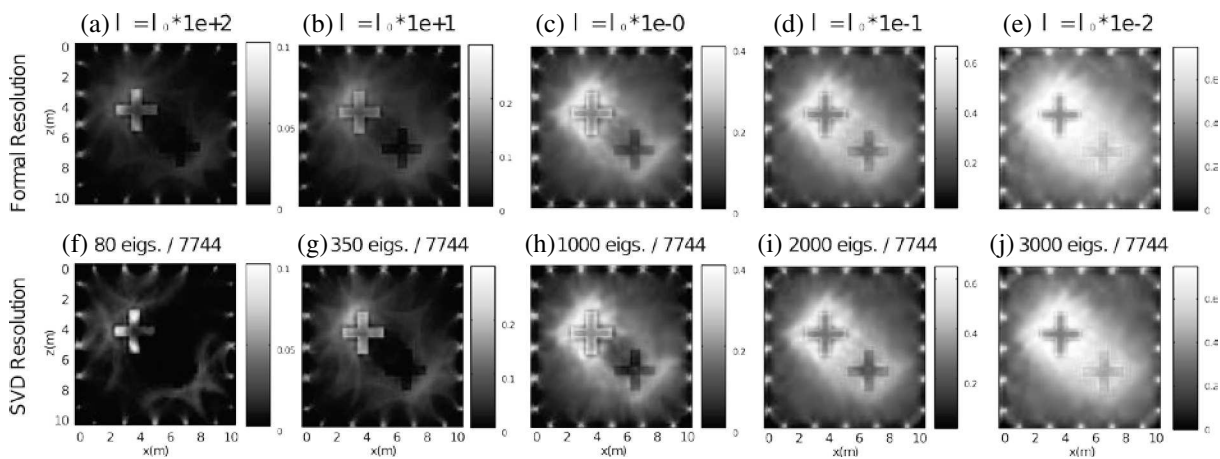


Figure 4: (a)–(e) Formal model resolution images for the 4-sided experiment and 5 different choices of the damping parameter λ . λ_0 corresponds to the median value of the diagonal elements of the pseudo-Hessian. (f)–(j) Corresponding SVD resolution for 5 different choices of the number of eigenvalues used for the definition of matrix A .

Moreover, the differences arising in these images as a function of the damping factor or the number of eigenvalues used suggests that particular care must be taken when analyzing resolution plots.

6. CONCLUSIONS

We have developed a new and effective method for calculating GPR sensitivity functions in the time domain. This allows us to assess the reliability of inverted images from GPR full waveform data and the adequacy of an experimental setup. The usual criterion of simple convergence in the data space is insufficient to appreciate the goodness of a model. Through an analysis based on the eigenvalue distribution of the Hessian, the cumulative sensitivity pattern and the formal or SVD resolution matrix it is possible to provide a meaningful estimate of the well resolved and poorly resolved features of an inverted image.

REFERENCES

1. Meles, G., S. A. Greenhalgh, J. van der Kruk, and A. G. Green, “Taming the non-linearity problem in GPR full-waveform inversion for high contrast media,” *J. App. Geophys.*, Vol. 73, 174–186, 2011.
2. Meles, G., S. A. Greenhalgh, A. G. Green, H. Maurer, and J. van der Kruk, “GPR full waveform sensitivity and resolution analysis using an FDTD adjoint method,” *IEEE Trans. Geosci. Remote Sensing*, Vol. 50, No. 5, 1881–1896, 2012.
3. Menke, W., *Geophysical Data Analysis. Discrete Inverse Theory*, Academic Press, London, 1989.
4. Pratt, R. G., C. Shin, and G. J. Hicks, “Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion,” *Geophys. J. Int.*, Vol. 133, No. 2, 341–362, 1998.
5. Tarantola, A., “Inverse problem theory and methods for model parameter estimation,” SIAM: Society for Industrial and Applied Mathematics, 2004.

Guaranteed Estimates of Linear Continuous Functionals of Solutions and Right-hand Sides of the Helmholtz Equation in the Domains with Infinite Boundaries under Uncertainties

Yury Podlipenko¹ and Y. Shestopalov²

¹Kiev National University, Ukraine

²Karlstad University, Sweden

Abstract— We consider the construction of guaranteed estimates [1, 2] of linear continuous functionals of the unknown solutions and right-hand sides of the Helmholtz equation [2, 3]; the boundary value problems under study are associated with the wave diffraction by a bounded body D situated in a domain $\Omega \in \mathbb{R}^n$, $n = 2, 3$, whose boundary $\partial\Omega$ stretches to infinity (e.g., a wedge or a layer) and Green's function $\Phi_k(x, y)$, $(x, y \in \Omega, x \neq y)$ corresponding to wave number k with $k > 0$ and boundary condition $B\Phi_k(x, y)|_{y \in \partial\Omega} = 0$ is known [4]. Here, for a function $u(y)$ defined in $\bar{\Omega}$ $Bu(y)|_{y \in \partial\Omega} := \alpha u(y)|_{y \in \partial\Omega} + \beta \frac{\partial u(y)}{\partial \nu}|_{y \in \partial\Omega}$, $\alpha, \beta = 0, 1, \alpha + \beta = 1$, ν is outward normal to $\partial\Omega$. We assume that right-hand sides of the equations entering the problem statement are not known; the only available information is that they belong to a bounded set of the space of square-integrable functions. In order to solve these estimation problems we need additional data: observations in the form of certain linear transformations of the solution distorted by noise. The latter are realizations of the random fields with the unknown second moment functions belonging to a given bounded set in the appropriate functional space. The approach set forth in [2, 3, 5] and developed in this study allows us to obtain optimal estimates of the unknown solution or right-hand sides of the equations and linear functionals, i.e., estimates sought in the class of functionals linear with respect to observations for which the maximal mean-square estimation error taken over all elements belonging to the aforementioned sets takes minimal value. Such estimates are called minimax or guaranteed estimates. We obtain representations for these estimates and estimation errors in terms of solutions to certain integro-differential or integral equations in bounded subdomains of domain $\Omega \setminus D$.

1. INTRODUCTION

Forward and especially inverse problems in electromagnetics aim at estimating model parameters of a system under study (antenna, waveguide-based device and the like) from available observations (data) of its response or output [6, 7]. Examples can be found in wave tomography or geophysics when the model can be formulated and the forward problem solved only if the (elastic or optical) properties of the medium are given in advance. However, they are usually not known which leads to the necessity of solving a certain type of inverse problem if measurements of wave amplitudes and phases at given points on the accessible sets are available. Using these data, an appropriate estimation algorithm can be formulated. In addition, data measurements are affected by errors depending on both controllable and uncontrollable factors such as precision of the adopted instrumentation or the environmental conditions. Using deterministic inverse algorithms one can hardly obtain error bounds on the parameter estimates given uncertainties in the data. On the other hand, probabilistic approaches, e.g., as in Kalman Filter estimation, that may allow identifying the propagation of uncertainties and provide error estimations have their own limitations since they require a prior assumption on the nature of uncertainties. It is desirable therefore to develop mathematically justified techniques that do not rely on the type of uncertainties and enable one to treat the estimation problem in its complete statement. This is the subject of the present study.

2. STATEMENT OF THE PROBLEM OF GUARANTEED ESTIMATION

Let $\varphi(x)$ be the state of the system governed by the Helmholtz problem

$$(\Delta + k^2) \varphi(x) = 0 \quad \text{in } \Omega \setminus \bar{D}, \quad (1)$$

$$\frac{\partial \varphi}{\partial \nu_A} = h \quad \text{on } \Gamma = \partial D, \quad B\varphi = 0 \quad \text{on } \partial\Omega, \quad (2)$$

where $h \in H^{1/2}(\Gamma)$, the function φ is a limit as $q \rightarrow k^2 + i0$ (in the sense of convergence in the space $H_{\text{loc}}^2(\Omega \setminus \bar{D})$) of a unique solution $\varphi_q \in L^2(\Omega \setminus \bar{D})$ of problem (1)–(2), in which k^2 is replaced by $q \in \mathbb{C}$ in Equation (1).¹

Let

$$G_0 := \left\{ \tilde{h} \in H^{1/2}(\Gamma) : \int_{\Gamma} |\tilde{h} - h_0|^2 g^2 d\Gamma \leq 1 \right\},$$

where $h_0 \in H^{1/2}(\Gamma)$ is a given function, g is a smooth function on Γ , that do not vanish there, and let $x'_k, k = \overline{1, N}$ and $x_k, k = \overline{1, m}$ be given systems of points belonging to domain $\Omega \setminus \bar{D}$. The problem is as follows: to estimate the expression

$$l(\varphi) = \sum_{i=1}^m \bar{a}_i \varphi(x_i), \quad (a_i \in \mathbb{C} \text{ are prescribed numbers}) \quad (3)$$

from the point observations of the form²

$$y_k = \varphi(x'_k) + \eta_k, \quad k = \overline{1, N}, \quad (4)$$

in the class of estimates

$$\widehat{l(\varphi)} = \sum_{k=1}^N \bar{u}_i y_k + c,$$

linear with respect to observations (4), where $u_i \in \mathbb{C}, i = \overline{1, N}, c \in \mathbb{C}$, under the assumptions that h and $\eta := (\eta_1, \dots, \eta_N)$ are not known exactly but it is only known that

$$h \in G_0 \quad \text{and} \quad \eta \in G_1.$$

Here, by G_1 we denote the set of random vectors $\tilde{\eta} = (\tilde{\eta}_1, \dots, \tilde{\eta}_N)$ with zero expectations, $\mathbb{E}\tilde{\eta}_i = 0, i = \overline{1, N}$, and finite second moments satisfying the condition

$$\sum_{i=1}^N r_i^2 \mathbb{E}|\tilde{\eta}_i|^2 \leq 1,$$

$r_i \in \mathbb{R}, r_i \neq 0, i = \overline{1, N}$, are given numbers, and \mathbb{E} denotes expectation of random variable.

Set $u := (u_1, \dots, u_N) \in \mathbb{C}^N$.

Definition 1. The estimate

$$\widehat{l(\varphi)} = \sum_{i=1}^N \bar{\hat{u}}_i y_i + \hat{c},$$

in which numbers \hat{u}_i and \hat{c} are determined from the condition

$$\inf_{u \in \mathbb{C}^N, c \in \mathbb{C}} \sigma(u, c) = \sigma(\hat{u}, \hat{c}),$$

where

$$\sigma(u, c) := \sup_{\tilde{h} \in G_0, \tilde{\eta} \in G_1} \mathbb{E} \left| l(\tilde{\varphi}) - \widehat{l(\tilde{\varphi})} \right|^2,$$

$$\widehat{l(\tilde{\varphi})} = \sum_{i=1}^N \bar{u}_i \tilde{y}_i + c,$$

$$\tilde{y}_i = \tilde{\varphi}(x_i) + \tilde{\eta}_i, \quad i = \overline{1, N},$$

and $\tilde{\varphi}(x)$ is the solution to the Neumann BVP at $h = \tilde{h}$, will be called the guaranteed estimate of expression (3).

The quantity $\sigma := (\sigma(\hat{u}, \hat{c}))^{1/2}$ will be called the error of the estimation of $l(\varphi)$.

Thus, the minimax estimate is an estimate minimizing the maximal mean-square estimation error calculated for the “worst” implementation of perturbations.

¹Here, we consider a class of domains $\Omega \setminus \bar{D}$, for which limiting absorption principle is justified (see, for example, [8–10]).

² η_i are errors of observations (4) that are realizations of random variables $\eta_i = \eta_i(\omega)$.

3. MAIN RESULTS

Theorem 1. *The guaranteed estimate of $l(\varphi)$ has the form*

$$\widehat{l(\varphi)} = \sum_{i=1}^N \widehat{u}_i y_i + \widehat{c} = l(\widehat{\varphi}),$$

where

$$\widehat{u}_k = r_k^2 p(x'_k), \quad k = \overline{1, N}, \quad \widehat{c} = \int_{\Gamma} \bar{z} h_0 d\Gamma,$$

the functions z , p , and $\widehat{\varphi}$ are determined, respectively, from the solution to the following problems:

$$-(\Delta + k^2)z(x) = \sum_{i=1}^m a_i \delta(x - x_i) - \sum_{k=1}^N r_k^2 p(x'_k) \delta(x - x'_k) \quad \text{in } \Omega \setminus \bar{D}, \quad (5)$$

$$\frac{\partial z}{\partial \nu} = 0 \quad \text{on } \Gamma, \quad Bz = 0 \quad \text{on } \partial\Omega, \quad (6)$$

$$\Delta p(x) + k^2 p(x) = 0 \quad \text{in } \Omega \setminus \bar{D}, \quad (7)$$

$$\frac{\partial p}{\partial \nu} = g^{-2} z \quad \text{on } \Gamma, \quad Bp = 0 \quad \text{on } \partial\Omega, \quad (8)$$

and

$$-(\Delta + k^2)\widehat{p}(x) = \sum_{i=1}^N r_i^2 [y(x'_i) - \widehat{\varphi}(x'_i)] \delta(x - x'_i) \quad \text{in } \Omega, \quad (9)$$

$$\frac{\partial \widehat{p}}{\partial \nu} = 0 \quad \text{on } \Gamma, \quad B\widehat{p} = 0 \quad \text{on } \partial\Omega, \quad (10)$$

$$\Delta \widehat{\varphi}(x) + k^2 \widehat{\varphi}(x) = 0 \quad \text{in } \Omega, \quad (11)$$

$$\frac{\partial \widehat{\varphi}}{\partial \nu} = g^{-2} \widehat{p} + h_0 \quad \text{on } \Gamma, \quad B\widehat{\varphi} = 0 \quad \text{on } \partial\Omega, \quad (12)$$

where the functions z , p and \widehat{p} , $\widehat{\varphi}$ are the limits as $q \rightarrow k^2 + i0$ of functions z_q , p_q and \widehat{p}_q , $\widehat{\varphi}_q$ ($z_q, p_q, \widehat{p}_q, \widehat{\varphi}_q \in L^2(\Omega \setminus \bar{D})$) satisfying problems (5)–(8) and (9)–(12)³ in which k^2 is replaced by \bar{q} in Equations (5), (9) and by $q \in \mathbb{C}$ in (7), (11), respectively⁴.

Problems (5)–(8) and (9)–(12) are uniquely solvable. The estimation error σ of $l(\varphi)$ is determined by

$$\sigma = [l(p)]^{1/2} = \left(\sum_{i=1}^m \bar{a}_i p(x_i) \right)^{1/2}. \quad (13)$$

Using the potential theory in Sobolev spaces we reduce problems (5)–(8) and (9)–(12) to problems of less dimensionality.

To obtain this, introduce the single- and double-layer potentials with density φ by

$$(\mathcal{V}_k \psi)(x) := \int_{\Gamma} \Phi_k(x, y) \psi(y) d\Gamma_y, \quad x \in \Omega \setminus \Gamma,$$

³Here, $z_q \rightarrow z$, $\widehat{p}_q \rightarrow \widehat{p}$ in $H_{\text{loc}}^2(\Omega \setminus \bar{D}_1)$ where $D_1 := D \cup \{x_1\} \cup \dots \cup \{x_m\} \cup \{x'_1\} \cup \dots \cup \{x'_N\}$, and $p_q \rightarrow p$, $\widehat{\varphi}_q \rightarrow \widehat{\varphi}$ in $H_{\text{loc}}^2(\Omega \setminus \bar{D})$.

⁴This requirement in the case when $\Omega = \{(r, \theta, \phi) \in \mathbb{R}^3: 0 < r < +\infty, 0 < \theta < \pi, 0 < \phi < \Phi\}$ is a wedge in spherical coordinate system $0 \leq r < +\infty, 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$ with the spread angle $\Phi, 0 < \Phi \leq 2\pi$, can be replaced by the following: the functions z , \widehat{p} and p , $\widehat{\varphi}$ must satisfy the Sommerfeld condition

$$\frac{\partial z}{\partial r} + ikz = o(1/r), \quad \frac{\partial \widehat{p}}{\partial r} + ik\widehat{p} = o(1/r), \quad r = |x|, \quad r \rightarrow \infty$$

and

$$\frac{\partial p}{\partial r} - ikp = o(1/r), \quad \frac{\partial \widehat{\varphi}}{\partial r} - ik\widehat{\varphi} = o(1/r), \quad r = |x|, \quad r \rightarrow \infty,$$

respectively.

$$(\mathcal{W}_k\chi)(x) := \int_{\Gamma} \frac{\partial\Phi_k(x,y)}{\partial\nu_y} \chi(y) d\Gamma_y, \quad x \in \Omega \setminus \Gamma,$$

where φ is a given integrable function on Γ .

We also introduce the single- and double-layer operators S_k and K_k , given by

$$\begin{aligned} (S_k\varphi)(x) &:= 2 \int_{\Gamma} \Phi_k(x,y) \varphi(y) d\Gamma_y, \quad x \in \Gamma, \\ (K_k\varphi)(x) &:= 2 \int_{\Gamma} \frac{\partial\Phi_k(x,y)}{\partial\nu_y} \varphi(y) d\Gamma_y, \quad x \in \Gamma, \end{aligned}$$

and the normal derivative operators K'_k and T_k , given by

$$\begin{aligned} (K'_k\varphi)(x) &:= 2 \int_{\Gamma} \frac{\partial\Phi_k(x,y)}{\partial\nu_x} \varphi(y) d\Gamma_y, \quad x \in \Gamma, \\ (T_k\varphi)(x) &:= 2 \frac{\partial}{\partial\nu_x} \int_{\Gamma} \frac{\partial\Phi_k(x,y)}{\partial\nu_y} \varphi d\Gamma_y, \quad x \in \Gamma. \end{aligned}$$

Denote by I identity operator. Then the following result is valid.

Theorem 2. *The guaranteed estimate of $l(\varphi)$ has the form*

$$\widehat{l(\varphi)} = \sum_{k=1}^N \hat{u}_k y_k + \hat{c} = l(\hat{\varphi}) = \sum_{i=1}^m \bar{a}_i \hat{\varphi}(x_i),$$

where

$$\begin{aligned} \hat{c} &= \frac{\hat{u}_k = r_k^2 p(x'_k), \quad k = \overline{1, N}}{\int_{\Gamma} \left[\psi + \left(\sum_{j=1}^m a_j \Phi_{-\bar{k}}(\cdot - x_j) - \sum_{l=1}^N r_l^2 p(x'_l) \Phi_{-\bar{k}}(\cdot - x'_l) \right) \Big|_{\Gamma} \right]} h_0 d\Gamma, \\ \hat{\varphi} &= \mathcal{W}_k \tilde{\chi} - \mathcal{V}_k \left[g^{-2} \left(\tilde{\psi} + \sum_{l=1}^N r_l^2 (y_l - \hat{\varphi}(x'_l)) \Phi_{-k}(\cdot, x'_l) \Big|_{\Gamma} + h_0 \right) \right] \quad \text{in } \Omega \setminus \bar{D}, \end{aligned}$$

the functions

$$\psi := z|_{\Gamma} - \left(\sum_{j=1}^m a_j \Phi_{-\bar{k}}(\cdot - x_j) - \sum_{l=1}^N r_l^2 p(x'_l) \Phi_{-\bar{k}}(\cdot - x'_l) \right) \Big|_{\Gamma},$$

$\chi := p|_{\Gamma}$, numbers $p(x'_l)$, $l = \overline{1, N}$, and the functions

$$\tilde{\psi} := \hat{p}|_{\Gamma} - \sum_{l=1}^N r_l^2 (y_l - \hat{\varphi}(x'_l)) \Phi_{-k}(\cdot, x'_l) \Big|_{\Gamma},$$

$\tilde{\chi} := \hat{\varphi}|_{\Gamma}$, numbers $\hat{\varphi}(x'_l)$, $l = \overline{1, N}$, are determined, respectively, from the solution of the problems (14)–(16) and (17)–(19):

$$\begin{aligned} &(I - K'_{-\bar{k}} + i\eta T_{-\bar{k}}) \psi \\ &= \left(-S_{-\bar{k}} + i\eta (I + K'_{-\bar{k}}) \right) \left(- \sum_{j=1}^m a_j \frac{\partial\Phi_{-\bar{k}}(\cdot - x_j)}{\partial\nu} \Big|_{\Gamma} + \sum_{l=1}^N r_l^2 p(x'_l) \frac{\partial\Phi_{-\bar{k}}(\cdot - x'_l)}{\partial\nu} \Big|_{\Gamma} \right), \quad (14) \end{aligned}$$

$$\begin{aligned} &(I - K_k - i\eta T_k) \chi \\ &= (-S_k - i\eta (I + K'_k)) g^{-2} \left(\psi + \sum_{j=1}^m a_j \Phi_{-\bar{k}}(\cdot - x_j) \Big|_{\Gamma} - \sum_{l=1}^N r_l^2 p(x'_l) \Phi_{-\bar{k}}(\cdot - x'_l) \Big|_{\Gamma} \right), \quad (15) \end{aligned}$$

$$p(x'_i) = \mathcal{W}_k \chi(x'_i) - \mathcal{V}_k g^{-2} \left(\psi + \sum_{j=1}^m a_j \Phi_{-\bar{k}}(\cdot - x_j) \Big|_{\Gamma} - \sum_{l=1}^N r_l^2 p(x'_l) \Phi_{-\bar{k}}(\cdot - x'_l) \Big|_{\Gamma} \right) \Big|_{x'_i}, \quad i = \overline{1, N}. \quad (16)$$

and

$$(I - K'_{-k} + i\eta T_{-k}) \tilde{\psi} = (-S_{-k} + i\eta (I + K'_{-k})) \left(\sum_{l=1}^N r_l^2 (\hat{\varphi}(x'_l) - y_l) \frac{\partial \Phi_{-k}(\cdot, x'_l)}{\partial \nu} \Big|_{\Gamma} \right), \quad (17)$$

$$(I - K_k - i\eta T_k) \tilde{\chi} = (-S_k - i\eta (I + K'_k)) \left[g^{-2} \left(\tilde{\psi} + \sum_{l=1}^N r_l^2 (y_l - \hat{\varphi}(x'_l)) \Phi_{-k}(\cdot, x'_l) \Big|_{\Gamma} \right) + h_0 \right], \quad (18)$$

$$\hat{\varphi}(x'_i) = \mathcal{W}_k \tilde{\chi}(x'_i) - \mathcal{V}_k \left[g^{-2} \left(\tilde{\psi} + \sum_{l=1}^N r_l^2 (y_l - \hat{\varphi}(x'_l)) \Phi_{-k}(\cdot, x'_l) \Big|_{\Gamma} \right) + h_0 \right] \Big|_{x'_i}, \quad i = \overline{1, N}. \quad (19)$$

in which η is an arbitrary positive number. These problems are uniquely solvable for all values of wave numbers $k > 0$.

Similar estimates are obtained for unknown right-hand side h from observations (4).

We generalized the obtained results for the case when linear functional (3) is replaced by a functional of the form

$$l(\varphi) = \int_{\Omega_0} \overline{l_0(x)} \varphi(x) dx$$

where $l_0 \in L^2(\Omega_0)$ is a given function defined in a bounded subdomain Ω_0 such that $\bar{\Omega}_0 \subset \Omega \setminus \bar{D}$ and the observations distributed on a finite system of surfaces or subdomains located in $\Omega \setminus \bar{D}$ are taken instead of point observations (4).

REFERENCES

1. Nakonechnyi, O., Y. Podlipenko, and Y. Shestopalov, "Estimation of parameters of boundary value problems for linear ordinary differential equations with uncertain data," arXiv:0912.2872, 72 p., 2009.
2. Podlipenko, Y., Y. Shestopalov, and V. Prishlyak, "Estimation under uncertainties of acoustic and electromagnetic fields from noisy observations," arXiv:0910.2331, 93 p., 2009.
3. Podlipenko, Y., Y. Shestopalov, and V. Prishlyak, "Estimation of solutions of helmholtz problems with uncertain data," *Proc. 2010 International Symposium on Electromagnetic Theory EMTS 2010*, 621–623, Berlin, Germany, Aug. 15–19, 2010.
4. Podlipenko, Y. and Y. Shestopalov, "On the electromagnetic scattering problem for an infinite dielectric cylinder of an arbitrary cross section located in the wedge," *J. Math. Phys.*, Vol. 40, No. 10, 4888–4902, 1999.
5. Podlipenko, Y. and Y. Shestopalov, "Guaranteed estimates of functionals from solutions and data of interior Maxwell problems under uncertainties," *Springer Proceedings in Mathematics & Statistics*, Vol. 52, 2013.
6. Tarantola, A., *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*, Elsevier, 1987.
7. Fedele, F. and R. Muhanna, "Interval solution for inverse problems under uncertainty," *Proc. 5th Int Conf. on Reliable Engineering Computing (REC 2012)*, LITERA, 139–149, 2012.
8. Eidus, D. M., "The principle of limiting amplitude," *Russian Math. Surveys*, Vol. 24, No. 3, 97–167, 1969.
9. Eidus, D. M., "On the principle of limiting absorption," *Mat. Sb.*, Vol. 57, 13–44, 1962.
10. Morgenrother, K. and P. Werner, "On the principles of limiting absorption and limit amplitude for a class of locally perturbed waveguides, Part 1: Time-independent theory," *Math. Meth. Appl. Sci.*, Vol. 10, No. 2, 125–144, 1988.

Calculating the Scattering from Periodic Conducting Surfaces without Using Evanescent Modes, Part I: Existence of a Solution

Dayalan Kasilingam and Christopher Goonan

Department of Electrical & Computer Engineering, University of Massachusetts Dartmouth
285 Old Westport Road, North Dartmouth, MA 02747-2300, USA

Abstract— In this paper, a new technique for solving the problem of scattering from rough surfaces without the need to calculate the evanescent modes is presented. In the new technique, it is shown that there exists a modal sub-space in which the boundary conditions are satisfied using only the incident propagating modes and the scattered propagating modes. This modal sub-space is chosen to be orthogonal to all the evanescent modes on the boundary. Thus by projecting the boundary conditions on to this sub-space, a set of equations which only involve propagating modes is derived. In a companion paper, it is shown how this sub-space can be estimated without calculating the evanescent modes. In this paper, the evanescent modes are calculated directly to show the existence of an appropriate modal sub-space which is orthogonal to all the evanescent modes. A unique solution is shown to exist. It is also shown that the known limitations of the Rayleigh hypothesis do not prevent one from solving for scattering from periodic surfaces, even when the surface roughness is large. A matrix form of a transfer function model, which describes the scattered fields in terms of the incident fields is derived. The transfer function model describes the scattering for all possible combinations of incident waves. The proposed method reduces the computational complexity of the scattering problem significantly by eliminating the need for calculating the evanescent modes. This newly developed technique has the potential to be a powerful computational method for calculating scattering from rough surfaces and gratings.

1. INTRODUCTION

In 1897, Lord Rayleigh published a classic paper in which he used periodic functions to describe the scattering from periodic surfaces [1]. Known as the Rayleigh hypothesis, this method not only used periodic plane wave representations to describe the fields away from the periodic surfaces, but also inside the corrugated region of the surface. However, in the mid-sixties through the early eighties, many researchers challenged Rayleigh's approach of using plane wave representations for the fields inside the corrugated region of the periodic surface [2, 3]. Several of these papers used extensive complex analysis to show that the solutions were unstable when the periodic surface become rough. However in the past decade, other researchers have shown that despite these apparent limitations, the Rayleigh hypothesis gives valid results even for very rough surfaces [4, 5].

In the Rayleigh method, the scattered fields are represented by an infinite spectrum of plane wave modes. However, only a small fraction of these modes are propagating plane waves. Most of the infinite set of modes are evanescent modes which exist primarily around the scattering surface and decay exponentially in the far-field. In most scattering problems, the scattered fields are measured in the far-field. The evanescent modes do not contribute to the far-field scattering since they occur at the boundary and fall off rapidly in the far-field. Since the evanescent modes are also reactive field components, they do not contribute to the power conservation calculations.

In this paper, a new technique where the rough surface scattering problem is solved without the need to include the evanescent modes is developed. In this new technique, it is shown that there exists a modal sub-space in which the boundary conditions are satisfied using only the incident propagating modes and the scattered propagating modes. This modal sub-space is chosen to be orthogonal to all the evanescent modes on the boundary. Thus by projecting the boundary conditions on to this sub-space, a set of equations which only involve propagating modes is derived. In a companion paper, it is shown how this model sub-space can be estimated without calculating the evanescent modes. In this paper, the evanescent modes are calculated directly to show the existence of an appropriate modal sub-space which is orthogonal to all the evanescent modes.

2. THEORY AND RESULTS

In this study, one considers scattering from a perfectly conducting, one-dimensional sinusoidal surface. For the purpose of clarity, the analysis will focus on horizontally polarized plane wave

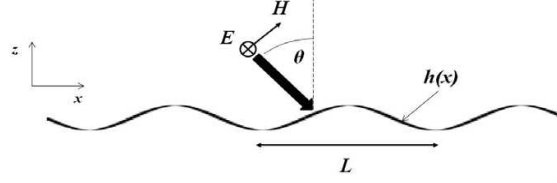


Figure 1: Geometry of the incident wave. The surface is a perfectly conducting sinusoidal surface. The periodicity is L .

incidence. Figure 1 shows the geometry of the incidence of the plane wave on the sinusoidal surface.

The analysis is not limited to a single plane wave incidence. It is assumed that the sinusoidal surface is illuminated by a spectrum of discrete plane waves in the directions of the Floquet modes [6]. The latter constraint allows one to analyze the scattering using a discrete set of plane waves. However, it does not limit the application of this analysis method to a set of fixed directions. The profile of the surface is given by

$$h(x) = h_0 \cos k_L x, \quad (1)$$

where h_0 is the amplitude of the sinusoidal surface and $k_L = 2\pi/L$, where L is the periodicity. As mentioned before, the incidence fields are represented by a spectrum of plane waves in discrete Floquet mode directions and can be written as

$$E_i(x, z) = \hat{y} \sum_{n=-N_0}^{N_0} a_n \exp[-j(\alpha_n x + \beta_n z)], \quad (2)$$

where a_n is the amplitude and α_n and β_n are the x - and z -directed components on the propagation constant of the n th mode. Using standard Floquet mode theory [6], α_n and β_n are given by $\alpha_n = 2\pi n/L$ and $\beta_n = \sqrt{k_0^2 - \alpha_n^2}$, respectively. $k_0 = 2\pi/\lambda_0$, where λ_0 is the electromagnetic wavelength. $2N_0 + 1$ is the total number of propagating incident modes (β_n is real) and depends on the magnitude of the periodicity, L , relative to λ_0 . Similarly, the scattered fields can be written as

$$E_s(x, z) = \hat{y} \sum_{n=-N_0}^{N_0} b_n \exp[-j(\alpha_n x - \beta_n z)] + \hat{y} \sum_{|n| > N_0} c_n \exp[-j\alpha_n x - \gamma_n z], \quad (3)$$

where b_n and c_n represent the amplitude of the scattered propagating modes and scattered evanescent modes, respectively. The evanescent modes are non-propagating modes where $|\alpha_n| > k_0$ and $\gamma_n = \sqrt{\alpha_n^2 - k_0^2}$. The problem is solved by calculating the unknown amplitudes b_n and c_n by applying the boundary conditions on the conducting surface. Since the tangential electric field everywhere on the conducting surface is zero, the Dirichlet boundary condition can be written in terms of the electric fields as

$$E_i(x, h(x)) + E_s(x, h(x)) = 0. \quad (4)$$

Equation (4) can be solved using either the point matching method or the method of moments [7]. In the point matching method, the boundary condition is enforced at M spatial observation points which are uniformly distributed across the surface, resulting in M linear equations. Using matrix representations these M equations can be written as

$$A\hat{a} + B\hat{b} + C\hat{c} = 0, \quad (5)$$

where \hat{a} , \hat{b} and \hat{c} are vectors representing the amplitudes of the incident propagating modes', scattered propagating modes' and scattered evanescent modes' power waves, respectively. The power wave is used in many microwave engineering applications to represent the generic wave amplitude, whose magnitude squared is equal to the power density per unit area in the z -direction. The amplitude of the power wave of a plane wave is obtained from the amplitude of the electric field by scaling the latter using a function of the cosine of the angle of incidence and the wave impedance. This factor is incorporated into the elements of the matrices A , B and C , which are determined by the appropriate mode values at each of the M observation points. In the method of

moments, a similar equation can be generated for the boundary conditions. By combining all the modes, Equation (5) can be written as

$$P\hat{e} = 0, \quad (6)$$

where $\hat{e} = [a_{-N_0} \dots a_{N_0} b_{-N_0} \dots b_{N_0} c_1 \dots]^T$ is the vector representing all the mode amplitudes. The matrix, P , is obtained by concatenating the matrices, A , B and C . In Equation (6), the vector \hat{e} represents the null space of P . Since there are only $2N_0 + 1$ incident modes, the nullity of matrix P should also be $2N_0 + 1$. If the nullity of P is greater than $2N_0 + 1$, then the solution will not be unique and if the nullity of P is less than $2N_0 + 1$, then the problem is under-defined and cannot be solved.

In this study, simulations are performed for scattering from perfectly conducting sinusoidal surfaces with a periodicity of $L = 5.5\lambda_0$ for a variety of different surface heights, h_0 . This value of L yields $2N_0 + 1 = 11$. The singular values of the matrix P are calculated for three different surface heights. Figure 2 shows the sorted singular values for the different surface heights. In all three cases, there were exactly 11 very small singular values ($< 10^{-8}$). This is an important result because it leads to the main result of this paper. Equation (5) can be written as

$$A\hat{a} + B\hat{b} = -C\hat{c} = \hat{c}_0. \quad (7)$$

If as shown in Figure 2, the nullity of P is $2N_0 + 1$, then the rank of the matrix generated by all the possible values of the vector \hat{c}_0 will also be $2N_0 + 1$, i.e., the solution vectors, \hat{c}_0 , will occupy a sub-space of dimension $2N_0 + 1$. However, the maximum dimension of the sub-space occupied by the propagating modes given by the vectors, \hat{a} and \hat{b} , is $2 \times (2N_0 + 1)$. This shows that there is a $2N_0 + 1$ dimensional sub-space which is occupied by the vectors, \hat{a} and \hat{b} , which is orthogonal to the solution vectors, \hat{c}_0 . If U is the basis representing this orthogonal sub-space, then projecting Equation (7) on to this sub-space will yield

$$U^H(A\hat{a} + B\hat{b}) = -U^H\hat{c}_0 = 0. \quad (8)$$

Equation (8) can be re-written as

$$\begin{aligned} U^H A\hat{a} &= -U^H B\hat{b} \\ \hat{b} &= W\hat{a} \end{aligned}, \quad (9)$$

where $W = -(U^H B)^{-1} U^H A$ is the $(2N_0 + 1) \times (2N_0 + 1)$ scattering matrix which relates the power waves of the scattered propagating modes to the power waves of the incident modes. Note the basis, U , has to be orthogonal to all the evanescent modes, if Equation (8) is true and the solutions to Equation (5) are unique. Since the surface is a perfect conductor, the power has to be conserved and thus W is unitary, i.e., $W^H W = I$. Note the vectors \hat{a} and \hat{b} represent power waves whose magnitude-squared yields power density. Reciprocity requires that W also be symmetric, i.e., $W = W^T$.

Equation (9) is the main result of this paper. However, the primary goal of this study is to find the basis, U , which represents the sub-space orthogonal to the evanescent modes. If U can be

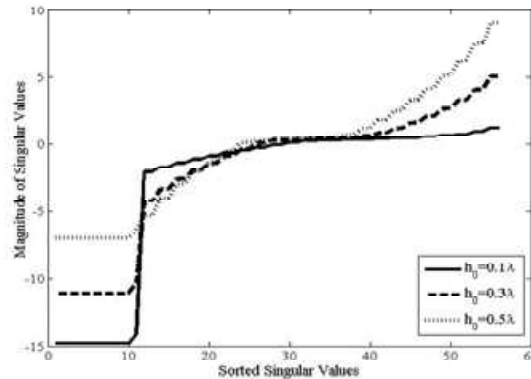


Figure 2: Sorted singular values of matrix \mathbf{P} .

found directly, then one may use Equation (9) to solve the scattering problem without calculating the evanescent modes. This is an important result. In most scattering problems such as in radar and remote sensing, the scattered fields are measured in the far-field. The evanescent modes do not contribute to the far-field scattering, since they decay exponentially with range. Since the evanescent modes are also reactive fields, they do not contribute to any power conservation calculations. Another important benefit in not calculating the evanescent modes is that they fluctuate rapidly with surface height because of their exponential dependence on z . Furthermore, the evanescent modes form an infinite set of modes. Generally in Equation (3), the number of evanescent modes is truncated to make the problem tractable. However, finding the number at which one truncates the series of evanescent modes is always a challenge. On the other hand, for this particular problem with periodic surfaces, the total number of propagating modes is finite and equal to $2 \times (2N_0 + 1)$.

The technique by which one estimates the basis, U , which is orthogonal to all the evanescent modes, without calculating the evanescent modes themselves, is given in the companion paper [8]. In this paper, the goal is to show the existence of this sub-space. The aforementioned arguments for the existence of a $2N_0 + 1$ dimensional sub-space which is orthogonal to the evanescent modes, is reinforced by performing simulations of scattering from a perfectly conducting, sinusoidal surface. However, in these simulations, the evanescent modes' functional form is calculated directly and then used to find the sub-space orthogonal to them. This approach does not produce any special advantages in the scattering calculations, but for buttressing the argument for the existence of a $2N_0 + 1$ dimensional sub-space.

A sinusoidal surface with periodicity of $L = 5.5\lambda_0$, is used in the simulations. The scattered fields are calculated by using the 'full blown' solution which is obtained by solving Equation (5) directly and by using the 'reduced' solution which is calculated using Equation (9). In the latter case, the sub-space U is first estimated by calculating the evanescent modes and finding the sub-space orthogonal to them.

Figure 3 shows the magnitude of the electric field in the region immediately adjacent to the sinusoidal surface for different surface heights, h_0 . These fields are calculated using the 'full blown' solution. Note the 'reduced' solution cannot be used in this case, because it explicitly excludes evanescent modes which are dominant in the near-field.

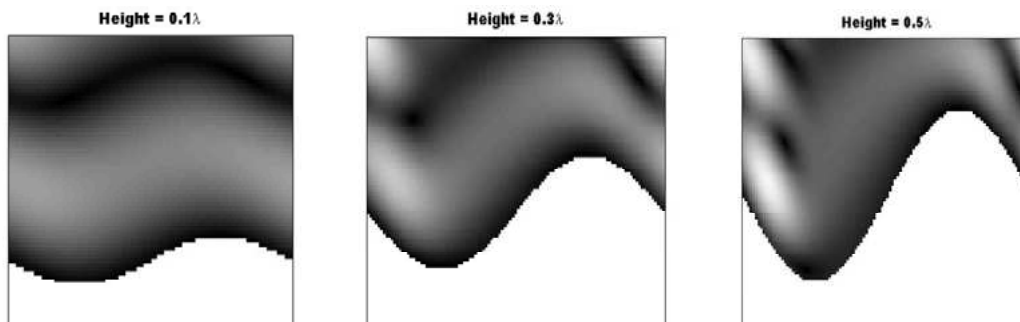


Figure 3: Magnitude of the near-field electric field strength calculated using the 'full blown solution' for different surface heights.

Figure 4 shows the scattering patterns for different surface heights, using the 'full blown' solution and the 'reduced' solution. Clearly the two sets of solutions agree completely for the three surface heights shown in Figure 4. This further reinforces the premise that there exists a sub-space of dimension $2N_0 + 1$ which is orthogonal to the evanescent modes and this allows one to solve the scattering problem using Equation (9).

Finally, the power conservation test is used to validate the 'reduced' solution approach. Figure 5 shows the power as a function of surface height. Shown in each figure is the total scattered power, the power in the forward scattering component and the power in the backscattered component relative to the total incident power. As expected when the surface height is small, the power in forward scattering component is almost equal to the total incident power and the backscattered power is negligibly small. As the surface height increases, the power in the forward scattering component decreases and the power in backscattered component increases. The total scattered power is always equal to the total incident power.

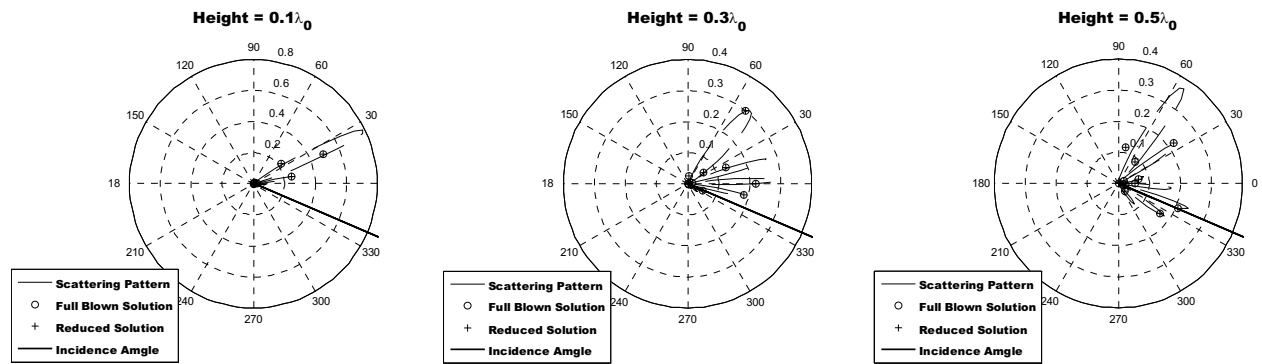


Figure 4: Scattering power patterns calculated using the ‘full blown’ and ‘reduced’ solutions for different surface heights.

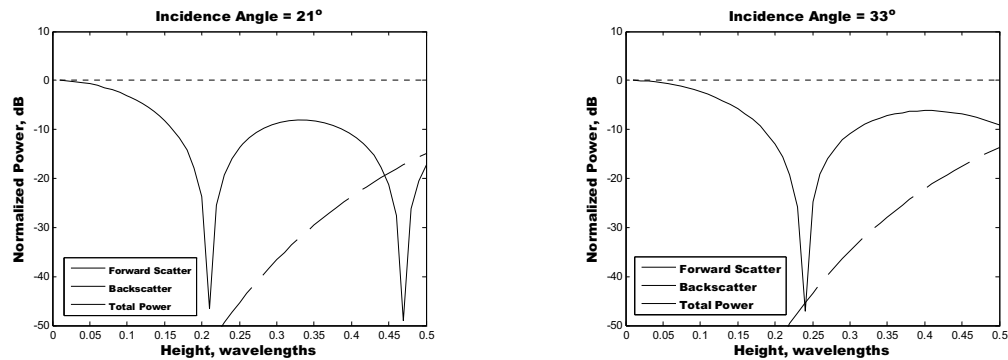


Figure 5: Scattered power relative to the total incident power as a function of surface height calculated using the ‘reduced’ solution.

3. CONCLUSIONS

In this paper, it is shown that for scattering from periodic surfaces, there exists a finite dimensional sub-space which is occupied by the propagating Floquet modes [8] and is orthogonal to all the evanescent modes. The dimension of the sub-space is equal to the total number of scattered (or incident) propagating modes. By projecting the boundary conditions on to this sub-space, a set of equations which involve only the scattered propagating modes and the incident scattered modes is derived. The scattered propagating mode amplitudes are found in terms of the incident mode amplitudes and represented by a unitary scattering matrix which is also symmetric. Simulations show that the ‘reduced’ solution using the sub-space method agrees completely with the ‘full blown’ solution which includes the evanescent modes. A fast technique for estimating the orthogonal sub-space without calculating the evanescent modes is given in a companion paper [9].

REFERENCES

1. Rayleigh, L., “On the dynamical theory of gratings,” *Proc. Royal Soc. A*, Vol. 79, 399–416, 1907.
2. Petit, R. and M. Cadilhac, “Sur la diffraction d’une onde plane par un réseau infiniment conducteur,” *C. R. Acad. Sci. B*, Vol. 262, 468–471, 1966.
3. Millar, R. F., “On Rayleigh assumption in scattering by a periodic surface,” *Proc. Camb. Phil. Soc.*, Vol. 65, 773–791, 1969.
4. Tishchenko, A. V., “Numerical demonstration of the validity of the Rayleigh hypothesis,” *Optic. Expr.*, Vol. 17, 17102–17117, 2009.
5. Kazandjian, L., “Rayleigh methods applied to electromagnetic scattering from grating in general homogeneous media,” *Phys. Rev. E*, Vol. 54, 6802–6815, 1996.
6. Eastham, M. S. P., “The spectral theory of periodic differential equations,” *Texts in Mathematics*, Scottish Academic Press, Edinburgh, 1973.
7. Garg, R., *Analytical and Computational Methods in Electromagnetics*, Artech House, 2008.
8. Chuang, S.-L. and J. A. Kong, “Scattering of waves from periodic surfaces,” *Proc. IEEE*, Vol. 89, 1132–1144, 1981.

9. Kasilingam, D. and C. Goonan, "Calculating the scattering from periodic conducting surfaces without using evanescent modes, part II: Formulation of the solution," *PIERS Proceedings*, Stockholm, Sweden, August 12–15, 2013.

FDTD Simulation of Waveguide with Non-uniform Dielectric Slab

A. P. Smirnov¹, A. N. Semenov¹, and Y. V. Shestopalov²

¹Lomonosov Moscow State University, Moscow, Russia

²Karlstad University, Karlstad, Sweden

Abstract— Scattering in the time domain of electromagnetic waves in the elongated waveguide with non-uniform dielectric slab is considered. Electromagnetic field components are computed and investigation of energy transport in the guide is performed by using Finite Difference Time Domain (FDTD) method for various frequency ranges. Computation for the non-stationary Maxwell equation system is performed by efficient 3D FDTD solver EMWSolver3D created by this paper authors. Simulation is performed for the H10-mode scattering from dielectric slab inclusions. Numerical computations for large-scale problems solution have been implemented on supercomputers of the last generation. The simulation of an empty waveguide without dielectric inclusions has shown that numerical dispersion arising during waves travelling in waveguide causes solution errors. Numerical phase velocity is shown to differ from the analytical phase velocity with the lapse of time that obstructs accurate finding of attenuation and propagation factors. In this respect method similar to Total Field/Scattered field has been proposed to specify waveguide mode with respect to numerical dispersion. The analytical solution of finite-difference equation for the H10-mode has been found for this purpose. Usage of the methods described above has allowed the authors to compute the values of waveguide attenuation and propagation factors for different configurations of dielectric slab.

1. INTRODUCTION

Investigation of electromagnetic waves scattering on dielectric bodies that have complicated geometry or structure is an important problem when composite or artificial materials and media are used as elements of various devices. However, the solution usually cannot be obtained directly, because of composite character of the material and small size of samples, which leads to the necessity of applying methods of mathematical modeling and numerical solution of the corresponding electromagnetic problems [1].

In this work, we are using 3-dimensional FDTD solver EMWSolver3D [2]. Based on Maxwell equations approximation in integral form on Yee lattice, it provides numerical solution in time domain. EMWSolver3D supports multi-core single processors machines and provides hybrid MPI/OpenMP support for IBM BlueGene/P series. Large-scale electromagnetic and optical problems with size of the order of 400 wavelengths in every dimension for a problem with arbitrary complex geometry structure can be solved. The parallel implementation on MSU BlueGene/P supercomputer, based on asynchronous operations, provides good scalability factor for large problems [3]. Solver supports Dirichlet and periodic boundary conditions on any boundary interface [1]. Uniaxial perfect matched layer conditions are also implemented in solver, so wide range of problems in unbounded region can be solved [1, 2, 4].

Pure scattered field method has been proposed to specify waveguide mode with respect to numerical dispersion. The analytical solution of finite-difference equation for the H10-mode have been calculated for this purpose. Using the methods described above has allowed to compute the values of waveguide transmission factors for different configurations of dielectric slab.

This work aims at propagation factors estimation of filters created on the basis of single-layered parallel-plane dielectric diaphragms in waveguides of rectangular cross section. This topic is interesting for the possibility of design of specific frequency selective filters. For instance, such filters can suppress a desired band of frequencies. The proposed method can be used for benchmark problems where electromagnetic parameters of materials obtained analytically or numerically reconstructed from inverse problem [5].

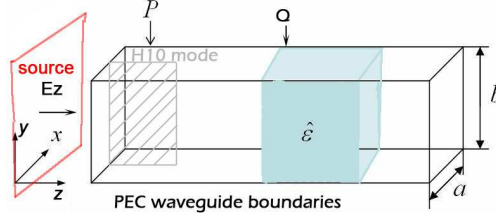


Figure 1: Geometry of the problem.

2. NUMERICAL METHOD

The problem is to find numerical solution of Maxwell's equations that describes scattering from dielectric body in waveguide. The numerical solution is base upon following system of equations:

$$\begin{cases} \text{rot} \vec{E}_t = -\frac{\partial \vec{B}}{\partial t} - \vec{M}_s \\ \text{rot} \vec{H}_t = \frac{\partial \vec{D}}{\partial t} + \vec{J}_s \\ \text{div} \vec{B} = 0 \\ \text{div} \vec{D} = 0 \\ \vec{B} = \mu \mu_0 \vec{H}_t \\ \vec{D} = \varepsilon \varepsilon_0 \vec{E}_t \end{cases} \quad (1)$$

where $\varepsilon = \varepsilon(\vec{r})$ scalar function, and \vec{J}_s and \vec{M}_s are sources. So the solution is considered for isotropic, non-dispersive media.

Lets consider geometry of the problem. Assume that a waveguide $P = \{x : 0 < x < a, 0 < y < b, -\infty < z < \infty\}$ of rectangular cross-section with the perfectly conducting boundary surface ∂P is given in the Cartesian coordinate system (see figure).

A three-dimensional body Q ($Q \subset P$ is a domain) with a constant magnetic permeability μ_0 and variable permittivity $\varepsilon(x)$ is placed in the waveguide. Function $\varepsilon(x)$ is bounded in \bar{Q} , $\varepsilon \in L_\infty(Q)$, and $\varepsilon^{-1} \in L_\infty(Q)$. The boundary ∂Q of domain Q is piecewise smooth. In the provided numerical experiment electric permittivity $\varepsilon(x)$ was set constant.

We assume that the electromagnetic field $\mathbf{E}_t = \mathbf{E}_s + \mathbf{E}_{inc}$, $\mathbf{H}_t = \mathbf{H}_s + \mathbf{H}_{inc}$ in the waveguide is excited by an external field with E_{inc} , H_{inc} the time dependence functions $e^{-i\omega t}$ and where E_s and H_s are scattered field components obtained from FDTD numerical scheme [1].

Transmission coefficient obtained from exact solution of finite-difference problem and E_ω field in frequency domain that is calculated using Fourier transform from time domain solution. So for the waveguide with rectangular cross-section solution for FDTD equations is:

$$E_y^n = E_{y0} \cos(\omega n \Delta t - \tilde{\gamma}_0 z) \sin\left(\frac{\pi x}{a}\right), \quad (2)$$

$$H_x^n = -\frac{\Delta t}{\mu \Delta z} E_{y0} \frac{\sin(\tilde{\gamma}_0 \Delta z / 2)}{\sin(\omega \Delta t / 2)} \cos(\omega n \Delta t - \tilde{\gamma}_0 z) \sin\left(\frac{\pi x}{a}\right), \quad (3)$$

$$H_z^n = \frac{\Delta t}{\mu \Delta z} E_{y0} \frac{\sin(\pi \Delta x / 2a)}{\sin(\omega \Delta t / 2)} \sin(\omega n \Delta t - \tilde{\gamma}_0 z) \cos\left(\frac{\pi x}{a}\right). \quad (4)$$

where n — time step number, Δt — time step size, Δx , Δy , Δz — spatial grid discretization size, $\tilde{\gamma}_0$ — numerical dispersion coefficient, E_{y0} — wave amplitude.

Dispersion relation can be written as:

$$\left[\frac{\sin(\tilde{\gamma}_0 \Delta z / 2)}{\Delta z} \right]^2 = \varepsilon \mu \left[\frac{\sin(\omega \Delta t / 2)}{\Delta t} \right]^2 - \left[\frac{\sin(\pi \Delta x / 2a)}{\Delta x} \right]^2 \quad (5)$$

from which we can obtain $\tilde{\gamma}_0$. Finally, we can calculate transmission \tilde{F} :

$$\text{Re} \tilde{F} = -(\sin(\tilde{\gamma}_0 z) E_{y \sin \omega} - \cos(\tilde{\gamma}_0 z) E_{y \cos \omega}) / \sin(\pi x / a) \cos(2\tilde{\gamma}_0 z), \quad (6)$$

$$\text{Im} \tilde{F} = -(\cos(\tilde{\gamma}_0 z) E_{y \sin \omega} - \sin(\tilde{\gamma}_0 z) E_{y \cos \omega}) / \sin(\pi x / a) \cos(2\tilde{\gamma}_0 z). \quad (7)$$

where $E_{y \sin \omega}$, $E_{y \cos \omega}$ is Fourier transform of E_y field component by $\cos(\omega t)$ and $\sin(\omega t)$ respectively.

3. NUMERICAL RESULTS

Choose the waveguide parameters so that $\pi/a < k_0 < \pi/b$, where k_0 is the free-space wave number, $k_0^2 = \omega^2 \epsilon_0 \mu_0$, and ω is the circular frequency. In this case, only one mode propagates in the waveguide (namely, the principal \mathbf{H}_{10} mode), and all other modes are evanescent (decaying). So let $a = 2b = 2\lambda$. It suits the condition that $\Lambda < \frac{2}{\sqrt{\frac{1}{a^2} + \frac{1}{b^2}}}$, where λ wavelength in cubic waveguide.

The width of dielectric slab is $l = 0.2a$ and permittivity $\epsilon = 1.9$. FDTD parameters are: $\Delta x = \Delta y = \Delta z = \lambda/20$. The result of simulation presented at Fig. 2. According to equation presented [5] $\text{Re}F = 0.904111$ and using proposed technique we evaluated $\text{Re}\tilde{F} = 0.90189$.

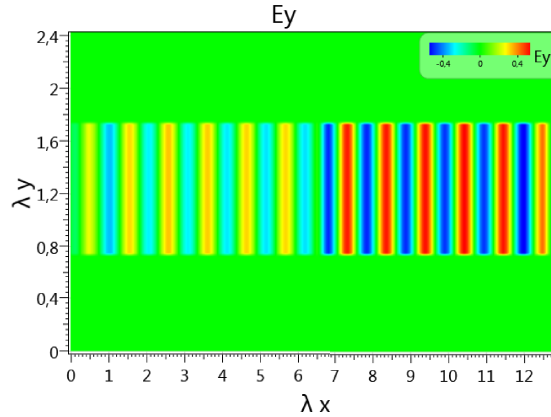


Figure 2: E_y field component 2d slice by x .

For the same problem with parameters: permittivity $\epsilon = 2.9$, $\Delta x = \Delta y = \Delta z = \lambda/40$ the result of simulation is analytical value of transmission $\text{Re}F = 0.631035$, while evaluated $\text{Re}\tilde{F} = 0.675053$.

4. CONCLUSION

Scattering in the time domain of electromagnetic waves in the waveguide of rectangular cross section with single-layered parallel-plane dielectric diaphragm is considered. Electromagnetic field components are computed and investigation of energy transport in the guide is effected by using Finite Difference Time Domain (FDTD) method for different frequency ranges. Computation for the nonstationary Maxwell equation system is performed by efficient 3D FDTD solver EMW-Solver3D created by authors of this paper. Simulation is performed for the \mathbf{H}_{10} -mode scattering from dielectric slab inclusions. Numerical computations for solving large-scale problems have been implemented on IBM BlueGene/P. Numerical phase velocity is shown to differ from the analytical phase velocity with the lapse of time that obstructs accurate finding of attenuation and propagation factors. In this respect pure scattered field method has been used to specify waveguide mode with respect to numerical dispersion. The analytical solution of finite-difference equation for the \mathbf{H}_{10} -mode have been calculated for this purpose. Comparison between numerical and analytical value of propagation coefficient is performed.

REFERENCES

1. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite Difference Time-domain Method*, Norwood, 2000.
2. Smirnov, A. P. and A. N. Semenov, "Full wave Maxwell's equations solver EMWSolver3D," *PIERS Proceedings*, 252–255, Moscow, Russia, Aug. 19–23, 2012.
3. Semenov, A. N., "Parallelnaya realizacia chislennogo resheniya uravneniy Maksvellia FDTD metodom dly bolshih zadach," *Nauchniy Servis v Seti Internet*, 530, Novorossiysk, Russian Federation, Sep. 2011.
4. Gedney, S. D., *IEEE Trans. Antennas Prop.*, Vol. 44, 1630–1639, 1996.
5. Smirnov, Y. G., Y. V. Shestopalov, and E. D. Derevyanchuk, "Permittivity reconstruction of layered dielectrics in a rectangular waveguide from the transmission coefficients at different frequencies," *Inverse Problems Workshop*, Karlstad, Sweden, May 2012.

Remarks on the Mathematical Solution of the Hollow Cavity Eigenvalue Problem

G. Costanza¹ and A. D. Ioannidis²

¹Lund University, Sweden

²Linnæus University, Sweden

Abstract— We discuss the eigenvalue problem for a perfectly conducting hollow cavity under a strict functional analytic point of view. We make use of a variant of the classical spectral theorem for compact selfadjoint operators and we pay extra attention on the null space of the Maxwell operator. We also discuss the corresponding inhomogeneous problem, where currents are present, even when they may depend on the fields.

1. NOTATION AND PRELIMINARIES

The purpose of this paper is to provide an efficient mathematical framework for the classical eigenvalue problem for a hollow, perfectly conducting cavity. Our motivation emerged from the recent research concerning the ESS accelerator project [3, 7]. The study of the cavity problem has been started as early as the 40's [10] and reconsidered many times then, even since our days, see [1, 4, 6, 11]. Nevertheless, there are still some dark points concerning mainly existence issues and the so-called completeness of the modes. Our aim is to give a clear and concise picture of the relevant mathematical problem and suggest the appropriate tools for its solution, in the spirit of [2].

The notation we use in this paper is as follows. Let $(X, \langle \cdot, \cdot \rangle)$ be an infinite dimensional separable Hilbert space. For a set $U \subset X$, we denote by \bar{U} the closure, by U^\perp the orthogonal complement and by $[U]$ the linear span of U ; the closed linear span is then $\overline{[U]}$. $B(X)$ stands for the Banach algebra of bounded operators in X and $K(X)$ for the ideal of compact operators. Given a linear operator $A : X \supset \mathfrak{D}(A) \rightarrow X$, we denote by $\mathfrak{R}(A)$ the range and by $\ker A$ the null space (kernel) of A . The graph norm on $\mathfrak{D}(A)$ is defined as

$$\|x\|_A := \sqrt{\|x\|^2 + \|Ax\|^2}.$$

When equipped with the graph norm, $\mathfrak{D}(A)$ will be denoted as $[\mathfrak{D}(A)]$ (not to be confused with the linear span notation). A^* stands for the adjoint operator. A is called selfadjoint if $A^* = A$, skew-adjoint if $A^* = -A$.

The resolvent set $\rho(A)$ consists of all $\lambda \in \mathbb{C}$ for which $R(\lambda; A) := (\lambda I - A)^{-1} \in B(X)$ (resolvent operator). If $\rho(A) \neq \emptyset$ then A is closed. The spectrum of A is defined as $\sigma(A) := \mathbb{C} \setminus \rho(A)$. $\lambda \in \sigma(A)$ is called an eigenvalue if $\lambda I - A$ is not injective and non-zero vectors of $\ker(\lambda I - A)$ are called the corresponding eigenvectors. The set of eigenvalues is denoted by $\sigma_p(A)$ (point spectrum).

Proposition 1. *The following are equivalent:*

- $R(\lambda; A) \in K(X)$ for some $\lambda \in \rho(A)$.
- $R(\lambda; A) \in K(X)$ for all $\lambda \in \rho(A)$.
- $\rho(A) \neq \emptyset$ and $[\mathfrak{D}(A)] \hookrightarrow X$ with a compact injection.

Definition 1. If A satisfies one of the equivalent conditions of Prop. 1, then it is called a discrete operator.

The following theorem provides the main tool for our exposition.

Proposition 2 (Spectral Theorem). *Let A be a discrete selfadjoint operator. Then $\sigma(A) = \sigma_p(A)$ and $\sigma_p(A)$ is a finite or unbounded countable set without accumulation point. The corresponding eigenspaces are finite dimensional and mutually orthogonal. Eigenvalues of A can be set as an increasing sequence (λ_n) , diverging at infinity if countable. Each non-zero eigenvalue is counted according to its multiplicity and the sequence (e_n) of the corresponding eigenvectors can be chosen as an orthonormal sequence. Moreover,*

- (e_n) is an orthonormal basis for $\mathfrak{R}(A)$ and $X = \ker A \oplus \overline{[(e_n)]}$.
- $Ax = \sum_n \lambda_n \langle x, e_n \rangle e_n$, $x \in \mathfrak{D}(A)$.

2. THE MAXWELL SYSTEM

As it is well known, every electromagnetic phenomenon is specified by four vector quantities: the electric field \mathbf{E} , the magnetic field \mathbf{H} , the electric flux density \mathbf{D} and the magnetic flux density \mathbf{B} , in the presence of electric and magnetic currents $\mathbf{J}_e, \mathbf{J}_m$, respectively. These quantities are considered as time-dependent vector fields on a domain $\Omega \subset \mathbb{R}^3$, so they are functions of the spatial variable $\mathbf{r} \in \Omega$ and the time variable $t \in \mathbb{R}$. All these fields are connected via the Maxwell system

$$\frac{\partial \mathbf{D}}{\partial t} = \text{curl} \mathbf{H} - \mathbf{J}_e, \quad \frac{\partial \mathbf{B}}{\partial t} = -\text{curl} \mathbf{E} + \mathbf{J}_m. \quad (1)$$

We have allowed existence of magnetic currents here because apertures in a cavity can be modeled this way [9]. The above are supplemented with the two Gauss laws

$$\text{div} \mathbf{D} = \rho_e, \quad \text{div} \mathbf{B} = \rho_m. \quad (2)$$

where ρ_e, ρ_m are the densities of the electric and magnetic charge, respectively. Currents and charges are not independent and obey equation of continuity

$$\frac{\partial \rho_e}{\partial t} + \text{div} \mathbf{J}_e = 0, \quad \frac{\partial \rho_m}{\partial t} + \text{div} \mathbf{J}_m = 0. \quad (3)$$

If one accepts (3) as part of the modeling, (2) become redundant and can be absorbed in the initial conditions.

We now assume that the domain Ω is a hollow cavity, *i.e.*, a vacuous bounded domain:

$$\mathbf{D} = \varepsilon_0 \mathbf{E}, \quad \mathbf{B} = \mu_0 \mathbf{H}. \quad (4)$$

Without loss of generality, we assume $\varepsilon_0 = \mu_0 = 1$. We further assume that the boundary Γ of Ω is Lipschitz and therefore an exterior normal $\hat{\mathbf{n}}$ is almost everywhere defined on it and the perfect electric conductor boundary condition applies

$$\hat{\mathbf{n}} \times \mathbf{E} = \mathbf{0}, \quad \text{on } \Gamma. \quad (5)$$

The above implies that $\hat{\mathbf{n}} \cdot \mathbf{H} = 0$ on Γ . In the six vector notation, (1) are read as follows:

$$\frac{\partial}{\partial t} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} = \begin{bmatrix} 0 & \text{curl} \\ -\text{curl} & 0 \end{bmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{H} \end{pmatrix} + \begin{pmatrix} -\mathbf{J}_e \\ \mathbf{J}_m \end{pmatrix}. \quad (6)$$

To make things more precise, let us denote by $\mathbf{e} := (\mathbf{E}, \mathbf{H})^T$ the electromagnetic (EM) field, by $\mathbf{j} := (-\mathbf{J}_e, \mathbf{J}_m)^T$ the EM current and by

$$\mathcal{M} := \begin{bmatrix} 0 & \text{curl} \\ -\text{curl} & 0 \end{bmatrix} \quad (7)$$

the formal Maxwell operator. Then (6) is written

$$\frac{\partial \mathbf{e}}{\partial t} = \mathcal{M} \mathbf{e} + \mathbf{j}. \quad (8)$$

We now assume $\mathbf{j} = \mathbf{0}$ (homogeneous problem) and apply a separation-of-variables technique, that is, we ask for a solution of (8) of the form $\mathbf{e}(\mathbf{r}, t) := \mathbf{e}(\mathbf{r})T(t)$ and thus

$$\mathcal{M} \mathbf{e} = \frac{T'(t)}{T(t)} \mathbf{e}.$$

Since the left hand side depends only on \mathbf{r} , the ratio $T'(t)/T(t)$ has to be a constant, say λ , and we conclude to the formal eigenvalue problem

$$\mathcal{M} \mathbf{e} = \lambda \mathbf{e}, \quad (9)$$

with $\mathbf{e} := (\mathbf{E}, \mathbf{H})^T$. Note that although we use the same notation, \mathbf{E}, \mathbf{H} now do not depend on time. We also have apparently $T(t) = e^{\lambda t}$.

3. REALIZATION OF THE EIGENVALUE PROBLEM

The exposition and notation in this section follows [2, 8]. Due to energy considerations, the fields \mathbf{E} , \mathbf{H} are taken to be square integrable, *i.e.*, they are vectors of the Hilbert space $L^2(\Omega; \mathbb{C}^3)$ with inner product

$$\langle \mathbf{U}, \mathbf{V} \rangle_0 := \int_{\Omega} \mathbf{U}(\mathbf{r}) \cdot \mathbf{V}(\mathbf{r}) d\mathbf{r}.$$

The *curl* operator is naturally realized in its weak sense in $L^2(\Omega; \mathbb{C}^3)$. More precisely, let $\mathbf{U} \in L^2(\Omega; \mathbb{C}^3)$. We say that \mathbf{V} is the (weak) rotation of \mathbf{U} , and we write $\mathbf{V} = \text{curl} \mathbf{U}$, if

$$\langle \mathbf{V}, \phi \rangle_0 = \langle \mathbf{U}, \text{curl} \phi \rangle_0$$

for every test function $\phi \in C_0^\infty(\Omega; \mathbb{C}^3)$. The maximal domain of definition of *curl* in \mathcal{X} is then the Sobolev space $H(\text{curl}; \Omega)$ and is a densely defined closed operator. Moreover, *curl* can be realized as a maximal selfadjoint operator in the subspace $H_0(\text{curl}; \Omega)$, which contains exactly the fields that satisfy (5) in a weak sense. The null spaces of these operators are denoted by $H(\text{curl}0; \Omega)$, $H_0(\text{curl}; \Omega)$ respectively. Analogous definitions apply for the weak divergence operator *div*, see the aforementioned references for details.

Consequently, the EM field \mathbf{e} is a vector of the product Hilbert space $\mathcal{X} := L^2(\Omega; \mathbb{C}^3) \times L^2(\Omega; \mathbb{C}^3)$ with inner product, for $\mathbf{u} := (\mathbf{U}_1, \mathbf{U}_2)^T$, $\mathbf{v} := (\mathbf{V}_1, \mathbf{V}_2)^T$,

$$\langle \mathbf{u}, \mathbf{v} \rangle := \langle \mathbf{U}_1, \mathbf{V}_1 \rangle_0 + \langle \mathbf{U}_2, \mathbf{V}_2 \rangle_0.$$

The Maxwell operator is realized in a weak sense in \mathcal{X} as follows: $\mathbf{e} \in \mathcal{X}$ is in the domain $\mathfrak{D}(\mathcal{M})$ of \mathcal{M} if there exists a (unique) vector $\mathbf{u} \in \mathcal{X}$ such that

$$\langle \mathbf{u}, (\phi_1, \phi_2)^T \rangle = \langle \mathbf{e}, (-\text{curl} \phi_2, \text{curl} \phi_1)^T \rangle, \quad (10)$$

for every choice of test functions $\phi_1, \phi_2 \in C_0^\infty(\Omega; \mathbb{C}^3)$. In this case, we set $\mathbf{u} := \mathcal{M}\mathbf{e}$. After this, problem (3) can be realized as an eigenvalue problem for such defined operator \mathcal{M} .

Proposition 3. $\mathfrak{D}(\mathcal{M}) = H_0(\text{curl}; \Omega) \times H(\text{curl}; \Omega)$ and \mathcal{M} is a densely defined, closed linear operator, represented by the operator matrix (7). Moreover, \mathcal{M} is skew-adjoint, *i.e.*, $\mathcal{M}^* = -\mathcal{M}$. Consequently, the spectrum of \mathcal{M} is purely imaginary.

That is to say, (3) has imaginary eigenvalues, if any. Following the usual practice, we let $\lambda := -i\omega$, $\omega \in \mathbb{R}$, and the problem is rewritten as

$$\mathcal{Q}\mathbf{e} = \omega\mathbf{e}, \quad (\star)$$

where $\mathcal{Q} := i\mathcal{M}$ is the selfadjoint Maxwell operator. So we are mainly interested in properties of operator \mathcal{Q} ; in view of the above proposition, $\mathfrak{D}(\mathcal{Q}) = \mathfrak{D}(\mathcal{M})$ and \mathcal{Q} is a selfadjoint operator with real spectrum.

Definition 2. Let (ω, \mathbf{e}) be a non-trivial solution of (\star) , *i.e.*, $\omega \in \mathbb{R}$ is an eigenvalue of \mathcal{Q} with corresponding eigenvector \mathbf{e} . ω is called an eigenfrequency of the cavity and \mathbf{e} the corresponding mode.

Proposition 4. The null space of \mathcal{Q} is $\ker \mathcal{Q} = H_0(\text{curl}0; \Omega) \times H(\text{curl}0; \Omega)$. For the range $\mathfrak{R}(\mathcal{Q})$ of \mathcal{Q} , we have $\mathfrak{R}(\mathcal{Q}) \subset H(\text{div}0; \Omega) \times H_0(\text{div}0; \Omega) := \mathcal{H}$.

We now consider the restriction $\mathcal{Q}_{\mathcal{H}}$ of \mathcal{Q} on \mathcal{H} , defined by $\mathfrak{D}(\mathcal{Q}_{\mathcal{H}}) = \mathfrak{D}(\mathcal{Q}) \cap \mathcal{H}$ and $\mathcal{Q}_{\mathcal{H}}\mathbf{e} = \mathcal{Q}\mathbf{e}$. Incidentally, $\mathcal{Q}_{\mathcal{H}}$ coincides with the part of \mathcal{Q} on \mathcal{H} .

Proposition 5. $\mathcal{Q}_{\mathcal{H}}$ is selfadjoint and $[\mathfrak{D}(\mathcal{Q}_{\mathcal{H}})]$ is compactly injected into \mathcal{H} . Consequently, $\mathcal{Q}_{\mathcal{H}}$ is discrete and its spectrum is an unbounded sequence of real eigenvalues with no accumulation point. The corresponding eigenspaces are finite dimensional and mutually orthogonal.

4. PROPERTIES OF EIGENFREQUENCIES AND MODES

Proposition 6. *Let $\omega \neq 0$ be an eigenvalue of $\mathcal{Q}_{\mathcal{H}}$ with corresponding eigenvector $(\mathbf{E}, \mathbf{H})^{\mathbf{T}}$. Then a) \mathbf{E}, \mathbf{H} satisfy the system*

$$\begin{cases} -\operatorname{curl}\mathbf{H} = i\omega\mathbf{E} \\ \operatorname{curl}\mathbf{E} = i\omega\mathbf{H} \end{cases} \quad (11)$$

b) $\|\mathbf{E}\|_0 = \|\mathbf{H}\|_0$.

c) $-\omega$ is an eigenvalue as well, with corresponding eigenvector $(\mathbf{E}, -\mathbf{H})^{\mathbf{T}}$.

That is, the eigenvalues of $\mathcal{Q}_{\mathcal{H}}$ can be ordered as a bilateral sequence $(\omega_n)_{n \in \mathbb{Z}}$, with $\omega_n > 0$ for $n > 0$, $\omega_n < 0$ for $n < 0$, $\omega_0 = 0$ and $\omega_{-n} = -\omega_n$. For $n \neq 0$, we count each eigenvalue ω_n as many times as its multiplicity, so we can assume that to each ω_n there corresponds exactly one normalized eigenvector $\mathbf{e}_n^p := (\mathbf{E}_n^p, \mathbf{H}_n^p)^{\mathbf{T}}$. The zero eigenvalue is counted once and we will discuss about it later. Namely, the eigenvalues are ordered as follows:

$$-\infty \leftarrow \dots \leq \omega_{-n} \leq \dots \leq \omega_{-1} < \omega_0 = 0 < \omega_1 \leq \dots \leq \omega_n \leq \dots \rightarrow \infty.$$

The sequence of eigenvectors (\mathbf{e}_n) is assumed to be orthonormal.

Proposition 7. *Let $n, m \in \mathbb{N}^*$, $\omega_n \neq \omega_m$. Then $\langle \mathbf{E}_n^p, \mathbf{E}_m^p \rangle_0 = \langle \mathbf{H}_n^p, \mathbf{H}_m^p \rangle_0 = 0$, i.e., both $(\mathbf{E}_n^p)_{n \in \mathbb{N}^*}$, $(\mathbf{H}_n^p)_{n \in \mathbb{N}^*}$ define orthogonal sequences in $L^2(\Omega; \mathbb{C}^3)$.*

Proposition 8. *$(\mathbf{e}_n^p)_{n \in \mathbb{Z}^*}$ is an orthonormal basis for $\overline{\mathfrak{R}(\mathcal{Q}_{\mathcal{H}})}$ and we have the decomposition*

$$\mathcal{H} = \ker \mathcal{Q}_{\mathcal{H}} \oplus \overline{[\dots, \mathbf{e}_{-n}^p, \dots, \mathbf{e}_{-1}^p, \mathbf{e}_1^p, \dots, \mathbf{e}_n^p, \dots]}$$

The closed subspace $\ker \mathcal{Q}_{\mathcal{H}}$ is finite dimensional. Moreover, for $\mathbf{e} \in \mathfrak{D}(\mathcal{Q}_{\mathcal{H}})$,

$$\mathcal{Q}_{\mathcal{H}}\mathbf{e} = \sum_{n \in \mathbb{Z}^*} \omega_n \langle \mathbf{e}, \mathbf{e}_n^p \rangle \mathbf{e}_n^p.$$

For a detailed description of $\ker \mathcal{Q}_{\mathcal{H}}$ we refer to [2, 8]. Let N_0 be the dimension of $\ker \mathcal{Q}_{\mathcal{H}}$ (a number depending on the geometry of Ω) and consider an orthonormal basis $\{\mathbf{e}_1^0, \dots, \mathbf{e}_{N_0}^0\}$ for $\ker \mathcal{Q}_{\mathcal{H}}$. Note that $\ker \mathcal{Q}_{\mathcal{H}}$ describes the source-free, static electromagnetism on Ω .

Proposition 9. *$\{\mathbf{e}_1^0, \dots, \mathbf{e}_{N_0}^0\} \cup (\mathbf{e}_n^p)_{n \in \mathbb{Z}^*}$ is an orthonormal basis for \mathcal{H} . $\{\mathbf{e}_1^0, \dots, \mathbf{e}_{N_0}^0\}$ can be completed to an orthonormal basis for $\ker \mathcal{Q}$, that is, there exist an orthonormal sequence $(\mathbf{e}_n^s)_{n \in \mathbb{A}_s} \subset \mathcal{X}$ such that $\{\mathbf{e}_1^0, \dots, \mathbf{e}_{N_0}^0\} \cup (\mathbf{e}_n^s)_{n \in \mathbb{A}_s}$ is an orthonormal basis for $\ker \mathcal{Q}$. Moreover, $\{\mathbf{e}_1^0, \dots, \mathbf{e}_{N_0}^0\} \cup (\mathbf{e}_n^p)_{n \in \mathbb{Z}^*} \cup (\mathbf{e}_n^s)_{n \in \mathbb{A}_s}$ is an orthonormal basis for \mathcal{X} .*

Note that \mathbb{A}_s is an infinite countable set, serving as the index set for (\mathbf{e}_n^s) . This analysis suggests the following classification for the cavity modes, see also [4]:

- Primary modes $(\mathbf{e}_n^p)_{n \in \mathbb{Z}^*}$ (solenoidal, non-irrotational).
- Static modes $\{\mathbf{e}_1^0, \dots, \mathbf{e}_{N_0}^0\}$ (solenoidal, irrotational).
- Secondary modes $(\mathbf{e}_n^s)_{n \in \mathbb{A}_s}$ (non-solenoidal, irrotational).

We also have that an arbitrary field $\mathbf{e} \in \mathcal{X}$ can be represented as

$$\mathbf{e} = \sum_{i=s,0,p} \sum_{n \in \mathbb{A}_i} \langle \mathbf{e}, \mathbf{e}_n^i \rangle \mathbf{e}_n^i, \quad (12)$$

where $\mathbb{A}_0 := \{1, 2, \dots, N_0\}$, $\mathbb{A}_p := \mathbb{Z}^*$. If, in addition, $\mathbf{e} \in \mathfrak{D}(\mathcal{Q})$, then

$$\mathcal{Q}\mathbf{e} = \sum_{n \in \mathbb{A}_p} \omega_n \langle \mathbf{e}, \mathbf{e}_n^p \rangle \mathbf{e}_n^p. \quad (13)$$

5. THE INHOMOGENEOUS PROBLEM

We now allow the presence of EM currents. In the frequency domain, this is modeled with the equation

$$\mathcal{Q}\mathbf{e} = \omega\mathbf{e} + \mathbf{j}, \quad (14)$$

where $\mathbf{j} = \mathbf{j}(\omega, \mathbf{r}) \in \mathcal{X}$. In the view of representations (4), (13), (14) reads

$$\sum_{n \in \mathbb{A}_p} \omega_n \langle \mathbf{e}, \mathbf{e}_n^p \rangle \mathbf{e}_n^p = \omega \sum_{i=s,0,p} \sum_{n \in \mathbb{A}_i} \langle \mathbf{e}, \mathbf{e}_n^i \rangle \mathbf{e}_n^i + \sum_{i=s,0,p} \sum_{n \in \mathbb{A}_i} \langle \mathbf{j}, \mathbf{e}_n^i \rangle \mathbf{e}_n^i,$$

which lead to the equations

$$(\omega_n - \omega) \langle \mathbf{e}, \mathbf{e}_n^p \rangle = \langle \mathbf{j}, \mathbf{e}_n^p \rangle, \quad n \in \mathbb{A}_p, \quad (15)$$

$$\omega \langle \mathbf{e}, \mathbf{e}_n^i \rangle = -\langle \mathbf{j}, \mathbf{e}_n^i \rangle, \quad i = s, 0, \quad n \in \mathbb{A}_i. \quad (16)$$

Equations (15), (16) lead to the following result, a variant of the Fredholm Alternative:

Proposition 10. *a) Let $\omega \neq \omega_n$, $n \in \mathbb{Z}$. Then (14) has a unique solution given by*

$$\mathbf{e} = -\frac{1}{\omega} \sum_{i=s,0} \sum_{n \in \mathbb{A}_i} \langle \mathbf{j}, \mathbf{e}_n^i \rangle \mathbf{e}_n^i + \sum_{n \in \mathbb{A}_p} \frac{1}{\omega_n - \omega} \langle \mathbf{j}, \mathbf{e}_n^p \rangle \mathbf{e}_n^p. \quad (17)$$

b) Let $\omega = \omega_m$ for some $m \in \mathbb{Z}^$. Then (14) has a solution if and only if \mathbf{j} is orthogonal to eigenspace $\ker(\omega_m I - \mathcal{Q}_{\mathcal{H}})$. In this case, a solution of (14) is of the form*

$$\mathbf{e} = -\frac{1}{\omega_m} \sum_{i=s,0} \sum_{n \in \mathbb{A}_i} \langle \mathbf{j}, \mathbf{e}_n^i \rangle \mathbf{e}_n^i + \sum_{\substack{n \in \mathbb{A}_p \\ \omega_n \neq \omega_m}} \frac{1}{\omega_n - \omega_m} \langle \mathbf{j}, \mathbf{e}_n^p \rangle \mathbf{e}_n^p + \mathbf{u} \quad (18)$$

for some $\mathbf{u} \in \ker(\omega_m I - \mathcal{Q}_{\mathcal{H}})$.

c) Let $\omega = 0$. Then (14) has a solution if and only if \mathbf{j} is orthogonal to the kernel $\ker \mathcal{Q}$. In this case, a solution of (14) is of the form

$$\mathbf{e} = \mathbf{u} + \sum_{n \in \mathbb{A}_p} \frac{1}{\omega_n} \langle \mathbf{j}, \mathbf{e}_n^p \rangle \mathbf{e}_n^p \quad (19)$$

for some $\mathbf{u} \in \ker \mathcal{Q}$.

REFERENCES

1. Aksoy, S. and O. A. Tretyakov, "The evolution equations in study of the cavity oscillations excited by a digital signal," *IEEE Trans. Antennas Prop.*, Vol. 52, No. 1, 2004.
2. Dautray, R. and J. L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology, Vol. 3: Spectral Theory*, Springer, Berlin, 1990.
3. Devanz, G. and J. Plouin, "Conceptual design of the beta = 0.86 cavities for the superconducting Linac of the ESS," *Proceedings of SRF*, Chicago, Illinois, USA, July 2011.
4. Geyi, W., "Time-domain theory of metal cavity resonator," *Progress In Electromagnetics Research*, Vol. 78, 219–253, 2008.
5. Helmborg, G., *Introduction to Spectral Theory in Hilbert Spaces*, North Holland, Amsterdam, 1969.
6. Kurokawa, K., "The expansions of electromagnetic fields in cavities," *IRE Trans. Microwave Theory Techniques*, Vol. 6, No. 2, 178–187, 1957.
7. Lindroos, M., S. Molloy, D. McGinnis, C. Darve, and H. Danared, "The ESS Linac design," *Proceedings of LINAC*, Tel-Aviv, Israel, December 2012.
8. Monk, P., *Finite Elements Method for Maxwell's Equations*, Clarendon Press, Oxford, 2003.
9. Omar, A., *Electromagnetic Scattering and Material Characterization*, Artech House, Boston, 2011.
10. Slater, J. C., "Microwave electronics," *Review Modern Physics*, Vol. 18, 441–512, 1946.
11. Teichmann, T. and E. P. Winger, "Electromagnetic field expansions in loss free cavities excited through holes," *J. Applied Physics*, Vol. 24, 262–267, 1953.

The Resolution of Identity: A Unifying Concept in Field Theory

A. R. Baghai-Wadji

University of Cape Town, South Africa

Abstract— The search for a unified regularization scheme of singularities arising in computational electromagnetics, along with the simultaneous quest for a systematic design of optimized analysis- and synthesis functions, have led to the insight that identity matrices and unity operators in finite- and infinite dimensional Hilbert space play an all-important fundamental role. The additive- and multiplicative factorization of the identity matrix \mathbf{I} , and the decomposition of unity operator $\hat{\mathbf{I}}$ in terms of exterior products in Hilbert space have been identified as a single unifying scheme in theoretical and computational electromagnetics. The latter is the focus in the present paper. Several related ideas promoted by the author in his recent publications have not been included in the discussion, e.g., eigenfunctions of hermitian operators as analysis- and synthesis functions and the construction of Green's functions-based wavelets and orthonormal bases. Interested reader is referred to the references and previous works of the author cited therein.

1. INTRODUCTION

Conventional Fourier-type transforms and their associated inverse transforms, less conventional scaling functions, wavelets, and multiresolution analysis [1], and little known concepts of frames and dual frames [2, 3], have been briefly discussed in this paper from a unifying perspective. The unity operator and its building blocks, exterior products of ket- and bra-vectors, allow clarifying Fourier analysis- and synthesis steps. Viewed this way, many seemingly disparate techniques can be considered as one and the same tool to just partition Hilbert space appropriately and meaningfully, by matching the techniques to the conditions and peculiarity of specific problems at hand. The message is that analysis- and synthesis functions (certain set of functions and their dual) can be designed, tailored and customized to boundary value problems under consideration, thereby, rendering the computations more accurate and at the same time more efficient.

Space limitation does now allow any of the significant details to be touched upon [4, 5]. Nonetheless, in order to establish some connection between the material presented in this paper and other related topics certain foundational ideas are briefly listed in the following subsection.

1.1. Traditional Concepts in Mathematical Physics Meet Modern Multiresolution Analysis

Notions of generalized integral (series) transforms and inverse integral (series) transforms of field vectors, as well as co-variant and contra-variant field vector- and tensor components can be established and incorporated into the present framework in a unified manner. The author is convinced that the proposed method paves the way and invites powerful concepts and ideas from signal processing, theoretical physics and computer science into the arena of computational electromagnetics. The notion of spinors in three and higher dimensions can be embraced in a natural way. The proposed methodology, inspired by physics, has its genesis in the modern functional analysis, in the form of multiresolution analysis. The proposed ideas are expected to find applications where classical electromagnetics, mesoscopic phenomena and quantum physics play a role, and thus physical phenomena in various scales must simultaneously be accounted for. More specifically, operators in classical electrodynamics and the corresponding dual operators can be considered, and factorized in their irreducible building blocks, utilizing factorization of differential unity operators. The Maxwell's equations in general bi-anisotropic and inhomogeneous media can be diagonalized in Cartesian coordinates. Maxwell equations in general curvilinear co-ordinates can be transformed to local Cartesian coordinate systems. Problem-specific basis- and weighting (analysis- and synthesis) functions can be constructed.

2. FOURIER- AND INVERSE FOURIER TRANSFORMS

2.1. Exponentially-harmonic Orthonormal Basis

Consider the infinite sequence of T -periodic functions $\{1/\sqrt{T}e^{jn\frac{2\pi}{T}t}|n \in \mathbb{Z}\}$ on the interval $[t_1, t_1+T]$ with $t_1 \in \mathbb{R}$ and $T > 0$. Introduce the bra- and ket-vectors $\langle 1/\sqrt{T}e^{jm\frac{2\pi}{T}t}|$ and $|1/\sqrt{T}e^{jn\frac{2\pi}{T}t}\rangle$. The inner product $\langle 1/\sqrt{T}e^{jm\frac{2\pi}{T}t}|1/\sqrt{T}e^{jn\frac{2\pi}{T}t}\rangle = \delta_{mn}$ states that the mentioned infinite set of functions constitute an orthonormal basis (ONB) on the interval $[t_1, t_1+T]$. Using intervals $[t_1, t_2]$, $[-\pi, \pi]$,

$[-\pi/2, \pi/2]$, $[-1, 1]$, or $[-1/2, 1/2]$ for $[t_1, t_1 + T]$ several familiar exponentially-harmonic ONB can be recognized.

Fourier synthesis. In the course of further discussion the interval $[-\pi, \pi]$ will be considered. Any function $f(t)$ satisfying the Dirichlet conditions¹ on the interval $[-\pi, \pi]$ can be written as the superposition of the functions $1/\sqrt{2\pi}e^{jnt}$ with $n \in \mathbb{Z}$.

$$|f(t)\rangle = \sum_{n=-\infty}^{\infty} \left| 1/\sqrt{2\pi}e^{jnt} \right\rangle f_n \quad (1)$$

Fourier analysis. Projecting $f(t)$ onto $|1/\sqrt{2\pi}e^{jmt}\rangle$, with an arbitrary $m \in \mathbb{Z}$, results in

$$\left\langle 1/\sqrt{2\pi}e^{jmt} | f(t) \right\rangle = \sum_{n=-\infty}^{\infty} f_n \delta_{mn} = f_m \implies f_n = \left\langle 1/\sqrt{2\pi}e^{jnt} | f(t) \right\rangle \quad (2)$$

Resolution of identity. Substituting for f_n from (2) into (1) gives:

$$|f(t)\rangle = \left\{ \sum_{n=-\infty}^{\infty} \left| 1/\sqrt{2\pi}e^{jnt} \right\rangle \left\langle 1/\sqrt{2\pi}e^{jnt} \right| \right\} |f(t)\rangle \quad (3)$$

where used was made of the convention $\langle \cdot | \cdot \rangle = \langle \cdot | | \cdot \rangle$. Furthermore, for greater clarity, and to more easily recognize the “identity operator” at the RHS curly brackets have been introduced.

Denote the identity operator by $\hat{\mathbb{I}}$ and write $\hat{\mathbb{I}}f(t)$ for $f(t)$ at LHS in (3). Since $f(t)$ is arbitrary, the following expression for the resolution of identity can be established (completeness condition):

$$\hat{\mathbb{I}} = \sum_{n=-\infty}^{\infty} \left| 1/\sqrt{2\pi}e^{jnt} \right\rangle \left\langle 1/\sqrt{2\pi}e^{jnt} \right| \quad (4)$$

This is an expression for the completeness of the involved ONB in terms of exterior products $|\cdot\rangle\langle\cdot|$.

2.2. Sinusoidally-harmonic Orthonormal Basis

In virtue of the fact that $\cos(nt)$ and $\sin(nt)$ can be expressed in terms of e^{jnt} and e^{-jnt} it is intuitive that the following set of 2π -periodic functions constitute a (complete) ONB on the interval $[-\pi, \pi]$: $1/\sqrt{2\pi}$, $\{1/\sqrt{\pi} \cos(nt) | n \in \mathbb{N}\}$, $\{1/\sqrt{\pi} \sin(nt) | n \in \mathbb{N}\}$.

Resolution of identity. The above ONB offers an alternative resolution of the identity operator $\hat{\mathbb{I}}$:

$$\hat{\mathbb{I}} = \left| \frac{1}{\sqrt{2\pi}} \right\rangle \left\langle \frac{1}{\sqrt{2\pi}} \right| + \sum_{n=1}^{\infty} \left| \frac{1}{\sqrt{\pi}} \cos(nt) \right\rangle \left\langle \frac{1}{\sqrt{\pi}} \cos(nt) \right| + \sum_{n=1}^{\infty} \left| \frac{1}{\sqrt{\pi}} \sin(nt) \right\rangle \left\langle \frac{1}{\sqrt{\pi}} \sin(nt) \right| \quad (5)$$

2.2.1. Fourier Analysis and Synthesis Steps

Apply the operator $\hat{\mathbb{I}}$ in (5) to any function $f(t)$ on the interval $[-\pi, \pi]$:

$$\begin{aligned} \hat{\mathbb{I}}|f(t)\rangle &= \left| 1/\sqrt{2\pi} \right\rangle \left\langle 1/\sqrt{2\pi} | f(t) \right\rangle + \sum_{n=1}^{\infty} \left| 1/\sqrt{\pi} \cos(nt) \right\rangle \left\langle 1/\sqrt{\pi} \cos(nt) | f(t) \right\rangle \\ &\quad + \sum_{n=1}^{\infty} \left| 1/\sqrt{\pi} \sin(nt) \right\rangle \left\langle 1/\sqrt{\pi} \sin(nt) | f(t) \right\rangle \end{aligned} \quad (6)$$

For odd functions $f(t)$: $\langle 1/\sqrt{2\pi} | f(t) \rangle = 0$ and $\langle 1/\sqrt{\pi} \cos(nt) | f(t) \rangle = 0$. For even functions $f(t)$: $\langle 1/\sqrt{\pi} \sin(nt) | f(t) \rangle = 0$. Consequently, the following special cases can be identified.

¹The Dirichlet conditions dictate that the function $f(t)$ on the interval $[-\pi, \pi]$ may only possess a finite number of finite jumps and no infinite jumps.

Resolution of identity for the space of even functions

$$\hat{\mathbb{I}}_e = \left| 1/\sqrt{2\pi} \right\rangle \left\langle 1/\sqrt{2\pi} \right| + \sum_{n=1}^{\infty} \left| 1/\sqrt{\pi} \cos(nt) \right\rangle \left\langle 1/\sqrt{\pi} \cos(nt) \right| \quad (7)$$

Resolution of identity for the space of odd functions

$$\hat{\mathbb{I}}_o = \sum_{n=1}^{\infty} \left| 1/\sqrt{\pi} \sin(nt) \right\rangle \left\langle 1/\sqrt{\pi} \sin(nt) \right| \quad (8)$$

2.2.2. Partitioning of \mathbb{L}_2 -space into Spaces \mathbb{L}_2^e and \mathbb{L}_2^o

Let \mathbb{L}_2 be partitioned into \mathbb{L}_2^e and \mathbb{L}_2^o , subspaces of quadratically integrable even- and odd-functions. Then \mathbb{L}_2 can be written as the direct sum of \mathbb{L}_2^e and \mathbb{L}_2^o . Denote the identity operators for \mathbb{L}_2 , \mathbb{L}_2^e and \mathbb{L}_2^o , by $\hat{\mathbb{I}}$, $\hat{\mathbb{I}}_e$, and $\hat{\mathbb{I}}_o$, respectively. Then: $\mathbb{L}_2 = \mathbb{L}_2^e \oplus \mathbb{L}_2^o \iff \hat{\mathbb{I}} = \hat{\mathbb{I}}_e + \hat{\mathbb{I}}_o$

Even- and odd-symmetric parts of an arbitrary function on $[-\pi, \pi]$. Let $f(t)$ be a function defined on $[-\pi, \pi]$. Then, $f(t) = f_e(t) + f_o(t)$ with $f_e(t) = 1/2f(t) + 1/2f(-t)$ and $f_o(t) = 1/2f(t) - 1/2f(-t)$. Here, the relationships $f_e(-t) = f_e(t)$ and $f_o(-t) = -f_o(t)$ hold valid.

2.2.3. On the Relationships $\mathbb{L}_2 = \mathbb{L}_2^e \oplus \mathbb{L}_2^o$ and $\hat{\mathbb{I}} = \hat{\mathbb{I}}_e + \hat{\mathbb{I}}_o$

Applying $\hat{\mathbb{I}}$ onto $f(t)$, introducing $f_e(t)$ and $f_o(t)$, and considering the mentioned properties associated with partitioning of \mathbb{L}_2 into \mathbb{L}_2^e and \mathbb{L}_2^o , the following relationships can be obtained:

$$\hat{\mathbb{I}}|f(t)\rangle = \left| 1/\sqrt{2\pi} \right\rangle \left\langle 1/\sqrt{2\pi} |f_e(t) \right\rangle + \sum_{n=1}^{\infty} \left| 1/\sqrt{\pi} \cos(nt) \right\rangle \left\langle 1/\sqrt{\pi} \cos(nt) |f_e(t) \right\rangle \quad (9a)$$

$$+ \sum_{n=1}^{\infty} \left| 1/\sqrt{\pi} \sin(nt) \right\rangle \left\langle 1/\sqrt{\pi} \sin(nt) |f_o(t) \right\rangle = \hat{\mathbb{I}}_e|f_e(t)\rangle + \hat{\mathbb{I}}_o|f_o(t)\rangle \quad (9b)$$

3. MULTIREOLUTION ANALYSIS AND WAVELETS

Let $\varphi(t)$ and $\psi(t)$ denote the scaling function and the wavelet constituting a Multiresolution Analysis (MRA) in Hilbert space \mathbb{L}_2 . Let the integer-translates of $\varphi(t)$ generate the function space ν_0 . Let $\psi(t)$ and its 2^k -dyadically compressed and 2^{-k} -translated copies generate spaces $\mathcal{W}_0, \mathcal{W}_1, \mathcal{W}_2, \dots$, with $\mathbb{L}_2 = \nu_0 \oplus \mathcal{W}_0 \oplus \mathcal{W}_1 \oplus \mathcal{W}_2 \oplus \dots$. Denote $\varphi_{0,n}(t) = \varphi(t-n)$ and $\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k)$. Assume an arbitrary function $f(t) \in \mathbb{L}_2$. Let the following synthesis formula be valid:

$$|f(t)\rangle = \sum_{n=-\infty}^{\infty} |\varphi_{0,n}(t)\rangle c_n + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |\psi_{j,k}(t)\rangle d_{j,k} \quad (10)$$

Assume orthonormality of the involved expansion functions. Project $|f(t)\rangle$ onto $|\varphi_{0,n}(t)\rangle$ and $|\psi_{j,k}(t)\rangle$, respectively, to obtain $c_n = \langle \varphi_{0,n}(t) | f(t) \rangle$, and $d_{j,k} = \langle \psi_{j,k}(t) | f(t) \rangle$. Write $\hat{\mathbb{I}}|f(t)\rangle$ for $|f(t)\rangle$ at the LHS and substitute for c_n and $d_{j,k}$ into (10) to obtain:

$$\hat{\mathbb{I}}|f(t)\rangle = \sum_{k=-\infty}^{\infty} |\varphi_{0,n}(t)\rangle \langle \varphi_{0,n}(t) | f(t) \rangle + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |\psi_{j,k}(t)\rangle \langle \psi_{j,k}(t) | f(t) \rangle \quad (11)$$

Since $|f(t)\rangle$ is arbitrary, deduce the following resolution of the identity operator $\hat{\mathbb{I}}$:

$$\hat{\mathbb{I}} = \sum_{k=-\infty}^{\infty} |\varphi_{0,n}(t)\rangle \langle \varphi_{0,n}(t)| + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} |\psi_{j,k}(t)\rangle \langle \psi_{j,k}(t)| \quad (12)$$

In particular let j run from $-\infty$ to ∞ . Then the first term at the RHS becomes the null operator and the following alternative resolution of identity can be obtained:

$$\hat{\mathbb{I}} = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |\psi_{j,k}(t)\rangle \langle \psi_{j,k}(t)| \quad (13)$$

4. FRAMES

Let \mathbb{V}_N denote an N dimensional vector space. Let $\mathcal{N} > N$. A set of \mathcal{N} ket-vectors $\{|f_n\rangle|n \in \mathcal{N}\}$ in \mathbb{V}_N is said to constitute a frame if for arbitrary $|f\rangle \in \mathbb{V}_N$ there are constants A and B (frame bounds), subject to the conditions $0 < A \leq B < \infty$, such that the following inequalities hold true:

$$A\|f\|^2 \leq \sum_{n=1}^{\mathcal{N}} |\langle f_n|f\rangle|^2 \leq B\|f\|^2 \quad (14)$$

Remarks. i) Since $|f_n\rangle \in \mathbb{V}_N$, $|f_n\rangle$ has N (complex valued) components. ii) Since $\mathcal{N} > N$ the vectors $|f_n\rangle$, with $n = 1, \dots, \mathcal{N}$, are over-complete in \mathbb{V}_N . iii) As shall be seen momentarily the middle term in (14) involves an operator ($\hat{\mathbb{S}}$). The inequalities in (14) ensure the existence of the inverse of $\hat{\mathbb{S}}$. $\hat{\mathbb{S}}^{-1}$ is crucial in constructing the dual frame corresponding to the assumed frame.

4.1. Interpretation of the Terms in (14) and the Frame Operator $\hat{\mathbb{S}}$

The term $\|f\|^2$ being the magnitude of $|f\rangle$ squared, is given by the inner product $\langle f|f\rangle$, or, $\langle f||f\rangle$. Inserting the identity operator $\hat{\mathbb{I}}$ between $\langle f|$ and $|f\rangle$ it results in $\langle f|\hat{\mathbb{I}}|f\rangle$. Thus with A being a constant: $A\|f\|^2 = \langle f|\{\hat{A}\hat{\mathbb{I}}\}|f\rangle$. Similarly, $B\|f\|^2 = \langle f|\{\hat{B}\hat{\mathbb{I}}\}|f\rangle$. In the middle term $\sum_{n=1}^{\mathcal{N}} |\langle f_n|f\rangle|^2$, the inner product $\langle f_n|f\rangle$ is in general complex-valued. The magnitude squared $|\langle f_n|f\rangle|^2$ of $\langle f_n|f\rangle$ can be written in the form $(\langle f_n|f\rangle)^* (\langle f_n|f\rangle)$. Thus, with $(\langle f_n|f\rangle)^* = \langle f|f_n\rangle$, and the conventions $\langle \cdot|\cdot\rangle = \langle \cdot||\cdot\rangle$:

$$\sum_{n=1}^{\mathcal{N}} |\langle f_n|f\rangle|^2 = \langle f|\left\{\sum_{n=1}^{\mathcal{N}} |f_n\rangle\langle f_n|\right\}|f\rangle = \langle f|\hat{\mathbb{S}}|f\rangle \quad (15)$$

In the above $|f_n\rangle\langle f_n|$ was recognized as an operator sandwiched between $\langle f|$ and $|f\rangle$, the linearity property was exploited, and the frame operator $\hat{\mathbb{S}}$ was introduced in the obvious manner.

Frame bounds. With the above equivalent representations, (14) takes the form:

$$\langle f|\{\hat{A}\hat{\mathbb{I}}\}|f\rangle \leq \langle f|\{\hat{\mathbb{S}}\}|f\rangle \leq \langle f|\{\hat{B}\hat{\mathbb{I}}\}|f\rangle \quad (16)$$

Since the choice of $|f\rangle$ was arbitrary, (16) implies: $\hat{A}\hat{\mathbb{I}} \leq \hat{\mathbb{S}} \leq \hat{B}\hat{\mathbb{I}}$. The validity of these inequalities ensures the existence of $\hat{\mathbb{S}}^{-1}$, the inverse of the frame operator $\hat{\mathbb{S}}$.

Remark. The frame operator $\hat{\mathbb{S}}$, in virtue of its definition, involves the frame vectors $|f_n\rangle$, and their hermitian conjugates $\langle f_n|$, which are constructed from $|f_n\rangle$ by transposition of $|f_n\rangle$ followed by complex conjugation of the individual components.

4.2. Interpretation of the Frame Operator $\hat{\mathbb{S}}$

In order to interpret $\hat{\mathbb{S}}$, apply it to an arbitrary function $|f\rangle \in \mathbb{V}_N$:

$$\hat{\mathbb{S}}|f\rangle = \sum_{n=1}^{\mathcal{N}} |f_n\rangle\langle f_n|f\rangle \quad (17)$$

Remarks. i) Construction of $\hat{\mathbb{S}}|f\rangle$ requires building the \mathcal{N} terms at the RHS of (17) and summing them up. ii) The generic term $|f_n\rangle\langle f_n|f\rangle$ involves the application of the operator $|f_n\rangle\langle f_n|$ to an arbitrarily chosen function $|f\rangle \in \mathbb{V}_N$. iii) To obtain $|f_n\rangle\langle f_n|f\rangle$ first $\langle f_n|f\rangle$ needs to be built. Then the resulting (complex) number $(\langle f_n|f\rangle)$ is multiplied by the ket-vector $|f_n\rangle$.

Consequently, to form $\hat{\mathbb{S}}|f\rangle$, proceed as follows:

1. Project $|f\rangle$ onto $|f_n\rangle$; i.e., build $\langle f_n|f\rangle$ for $n \in \mathcal{N}$. *This step is akin to determining Fourier coefficients (the analysis-step). This analogy should not be stretched too much though!*
2. Having determined $\langle f_n|f\rangle$, it is subsequently multiplied by $|f_n\rangle$ to form $|f_n\rangle(\langle f_n|f\rangle)$. Adding the resulting terms for $n \in \mathcal{N}$, the RHS in (17) is obtained. *This step is akin to synthesis step in Fourier transform. Again this analogy should not be stretched excessively.*

Remarks. Obviously it cannot be expected that the RHS side in (17) reconstructs $|f\rangle$, as this would be the case for an ONB. In virtue of (17) it is seen that the above “analysis” and “synthesis” steps, after being applied to $|f\rangle$, result in $\hat{\mathbb{S}}|f\rangle$, rather than $|f\rangle$.

Having identified $\hat{\mathbb{S}}$, the inequalities in (16) read:

$$A\langle f|f\rangle \leq \langle f|\hat{\mathbb{S}}|f\rangle \leq B\langle f|f\rangle \quad (18)$$

The term $\langle f|\hat{\mathbb{S}}|f\rangle$, the inner product of $\hat{\mathbb{S}}|f\rangle$ and $|f\rangle$, is a measure for the similarity between $\hat{\mathbb{S}}|f\rangle$ and $|f\rangle$. For $\mathcal{N} = N$, and $\{|f_n\rangle|n \in N\}$ being an ONB, $\hat{\mathbb{S}}|f\rangle$ equals $|f\rangle$, $A = B = 1$, and, consequently, the equalities hold valid in (18). *The frame concept embeds spectral analysis and synthesis, as a special case.*

4.3. Dual Frames and the Resolution of Identity

The RHS of (17) provides an expression for $\hat{\mathbb{S}}|f\rangle$. To obtain $|f\rangle$ both sides of (17) are multiplied by $\hat{\mathbb{S}}^{-1}$, which in virtue of the frame inequalities exists. Thus:

$$|f\rangle = \sum_{n=1}^{\mathcal{N}} \hat{\mathbb{S}}^{-1}|f_n\rangle\langle f_n|f\rangle \quad (19)$$

Introducing the dual frame vectors $|\widetilde{f}_n\rangle$ according to $|\widetilde{f}_n\rangle = \hat{\mathbb{S}}^{-1}|f_n\rangle$, Eq. (19) reads:

$$|f\rangle = \sum_{n=1}^{\mathcal{N}} |\widetilde{f}_n\rangle\langle f_n|f\rangle \quad (20)$$

Resolution of identity. Writing the LHS as $\hat{\mathbb{I}}|f\rangle$, and remembering that $|f\rangle$ is arbitrary, result in the resolution of identity in terms of frames and dual frames:

$$\hat{\mathbb{I}} = \sum_{n=1}^{\mathcal{N}} |\widetilde{f}_n\rangle\langle f_n| \quad (21)$$

5. CONCLUSION

The concepts of orthonormal bases, multiresolution analysis (scaling functions and wavelets) and over-complete sets of analysis- and synthesis functions (frame and dual frames) were discussed in a unified form, by resolving the identity operator. The orthonormal eigenfunctions of hermitian operators and Green’s function based wavelets and orthonormal bases were briefly mentioned.

ACKNOWLEDGMENT

This work is based on the research supported in part by the National Research Foundation (UID: 85889). A start-up grant, provided by the Engineering and Built Environment (EBE) Faculty at UCT, is gratefully acknowledged. The initial stage of the work was carried out under an Australian Research Council (ARC) Linkage Grant: LP0775463.

REFERENCES

1. Burrus, C. S., R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, Prentice Hall, Englewood, NJ, 1998.
2. Christensen, O., *An Introduction to Frames and Riesz Bases*, Birkhauser, 2002.
3. Baghai-Wadji, A. B., “A gentle introduction to the theory of frames and dual frames for engineers in electromagnetics,” *EMC News Letters*, Melbourne, Australia, 2012.
4. Baghai-Wadji, A. B. and G. G. Walter, “Green’s function-based wavelets,” *Proceedings of International Ultrasonics Symposium*, Puerto Rico, 2000.
5. Baghai-Wadji, A. B. and G. G. Walter, “Green’s function-based wavelets: Selected properties,” *Proceedings of International Ultrasonics Symposium*, Puerto Rico, 2000.

Asymptotic Analysis of the Whistler Waves Propagation in Space Plasma Thrusters

D. Melazzi¹, A. Cardinali², M. Manente³, and D. Pavarin^{1,4}

¹CISAS “G. Colombo” Center of Studies and Activities for Space, University of Padua, Italy

²ENEA-EURATOM Association, Frascati, Rome, Italy

³Hit09, S.r.l., Padova, Italy

⁴Department of Industrial Engineering, University of Padua, Italy

Abstract— We present a three-dimensional Ray-Tracing solver, called RAYWh (RAY-tracing Whistler), for the electromagnetic propagation and power deposition in plasma sources for space thrusters, where high density plasmas (ranging from 10^{17} to 10^{19} particles/m³) are confined by general magnetic configurations with magnitude below < 0.15 T. The 3D Maxwell-Vlasov equations are solved by means of a WKB asymptotic expansion, to investigate the propagation and absorption of whistler waves (excitation frequency is 13.56 MHz) under the influence of general confinement magnetic field, and axisymmetric realistic density profiles. The reduced set of the WKB equations for the wave phase and for the square amplitude of the electric field are solved numerically by means of Hamming’s modified predictor-corrector method; the verification of WKB hypothesis are monitored during the simulation.

The Ray Tracing approach is employed — for the first time — in the analysis of plasma sources for space plasma thrusters. A direct comparison between common helicon sources with axial, constant and uniform confinement magnetic field, and plasma sources with actual confinement magnetic field lines revealed a propagative picture with unconventional mode conversions, cut-offs and resonances inside the source that affect the power deposition.

1. INTRODUCTION

Recent advances in plasma-based propulsion systems have led to the development of electromagnetic (EM) Radio-Frequency (RF) plasma generation and acceleration systems, called Helicon Plasma Thruster (HPT) and derived from high density industrial helicon plasma sources [1]. HPT main components are: a gas feeding system, an RF antenna, and magnetic coils. The feeding system injects a neutral gas into a cylindrical vessel, wrapped by a RF antenna system working in the MHz range, ionizing the neutral gas and heating the plasma. The magnetic coils provide the axial magnetic field allowing for the propagation of whistler waves, and the confinement of plasma inside the cylindrical source. Additionally, the magnetic field lines at the exhaust section have to become divergent providing a magnetic nozzle effect on the magnetized plasma. The HPT has been derived from industrial helicon sources and its propulsive figures of merit (i.e., specific impulse, thrust efficiency) depend on the EM energy deposited into the plasma [2].

Among the available plasma sources, helicon sources have been found much more efficient at depositing EM power, and thus at generating dense plasmas. Plasma densities up to 10^{19} particles/m³ can be reached by using moderate magneto-static fields (below < 0.1 T). A helicon source consists of a dielectric tube surrounded by coils which generate a weak magneto-static field (up to 0.15 T) and an RF antenna working in the range of frequencies $\Omega_{ci} \ll \omega_{lh} \ll \omega \ll \Omega_{ce}$, where Ω_{ci} (Ω_{ce}) is the ion (electron) cyclotron angular frequency, and ω_{lh} is the lower-hybrid frequency. Different models have been developed to study, design and optimize such a plasma propulsion system [3, 4]. All these models rely on the assumption that the confinement magnetic field is purely axial, constant and uniform as expected in helicon sources; however, experimental setups for HPTs can depart from this configuration due to dimension, mass and power budget limitations for space application purposes, leading to general confinement magnetic field lines. By means of Ray-Tracing technique [5], we can include with a minimum effort the 3D characteristics of the wave propagation and absorption in a plasma source of finite dimensions, with a general confinement magnetic field and plasma density profiles. The paper is organized as follows. Section 2 gives a brief review of the methodology adopted. Section 3 presents the numerical results. A time dependence in the form $\exp(i\omega t)$ for wave quantities is assumed and suppressed throughout the rest of this paper.

2. WKB ASYMPTOTIC EXPANSION OF WHISTLER WAVE EQUATION

As far as the the whistler propagation is concerned, the plasma wave interaction described by the Maxwell-Vlasov system of equations can be simplified by making the hypothesis that the field amplitude is sufficiently small to justify the linearization of the kinetic equation [6]; additionally, the physics of wave propagation and absorption is well described in the cold plasma limit, thus the wave equation for the electric field reads:

$$\nabla(\nabla \cdot \mathbf{E}(\mathbf{r})) - \nabla^2 \mathbf{E}(\mathbf{r}) - \frac{\omega^2}{c^2} \varepsilon^H \cdot \mathbf{E}(\mathbf{r}) = \frac{\omega^2}{c^2} i\varepsilon^A \cdot \mathbf{E}(\mathbf{r}) \quad (1)$$

where $\mathbf{r} = (r, \theta, z)$ is the position, ε^H and ε^A are the Hermitian and anti-Hermitian parts, respectively, of the dielectric tensor [6]. It is worth recalling that the dielectric tensor depends upon the local confinement magnetic field, plasma density, and collision frequency.

Equation (1) is solved asymptotically via the WKB method. At the lowest order in the expansion, we have the Ray-Tracing equations

$$\begin{aligned} \frac{dx}{d\tau} &= \delta_0^{-1} \frac{\partial \mathcal{H}}{\partial n_x}; & \frac{d\theta}{d\tau} &= \delta_0^{-1} \frac{\partial \mathcal{H}}{\partial m}; & \frac{d\hat{z}}{d\tau} &= \delta_0^{-1} \frac{\partial \mathcal{H}}{\partial n_z} \\ \frac{dn_x}{d\tau} &= -\delta_0^{-1} \frac{\partial \mathcal{H}}{\partial x}; & \frac{dm}{d\tau} &= -\frac{\partial \mathcal{H}}{\partial \theta}; & \frac{dn_z}{d\tau} &= -\delta_0^{-1} \frac{\partial \mathcal{H}}{\partial \hat{z}} \\ \frac{d\mathbf{S}_0}{d\tau} &= n_x \frac{\partial \mathcal{H}}{\partial n_x} + m \frac{\partial \mathcal{H}}{\partial m} + n_z \frac{\partial \mathcal{H}}{\partial n_z} \\ \frac{dt}{d\tau} &= \frac{\partial \mathcal{H}}{\partial \omega}; & \frac{d\omega}{d\tau} &= -\frac{\partial \mathcal{H}}{\partial t} = 0 \Rightarrow \omega = \text{const} \end{aligned} \quad (2)$$

where $\delta_0 = \omega a c^{-1}$ is the expansion parameter, a is the plasma radius, $\mathbf{k}c/\omega = (n_x, m, n_z)$ are the components in cylindrical coordinate of the wave-number, the normalized coordinates along the radial and axial directions are ($x = r/a$; $\hat{z} = z/a$), τ is the variable along the wave trajectory, ω is the frequency given by the exciting antenna, \mathbf{S}_0 is the wave phase.

Equation (2) provides the characteristic curves that define a bundle of trajectories in the phase space (\mathbf{r}, \mathbf{k}) , with \mathbf{k} the wave vector. The function $\mathcal{H}(\mathbf{r}, n_{\parallel}, n_{\perp})$

$$\mathcal{H}(\mathbf{r}, n_{\parallel}, n_{\perp}) = A(\mathbf{r}, n_{\parallel}) n_{\perp}^4 + B(\mathbf{r}, n_{\parallel}) n_{\perp}^2 + C(\mathbf{r}, n_{\parallel}) = 0 \quad (3)$$

is the cold electromagnetic dispersion relation, where $n_{\perp} = \nabla_{\perp} \mathbf{S}_0$ and $n_{\parallel} = \nabla_{\parallel} \mathbf{S}_0$ are the perpendicular and parallel (with reference to the external magnetic field) components of the wave-number, respectively. Coefficients in Eq. (3) are a combination of the dielectric tensor elements.

At the next order in the WKB expansion, we have a partial differential equation for the wave amplitude $|\mathbf{A}_0|$, which is cumbersome and difficult to solve. As shown in [5], it is more convenient to write the Poynting Theorem for the wave energy conservation starting from the WKB amplitude equation, then integrating it over the plasma volume; we have the Power damping equation

$$\begin{aligned} \frac{dP}{dt} &= -2\gamma(\mathbf{r}, \mathbf{k}, \omega) P, \quad P = \frac{\omega}{16\pi} \int d\Sigma \cdot \frac{\partial \mathcal{H}(\mathbf{k}, \omega)}{\partial \mathbf{k}} |_{\mathcal{A}_0}|^2 \\ \gamma(\mathbf{r}, \mathbf{k}, \omega) &= \frac{\mathbf{e}_0^* \cdot \varepsilon^A(\mathbf{r}) \cdot \mathbf{e}_0}{\partial \mathcal{H}(\mathbf{k}, \omega) / \partial \omega} \end{aligned} \quad (4)$$

where P is the power carried by the single ray, γ is the power damping rate (in s^{-1}) accounting for Landau and collisional damping, $\mathbf{e}_0 = \mathbf{E}/|\mathbf{E}|$ is the electric field unit vector. Eqs. (2)–(4) represent the Ray-Tracing and Power damping equations, and they are an ensemble of independent initial value problems, whose initial conditions give the starting wave vector and power at each point of a reference surface of the propagating wave. The integration of these equations requires initial conditions for (\mathbf{r}, \mathbf{k}) on an initial reference surface $(\Gamma, \mathbf{S}_0|_{\Gamma})$, which is the image of the launching wave antenna on the plasma boundary. We solved Ray-Tracing and Power damping equations numerically by means of Hamming's modified predictor-corrector method; we used a fourth order Runge-Kutta method suggested by Ralston to adjust the initial increment, and to compute the starting values for the non self-starting predictor-corrector method. During the simulation, we

verified that WKB hypotheses are satisfied, namely: (i) the wavelength of the propagating mode is much less than the characteristic scale length, (ii) absence of diffraction effects in the wave propagation.

The solution of Eqs. (2)–(4) allows for information on the wave propagation and power deposition in a general 3D geometry which accounts for the 3D coordinate dependence of the external confinement magnetic field structure and 2D (radial and axial) plasma density profiles.

3. WAVE PROPAGATION AND POWER DEPOSITION ANALYSIS

We considered an actual plasma thruster made of a cylindrically-shaped plasma source with plasma radius $a = 10$ cm, and axial length $L = 40$ cm. The confinement magnetic field is provided by means of a solenoid wrapped around the plasma cylinder, and fed by a current resulting in a magnetic field on the axis $B_0 = 0.025$ T. We assumed a parabolic density profile along the radial direction, and a Gaussian density profile along the axial direction

$$n(x, \hat{z}) = n_0 \left(1 - (f_a x)^2\right) e^{-\hat{z}^2} \quad (5)$$

where $f_a = \sqrt{1 - \frac{n_{edge}}{n_0}}$, with $n_0 = 10^{18} \text{ m}^{-3}$, $n_{edge} = 5 \times 10^{17} \text{ m}^{-3}$. The electron (ion) temperature is uniform $T_e = 3 \text{ eV}$ ($T_i = 0.1 \text{ eV}$); the neutral pressure is $p_n = 10 \text{ mTorr}$ (introduced via Krook model). The antenna excited an azimuthal mode $m = 0$ at a frequency $f = 13.56 \text{ MHz}$, and $30 \leq n_z < 90$ is the propagative axial spectrum allowed by the plasma parameters chosen. The analytical expression available for plasma density profiles and magneto-static field allow for the evaluation of derivatives in Eq. (2).

We considered three different configurations for the wave propagation and power deposition analysis: (a) 2D confinement magnetic field and radial density profile, (b) 2D confinement magnetic field and 2D density profile, (c) axial, uniform and constant confinement magnetic field and radial density profile (helicon case). These configurations have been reported in Table 1. Each test label is used as label for the curves in the following pictures.

Table 1: Test cases considered. \mathbf{e}_r and \mathbf{e}_z are the unit vectors along the radial and axial directions.

Label	Density profile	Confinement magnetic field
a)	radial	$B_r(r, z) \mathbf{e}_r + B_z(r, z) \mathbf{e}_z$
b)	radial & axial	$B_r(r, z) \mathbf{e}_r + B_z(r, z) \mathbf{e}_z$
c)	radial	$B_0 \mathbf{e}_z = \text{const}$

3.1. Discussion on Wave Propagation Properties

The axial wavenumber n_z , the parallel wavenumber $n_{||}$, the perpendicular wavenumber n_{\perp} have been plotted as a function of the radial normalized variable x in Figs. 1(a), (b), (c), when the fast

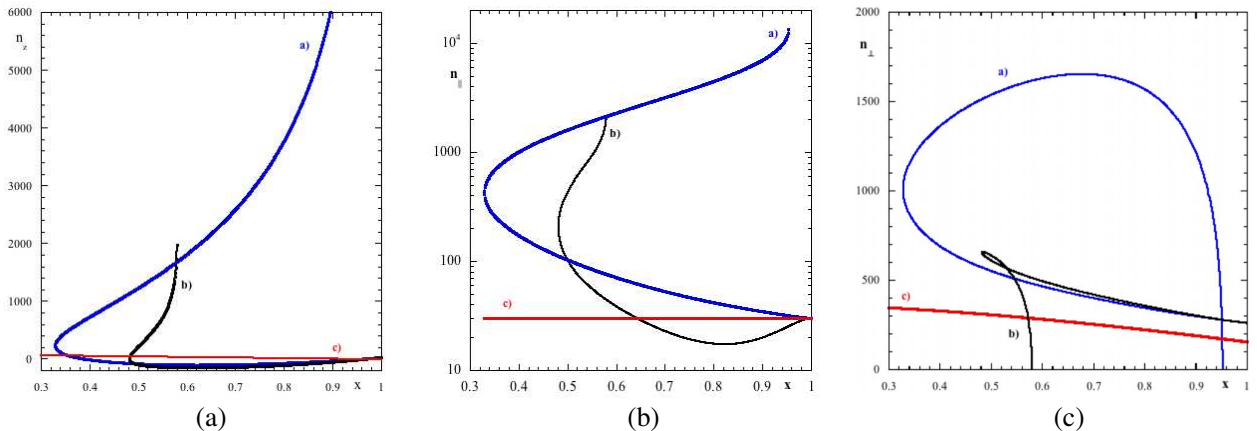


Figure 1: (a) axial, (b) parallel, (c) perpendicular components of the wavenumber along the radial normalized variable x for cases a), b), c).

wave with $m = 0$ and $n_z = 30$ is launched at the plasma edge. Consider the axial component of the wavenumber as shown in Fig. 1(a). In the c) case, the axial wavenumber remained constant inside the plasma as expected by Eq. (2). In a) and b) the magnetic field induces a variation of the axial wavenumber along the ray. In the case a), after a radial reflection, an axial resonance appears near the plasma edge, while in the case b) the resonance appears just after the radial reflection.

Consider the parallel and the perpendicular (to the external magnetic field) components of the wavenumber as shown in Figs. 1(b), (c). The waves corresponding to curves a) and b) feel the parallel resonance $n_{\parallel} \rightarrow \infty$, while on the same location the perpendicular wavenumbers show a cut-off $n_{\perp} \rightarrow 0$. The occurrence of the resonance and cut-off is located around $x \approx 0.95$ for the case a) and $x \approx 0.6$ for the case b), respectively. No resonance or cut-off occurs for the case c), where the excited whistler wave propagates towards the plasma center without changing the polarization, and no singular points are met.

3.2. Discussion on Power Deposition Properties

We discuss the wave absorption features related to cases a) and b). At the plasma edge, the fast wave is launched with $m = 0$ and $n_z = 30$ with P_0 power; along the trajectory, the power damping P/P_0 gives the percentage of the power the wave is actually carrying, while the amount of power missing from the edge value is the percentage of the power absorbed by the plasma. The power deposition profile gives the power density deposited by the wave along the trajectory. In Fig. 2(a), the power damping is pictured for a ray path such that the wave gets reflected radially and gets back to $x \approx 0.94$. Two particular points can be identified. The first one is around $x \approx 0.33$ where the wave is reflected radially. The second point is at $x \approx 0.45$ where the fast wave couples to a slow wave. At this point, the power damping curve changes its slope, and this is related to the different ways the slow and the fast wave modes deposit energy into the plasma. This feature becomes even more evident in Fig. 2(c), where the power deposition profile is pictured for the same case a). The fast mode couples more and more power into the plasma travelling toward the axis up to $x \approx 0.45$, where it is converted to the slow mode. The mode conversion is clearly represented by the jump

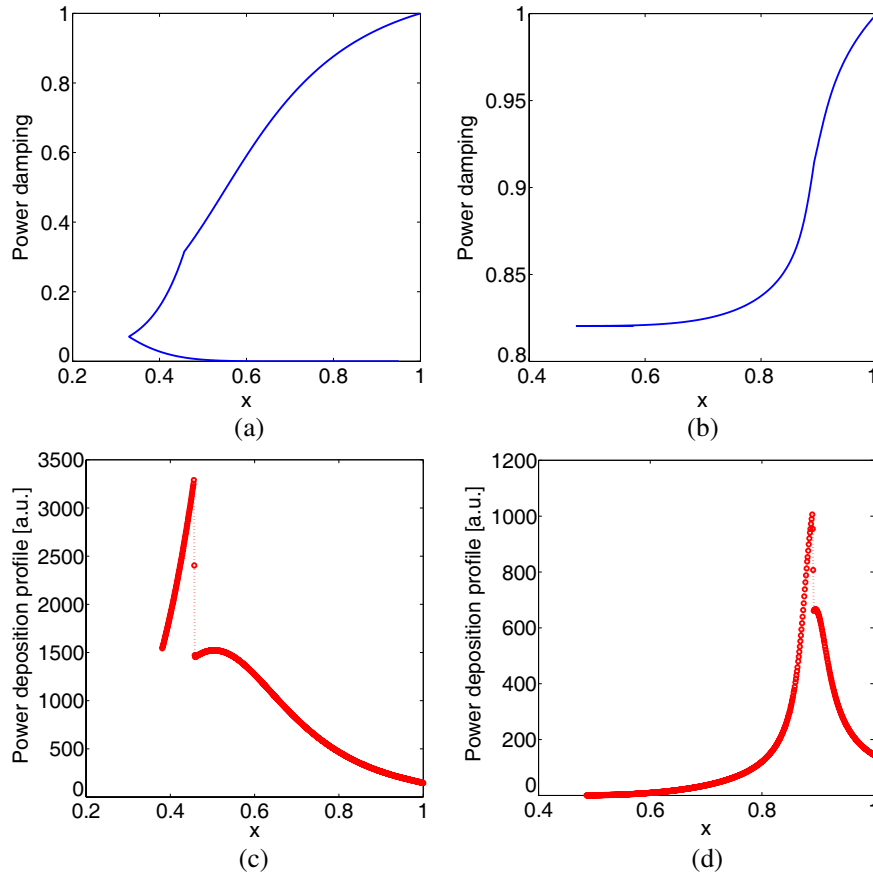


Figure 2: Power damping along x , for (a) case a), (b) case b). Power deposition profile in arbitrary unit along x , for (c) case a), (d) case b).

in the power deposition profile; then the slow mode keeps propagating and depositing energy up to $x \approx 0.38$, where the wave reaches the axial boundary of the cylinder. It is noteworthy that the slow mode couples power into the plasma due to collisional processes more efficiently than the fast one, being rapidly damped after the mode conversion.

In Fig. 2(b), the power damping has been pictured for a ray path such that the wave gets reflected radially and gets back up to $x \approx 0.57$. From Fig. 2(b) just one particular point can be identified, which is at $x \approx 0.48$, where the wave is reflected radially. There is actually another point at $x \approx 0.9$, where the fast wave launched at the edge couples to a slow wave. This feature is not easily visible in Fig. 2(b), where there is no clear change in the slope of the power damping curve; however, it can be clearly recognized in the jump in the power deposition profile showed in Fig. 2(d). After the mode conversion, the slow wave is rapidly damped by collisions before reaching the axis.

4. CONCLUSIONS

We developed a three-dimensional Ray-Tracing solver called RAYWh, and we used it to study — for the first time — the electromagnetic propagation and power deposition in a cylindrically-shaped plasma source for space plasma thrusters. Previous approaches [3, 4] (to mention a few) cannot provide accurate information on the power deposited into the plasma, when real thruster setups are considered. Indeed, magnetic confinement configurations in actual HPTs can depart from the simplified helicon one, due to dimension, mass and power budget limitations, leading to different power coupling levels into the plasma, and resulting in different propulsive characteristics. The approach implemented in RAYWh provides the evolution of the wave vector and power deposition profiles inside the plasma source, where realistic density profiles and confinement magnetic field lines can be readily included without any approximation. Unlike the helicon case, parallel and perpendicular wavenumbers changed during the wave trajectory leading to a completely different propagative picture, when actual confinement magnetic configurations and plasma density profiles are considered. Unexpected cut-offs, resonances, radial reflections, and mode conversions of the excited waves have been found, as a result of the confinement magnetic field along with variation in the plasma density the waves encountered. These in turn influenced the power deposition phenomena. The results are relevant for space thruster applications, but they can be fruitfully employed in industry plasma sources for the identification of the best source configuration (in terms of confinement magnetic field lines in addition to actual plasma density profiles), thus providing the maximum power transfer from the RF antenna to the plasma.

REFERENCES

1. Pavarin, D., et al., “Design of 50 W Helicon plasma thruster,” *Proceedings of 31th International Electric Propulsion Conference, IEPC-2009-205*, Ann Arbor, Michigan, USA, 2009.
2. Ahedo, E., “Plasmas for space propulsion,” *Plasma Phys. Control. Fusion*, Vol. 53, 2011.
3. Mouzouris, Y. and J. E. Scharer, “Modeling of profile effects for inductive helicon plasma sources,” *IEEE Trans. Plasma Science*, Vol. 24, 152–160, 1996.
4. Melazzi, D., D. Curreli, M. Manente, J. Carlsson, and D. Pavarin, “SPIREs: A finite-difference frequency-domain electromagnetic solver for inhomogeneous magnetized plasma cylinders,” *Comp. Phys. Comm.*, Vol. 183, No. 6, 1182–1191, 2012.
5. Brambilla, M. and A. Cardinali, *Plasma Physics and Controlled Nuclear Fusion*, Vol. 24, 1187–1218, 1982.
6. Brambilla, M., *Kinetic Theory of Plasma Waves*, Oxford Science Publication, Clarendon Press, Oxford, 1998.

Electromagnetic and Gravitational Origin of Dark Energy in Kaluza-Klein $D = 5$ Spacetime

Mohamed S. El Naschie

Department of Physics, Faculty of Science, Alexandria University, Alexandria, Egypt

Abstract— Kaluza-Klein (K-K) theory is generally accepted as a partial unification of two of the fundamental interactions in Nature namely electromagnetism and gravity. In the present work against conventional expectation, we show using transfinite quantization (i.e., fractals) that K-K theory in conjunction with Penrose’s non-commutative (fractal) space tiling leads to dual quantum relativity equations. The first is the ordinary energy equation of a quantum particle $E(O) = (\phi^5/2)mc^2$ where E is the energy, m is the mass, c is the speed of light and $\phi = 2/(\sqrt{5}+1)$. On the other hand, the second dual equation corresponding to the dark energy of the quantum wave, i.e., of the empty set is given by $E(D) = (5\phi^2/2)mc^2$.

Finally a third equation resulting from summing up the ordinary energy of the quantum particle and the dark energy of the quantum wave namely

$$E(O) + E(D) = (\phi^5 + 5\phi^2) mc^2/2 = 2mc^2/2 = mc^2$$

is nothing else but Einstein’s famous equation.

Recalling that $E(O) = (\phi^5/2)mc^2 \simeq mc^2/22$ is almost exactly equal to the value found from the WMAP and supernova cosmological measurement which was awarded the 2011 Nobel Prize in Physics, we conclude that $E = mc^2$ is blind to any distinction between the ordinary energy which we can measure and the dark energy presumed so far to be “missing”. It is further reasoned that ordinary energy is of opposite sign to dark energy which could be mathematically traced to the negative sign of the empty set of the quantum wave or physically to the compactified dimension which is behind negative gravity and the recently measured acceleration of cosmic expansion.

1. INTRODUCTION: DARK ENERGY IN A NUTSHELL?

Due to space limitation the present work is a very condensed paper giving only a glimpse into our attempt to resolving the fundamental problem of the missing dark energy of the cosmos. Take a line segment representing a unit interval which in turn represents a one dimensional “spacetime” (see Fig. 1). We start by randomly removing parts of this line except for the end points in a manner reminiscent of what we do when we construct a deterministic middle third Cantor set but adding randomness to the iteration [1, 2]. In the case of the classical Cantor set we end up, after infinitely many iterations with a Hausdorff dimension amounting to $\ln 2/\ln 3 \simeq 0.63$ [1–4]. Noting that we have nothing left except unaccountably many points of topological dimension equal zero [1–4], a dimension equal 0.63 is relatively speaking quite substantial [2–4]. For the random Cantor set at hand, the situation is quite similar but instead of $\ln 2/\ln 3$ as a Hausdorff dimension we end up with the remarkable golden mean value $(\sqrt{5} - 1)/2 = \phi \simeq 0.618033$ as was shown some time ago by American mathematicians Mauldin and Williams [5–7]. Considering now that for the original line segment both the topological and the Hausdorff dimensions coincide and are equal to $D_T = D_H = 1$, then it follows that for the gaps left representing by definition and construction regions of No space and No time we have a Hausdorff dimension equal to $1 - \phi = \phi^2$ as shown in Fig. 1 [3–6]. To summarize the above thus far we have two things:

- (a) An uncountable infinite number of zero dimensional points with zero measure [2, 7], i.e., zero length and a points set possessing a Hausdorff dimension equal to $D_H = (\sqrt{5} - 1)/2$ [3–7].
- (b) An infinite but countable number of gaps with a complementary Hausdorff dimension equal to $D_H = 1 - \phi = \phi^2 = 0.381966011$ [8–11]. By contrast the measure, i.e., the length of the complementary set is still equal to $1 - 0 = 1$.

In the present short paper will show that by lifting both the ϕ points thin Cantor set and the ϕ^2 collection of gaps, i.e., fat Cantor set to a Kaluza-Klein five dimensional spacetime we can resolve the problem of dark energy [12, 13].

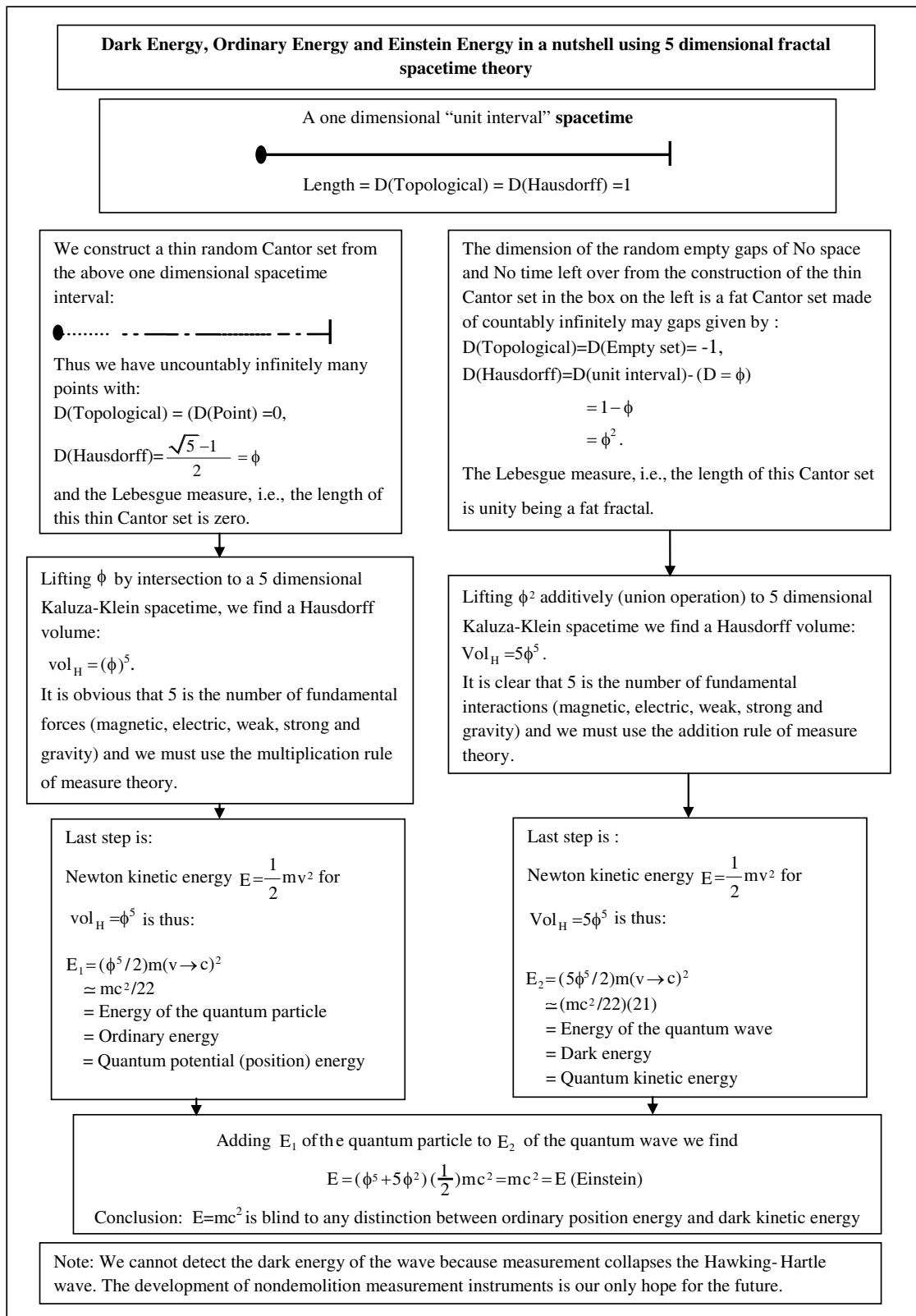


Figure 1: Dark energy as fat (thick) fractal quantum wave energy and ordinary energy as thin fractal quantum particle energy.

2. ANALYSIS

Proceeding in the way indicted in the introduction we encounter two distinct situations:

- i) For the ϕ zero measure thin Cantor zero set we calculate a quasi-Hausdorff hyper volume [8–

11] by extending our familiar area and volume definition to formally five dimensions. That means $\text{vol}_H^{(5)} = \phi\phi\phi\phi\phi = \phi^5$ for a Kaluza-Klein $D = 5$ spacetime as indicated in Fig. 1 [12, 13]. Since the topological probability of finding a ‘‘Cantor point’’ in our set is ϕ , then ϕ^5 can be understood as an application of the intersection rule of sets or alternatively as the multiplication theorem of probabilistic events [2, 6, 7, 8].

- ii) For the ϕ^2 measure 1 fat cantor empty minus one set on the other hand, we have a dual additive quasi-Hausdorff measure representing the total length of the circumference of a pentagon with the length of each side being equal to ϕ^2 . In other words lifting additively the Hausdorff dimension of the complementary set, i.e., $D_H = 1 - \phi = \phi^2$ to $D = 5$ would give us a quasi-hyper volume $\text{vol}_{H(5)} = \phi^2 + \phi^2 + \phi^2 + \phi^2 + \phi^2 = 5\phi^2$ (see Fig. 1). Similar to ϕ , the result may be interpreted as an application of the addition rule of events of probability theory or equivalently as the union rule of sets [2, 6, 7, 8].

Next we look at the magnitude of Newtonian kinetic energy $E = \frac{1}{2}mv^2$ inside $\text{vol}_H^{(5)} = \phi^5$ and $\text{vol}_{H(5)} = 5\phi^2$ when $v \rightarrow c$ where m is mass, v is velocity and c is the speed of light [8, 14]. In a similar manner to the above we find two different types of energies:

- 1) The energy associated with density ϕ^5 is obviously $E = (\phi^5/2)mc^2$. Noting that ϕ^5 is equal to the celebrated Hardy probability for quantum entanglement [8, 15] where ϕ is the Hausdorff dimension for the zero set fractal (i.e., a set of only zero in it) modeling the quantum particle given by $D_0 \equiv (0, \phi)$, then we see that this is the ordinary energy of a quantum particle which will be denoted by $E(O) = (\phi^5/2)mc^2 \simeq mc^2/22$ [14]. This energy density is only 45% of what Einstein’s relativity theory predicts classically and is thus equal to the energy density found from the WMAP measurement and Supernova cosmological data analysis [14, 17]. In fact from $E_1 = \frac{1}{2}mc^2\alpha^2$ where $\alpha = 1/137$ of the ground state of the hydrogen atom and since α^2 is the square of a probability namely a geometrical electromagnetic one, then changing α^2 to $\phi^2\phi^3 = \phi^5$ quantum entanglement probability one finds $E = (\phi^5/2) m (c)^2$.
- 2) By contrast for the energy due to the density $5\phi^2$ associated with pentagonal surface or halo of the $D = 5$ particle connected to a set with nothing in it (i.e., the empty set $(-1, \phi^2)$ by which classical set theory starts or the quantum wave in $D = 5$ we have $E = (5\phi^2/2)mc^2$ [5]. It is easily shown that this is the value of the supposedly missing dark energy of the cosmos [14, 16, 17] and will therefore be denoted by $E(D) = (5\phi^2/2)mc^2 \simeq mc^2(21/22)$. This energy obviously has a different sign to $E(O)$ and therefore produces a form of antigravity force [17] which explains the increased rate of cosmic expansion observed in relatively recent accurate cosmological measurements [17] and attributes it to the negative anticlastic curvature of the compactified 22 bosonic dimensions of spacetime as distinct from the normal 4 large dimensions. (i.e., 3 space and one time dimension fused together relativistically). In addition being the energy of the Hawking-Hartle quantum wave of the universe, it collapse on measurement and that is why we cannot detected dark energy.

From the above we must conclude that when summing up $E(O)$ and $E(D)$ we must obtain $E(\text{Einstein})$ which is easily verified using elementary arithmetic based on the fact that $\phi + \phi^2 = 1$ and $\phi^5 + 5\phi^2 = 2$. Thus we have the wonder which is no wonder namely that the total energy is given by:

$$E(O) + E(D) = \frac{1}{2} (\phi^5 + 5\phi^2) mc^2 = \frac{1}{2} (2)mc^2 = mc^2 = E(\text{Einstein}) \quad (1)$$

Our most important and final conclusion is thus that Einstein $E = mc^2$ is blind to any distinction between ordinary and dark energy The entire analysis is summarized in Fig. 1.

3. CONCLUSION

Dark energy is the negative energy of the quantum surface or the outside of the quantum particle, i.e., the quantum wave while ordinary energy is the energy of the inside core of the quantum wave which is the quantum particle. In Set Theoretical Terminology [5], ordinary energy is the energy of the zero set while dark energy is the complementary energy of the empty set. An equivalent string theoretical explanation leading to the same conclusion is to consider dark energy to be due to the anticlastic curvature caused by the 22 compactified dimensions of the 26 dimensions of bosonic strings’ spacetime theory.

Einstein's energy is thus blind to any distinction between dark energy and ordinary energy density. However our apparatus feels the difference and can register only ordinary energy and that is the explanation for the result of the cosmological measurement be it that only 4.5% of the energy predicted by the Theory of Relativity [17] is present or that the universe is pushed apart rather than being pulled together as we previously presumed and which new accurate measurements have now contradicted [14, 16, 17]. To detect and measure the dark energy of the cosmos the only hope is a future nondemolition instrument which does not collapse the Hawking-Hartle quantum wave of the cosmos.

We note that because the axiomatic structure of set theory and mathematical consistency could not be guaranteed without the introduction of the empty set as well as the zero set, then by the same token it follows that fundamental quantum physics and quantum gravity could not be consistent nor understood except by embracing the vital role of the particle-wave duality and its connection to the zero set-empty set duality. In a nutshell traditional physics does not recognize the empty set and equates it to an absolute nothingness. It is imperative to recognize that physics depends upon logical structure and it should never confuse the zero set with the empty set. It is equally imperative not to confuse either the zero set or the empty set with insubstantial total nothingness.

One should not be entirely surprised that an empty set has a physical effect because a quantum wave which is merely a probability wave devoid of ordinary matter or energy also has a physical effect. It is an elementary fact of quantum physics discovered long ago by Max Born that the square of the probability wave function gives the probability of finding the spatial location of a quantum particle [6]. In a way dark energy discloses the mystery of the quantum wave function and vice versa. That may be a circulatory explanation or worse still tautology. However logical understanding is partially achieved by reducing the number of concepts. At a minimum this is what we have done here

REFERENCES

1. Mandelbrot, B., *The Fractal Geometry of Nature*, Freeman, New York, 1983.
2. Morgan, F., *Geometric Measure Theory*, Elsevier, Amsterdam, 2009.
3. Stakhov, A., *The Mathematics of Harmony*, World Scientific, New Jersey, 2009.
4. El Naschie, M. S., O. E. Rossler, and I. Prigogine, *Quantum Mechanics, Diffusion and Chaotic Fractals*, Pergamon Press, Elsevier, Oxford, UK, 1995, ISBN0080420273.
5. He, J. H. and M. S. El Naschie, "On the monadic nature of quantum gravity as a highly structured golden ring spaces and spectra," *Fractal Spacetime and Non-commutative Geometry in Quantum and High Energy Physics*, Vol. 2, No. 2, 94–98, Asian Academic Publisher Ltd., Hong Kong, 2012, ISSN: 2221-0652.
6. El Naschie, M. S., "A review of E-infinity and the mass spectrum of high energy particle physics," *Chaos, Solitons & Fractals*, Vol. 19, No. 1, 209–236, 2004.
7. El Naschie, M. S., "The theory of cantorion spacetime and high energy particle physics, (an informal review)," *Chaos, Solitons & Fractals*, Vol. 41, No. 5, 2635–2646, 2009.
8. Penrose, R., *The Road to Reality*, Jonathan Cape, London, 2004.
9. Coxeter, H., *The Beauty of Geometry*, Dover Publications, New York, 1999.
10. Coxeter, H., *Regular Polytopes*, Dover Publication, New York, 1973.
11. Bengtsson, I. and K. Zyczkowski, *Geometry of Quantum States*, Cambridge, 2006.
12. Halpern, P., *The Great Beyond, Higher Dimensions, Parallel Universes and the Extraordinary Search for a Theory of Everything*, John Wiley, New Jersey, 2004.
13. El Naschie, M. S., "Kaluza-Klein unification. Some possible extensions," *Chaos, Solitons & Fractals*, Vol. 37, No. 1, 16–22, 2008.
14. El Naschie, M. S. and L. Marek-Crnjac, "Deriving the exact percentage of dark energy using a transfinite version of Nottale's scale relativity," *Int. J. of Modern Nonlinear Theory & Applications*, Vol. 1, 118–124, 2012.
15. El Naschie, M. S., "Entanglement as a consequence of a Cantorian micro-space geometry," *J. of Quant. Info. Sci.*, Vol. 1, 50–53, Published Free Access Online on Sep. 2011, 2011.
16. Amendola, L. and S. Tsujikawa, *Dark Energy Theory and Observations*, Cambridge University Press, Cambridge, 2010.
17. Copeland, E. J., M. Sami, and S. Tsujikawa, *Dynamics of Dark Energy*, Jun. 16, 2006, arxiv: hep-th/0603057V3.

Synthesis and Properties of the LiNbO_3 Thin Films Intended for Nanogradient Structures

R. N. Zhukov, S. V. Ksenich, A. S. Bykov, D. A. Kiselev,
M. D. Malinkovich, and Yu. N. Parkhomenko
National University of Science and Technology “MISIS”
Leninskiy pr. 4, Moscow 119049, Russian Federation

Abstract— In terms of development of the methods of producing of the electro-optic lithium niobate thin films on dielectric layers, experiments on the synthesis and studied the domain structure of the films grown on silicon oxide substrate (111) and (100) orientation. The films were synthesized by RF magnetron sputtering, and the primary method of research — atomic force microscopy. The results show the high degree of grains orientation in polycrystalline structure. From the measured amplitude of the piezoelectric response signal vs applied AC voltage between the cantilever tip and the substrate are calculated values of piezoelectric modules for both orientation the lithium niobate thin films.

1. INTRODUCTION

Among the most interesting and promising directions in the physics of optics and materials science is the study of the synthesis and characterization of thin film gradient optical structures [1]. The results already obtained in this area are of great practical importance. In particular, created by ultra-wide range (in the visible, near and mid-IR) and deep modulation on the beam splitter, reflectorless optical phase shifters, compressors femtosecond optical pulses that work in non-prism mode transient tunneling narrow-band filters in the visible and near — infrared regions of the spectrum etc. [2].

As the next step in the development of this area is important on the basis of gradient thin-film coatings to develop a new class of optical elements — gradient optical “Transformers”, i.e., the gradient of thin-film multilayer structures, the characteristics of which may vary considerably during their work by changing the optical thickness of a certain the number of sub-elements. For example, the reflection of a certain wavelength range will change his passing, etc.. This will greatly simplify the optical circuit devices, in some cases, remove them multiple channels to reduce the weight and dimensions. To carry out this possible by creating layers in multilayer structures of electro-optic materials, such as lithium niobate. However, this procedure requires the development of synthesis of dielectric substrates oriented both crystallographic direction and the polarization vector by fine film of lithium niobate to 100 nm thick.

Oriented polycrystalline lithium niobate (LiNbO_3 , LN) thin film are also of interest for a variety of electro-optical and acousto-optical applications, including integrated device structure containing micro- and opto-electronic components. The ferroelectric properties of LN films were investigated by piezoresponse force microscopy (PFM), which allows to identify the polarization direction and the local domain distribution [3]. Previously [4–6], this method has already been used to characterize of the polarization domain state in films deposited on conductive silicon substrates.

In order to directly address the domain structures and local piezoelectric properties of LiNbO_3 thin films we apply the PFM method which has been successfully used for the investigation the local piezoresponse properties in this films on the nanoscale.

2. EXPERIMENTAL DETAILS

The studied LiNbO_3 films thickness 100 nm were deposited by RF magnetrom sputtering of the single-crystalline target in $\text{Ar}/\text{O} = 1$ atmosphere (0.6 Pa) on n-type Si (100) and (111) substrates ($\rho = 2 \Omega \cdot \text{cm}$) with SiO_2 buffer layer. The subsequent thermal annealing of the obtained structures has been done in air at 600°C . Atomic force microscopy measurements indicate that the surface roughness of the LiNbO_3 thin films was 4–10 nm, which meets the demands for practical wave guiding devices. The ferroelectric properties have been studied by visualization of the as-growth domain structure recording induced ferroelectric states and the hysteresis loops by piezoresponse force microscopy (PFM) using Scanning probe laboratory NTEGRA-Prima (NT MDT, Russia). Out-of-plane PFM images of the samples were obtained by applying AC voltage (5 V, peak-to-peak) with a frequency of 150 kHz. Piezoresponse versus DC voltage [effective d_{zz} (V)] were measured in the pulse DC mode, so-called remnant piezoloop.

3. RESULTS AND DISCUSSIONS

Figures 1(a), (d) shown surface images LiNbO₃ films 100 nm thick synthesized on silicon substrates (100) and (111) orientations. The value of surface roughness was about 6 nm for LiNbO₃/SiO₂/Si (100), and 8 nm for LiNbO₃/SiO₂/Si (111).

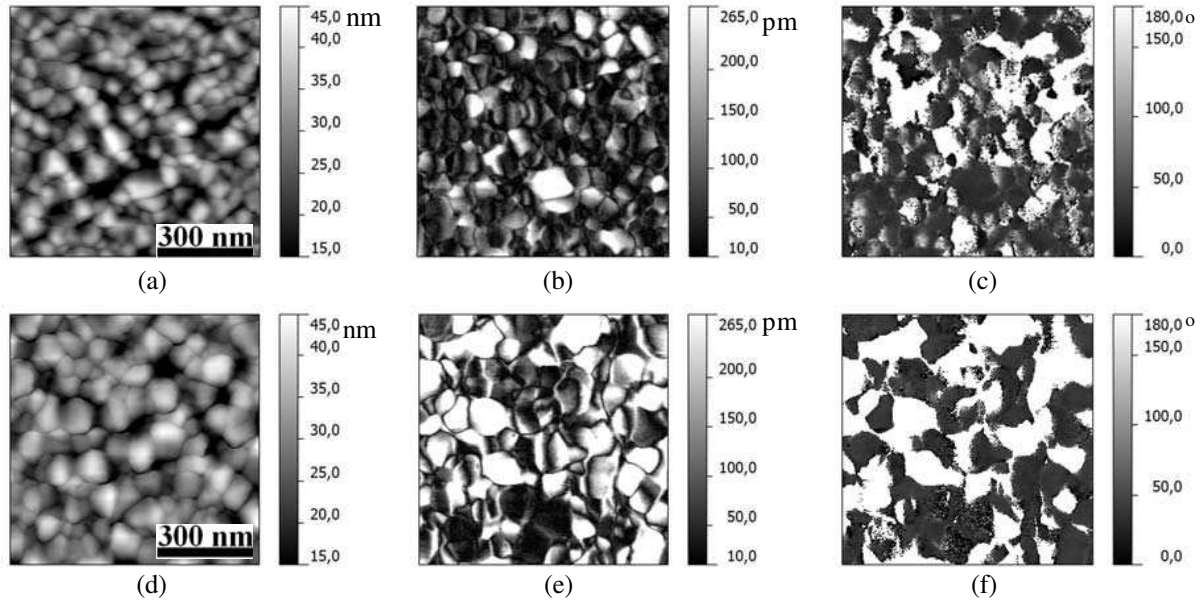


Figure 1: (a), (d) Topography, (b), (e) amplitude and (c), (f) phase images for LN thin films, synthesized on (100) and (111) silicon substrates, respectively.

Lateral relief correspondence grain sub-structure films with an average grain size of 60 nm on the Si (100) and 80 nm on Si (111) substrates. Along with the topographic image obtained independently amplitude (Figs. 1 (b), (e)) and phase images (Figs. 1(c), (f)) the vertical component of the piezoelectric response. In this case piezoelectric response amplitude values correspond to a deformation of the sample to an alternating electric field. From these phase images shows that the grains have a “bright” and “dark” contrast that corresponds to a particular orientation of the polarization vector in the studied films.

The dependence of the amplitude of the local deformation of the LN thin film under the influence of an electric field created by the voltage applied between the cantilever and the substrate (inverse piezoelectric effect) for both orientations. Figs. 2(a)–(c) are piezoelectric response amplitude ratio of the signal on the same scale by an alternating voltage amplitude of 3 V, 7 V and 11 V, respectively. It is clear that contrast of the piezoresponse amplitude increase with increasing magnitude of the AC voltage applied to the conductive cantilever.

Figure 2(d) shows a histogram distribution of the signal amplitude for the PFM image obtained by registering the variable voltage with 1–11 V. The average amplitude of the piezoelectric response signal is determined from the histogram, is a measure of the deformation induced by the applied voltage of the film. Fig. 3 shows the dependence of the mean amplitude of the piezoelectric response of the applied voltage for the films grown on substrates with different crystallographic orientations. As expected, the piezoresponse amplitude increases linearly with an alternating voltage from 1 to 11 V.

Approximating a linear function of the resulting dependence on the slope of the curve to determine the value of piezoelectric module which is of formula [7]:

$$d_{zz} = \frac{A}{V_{AC}}, \quad (1)$$

where A — piezoresponse amplitude, V_{AC} — AC voltage amplitude.

Thus, we have experimentally that the value for LiNbO₃/SiO₂/Si (111) heterostructures was 19 pm/V, and for LiNbO₃/SiO₂/Si (100) is slightly less — 16 pm/V. These values are in good agreement with literature data for lithium niobate single crystal [8], and nanocrystalline particles based on LiNbO₃ [9].

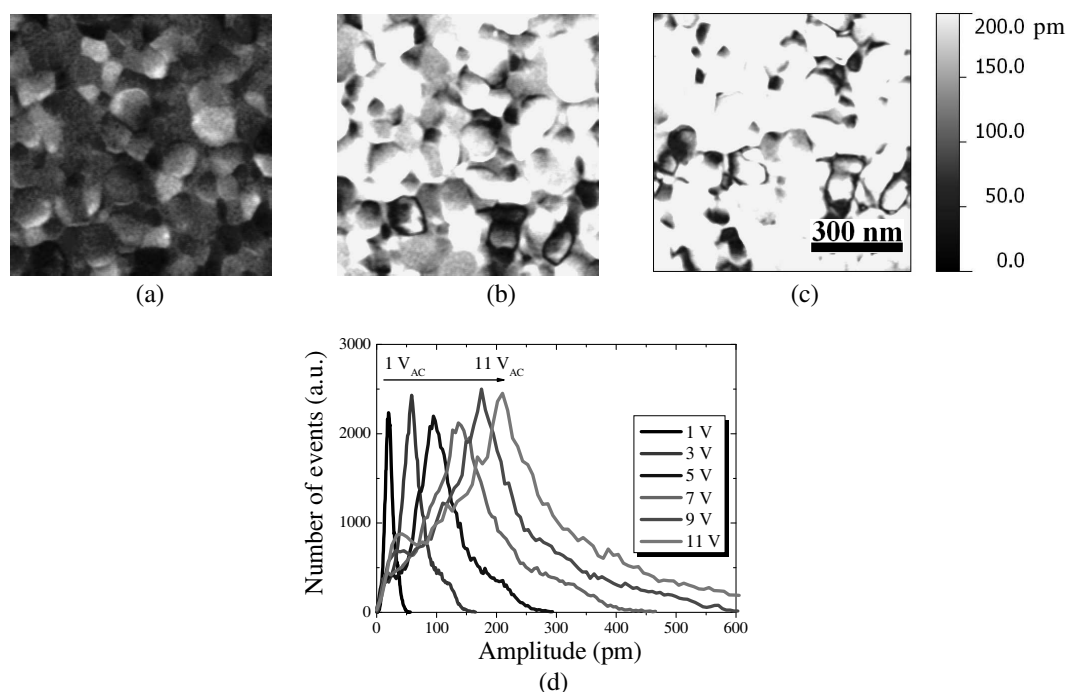


Figure 2: Out-of-plane PFM amplitude images taken at (a) 3 V, (b) 7 V and (c) 11 V, respectively, for LiNbO_3 thin films, and (d) corresponding histograms for PFM amplitude signals.

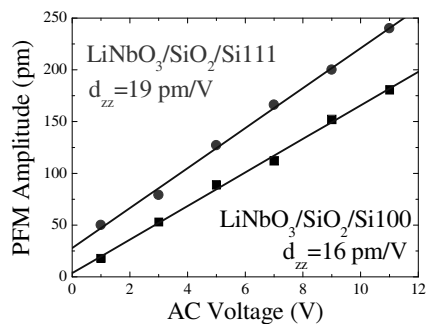


Figure 3: Average piezoresponse from PFM amplitude images obtained at each voltage step versus AC drive voltage for the LN thin films.

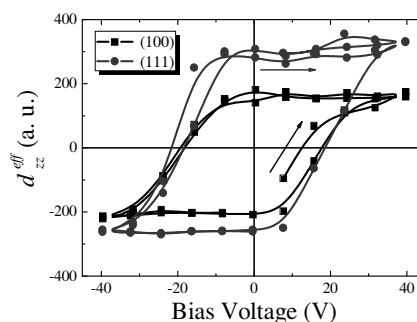


Figure 4: REM piezoelectric hysteresis loops for LiNbO_3 films vs Si substrate orientations.

To confirm the presence of differences in the values of piezoelectric modules for films of lithium niobate were synthesized by varying the orientation of the substrate, we obtained the piezoelectric hysteresis loop (Fig. 4) in the mode of remnant piezoelectric response (REM hysteresis) for the two orientations of the substrate.

Measurements were carried out on the free surface of the film after removal of the applied voltage (the time of electrical stimulation — 1 s exposure time after removal of the stress — 1), the beginning of the directions indicated by the arrows bypass. The maximum voltage applied to the sample, ± 40 V. It is clear that the remnant piezoelectric response, measured after removal of the polarizing voltage for LN films synthesized on Si (111) more than for the films synthesized on Si (100) orientation.

4. CONCLUSIONS

The piezoelectric properties of the films thus depend strongly on the crystallographic orientation of the silicon substrate with SiO_2 buffer layers. Characterized in particular, the values of piezoelectric modules which are 19 pm/V and 16 pm/V for structures $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$ (111) and $\text{LiNbO}_3/\text{SiO}_2/\text{Si}$ (100), respectively. The results indicate that one of the problems the solution of which is necessary for the creation of electro-optic lithium niobate thin films on dielectric layers, namely — high crystallographic orientation of the structure can be solved by means of RF

sputtering technology.

ACKNOWLEDGMENT

The research was financially supported by the Ministry of Education and Science of the Russian Federation using the equipment of Common-Use Scientific Center “Material Science and Metallurgy” at NUST “MISiS”. DAK acknowledges German Academic Exchange Service (DAAD project A/12/00679) for financial support during his research stay at the University of Duisburg-Essen.

REFERENCES

1. Volpian, O. D. and A. I. Kuzmichev, “Negative refraction of waves. Keeping to the physics and technology of electromagnetic metamaterials,” Avers, Kiev, 2012.
2. Volpian, O. D., S. I. Shkatula, and M. D. Malinkovich, “The defense industry — Scientific and technical progress Russia,” No. 4, 22, 2012.
3. Bornand, V., B. Gautier, and P. Papet, “Growth and nanoscale ferroelectric investigation of radio frequency-sputtered LiNbO_3 thin films,” *Mater. Chem. and Phys.*, Vol. 86, 340–346, 2004.
4. Zhukov, R. N., A. S. Bykov, D. A. Kiselev, M. D. Malinkovich, and Y. N. Parkhomenko, “Piezoelectric properties and surface potential behavior in LiNbO_3 thin films grown by the radio frequency magnetron sputtering,” *J. Alloys Comp.*, <http://dx.doi.org/10.1016/j.jallcom.2013.01.116>, 2013.
5. Zhukov, R. N., D. A. Kiselev, M. D. Malinkovich, Y. N. Parkhomenko, E. A. Vygovskaya, and O. V. Toropova, “Propagation of polarization of ferroelectric grains in electrically isolated lithium niobate films,” *Russian Microelectronics*, Vol. 41, No. 8, 459–463, 2012.
6. Kiselev, D. A., R. N. Zhukov, A. S. Bykov, M. D. Malinkovich, and Y. N. Parkhomenko, “Growth and investigation of LiNbO_3 thin films at nanoscale by scanning force microscopy,” *PIERS Proceedings*, 520–523, Moscow, Russia, Aug. 19–23, 2012.
7. Pintilie, L., C. Dragoi, Y. H. Chu, L. W. Martin, R. Ramesh, and M. Alexe, “Orientation-dependent potential barriers in case of epitaxial $\text{Pt-BiFeO}_3\text{-SrRuO}_3$ capacitors,” *Appl. Phys. Lett.*, Vol. 94, 232902–232904, 2009.
8. Shaskol'skaya, M. P., *Acoustic Crystals*, Nauka, Moscow, 1982.
9. Mohanty, D., G. S. Chaubey, A. Yourdkhani, S. Adireddy, G. Caruntu, and J. B. Wiley, “Synthesis and piezoelectric response of cubic and spherical LiNbO_3 nanocrystals,” *RSC Advances*, Vol. 2, 1913–1916, 2012.

Quantum Blooming: The Influence of the Detuning of the Antireflective Wells and the Phenomenon of Nonsymmetric Transmission

A. Kaklyugin

Atmospheric Plasma Instant Technology Corp., Hangenbieten 67980, France

Abstract— We formulate a simple model of tunneling allowing analytical solution. It is shown that there is a possibility of the potential barrier blooming by supplementing it with additional wells, resulting in the barrier is clear: there are no reflected waves and the passing wave propagates without attenuation. We consider the possibility of blooming detuning of the quantum wells. The dependences of the reflection and transmission coefficients are found as a function of the degree of detuning. It is shown that asymmetric detuned barrier loses the property of reciprocity: the properties of the barrier transparency are significantly dependent on the direction of propagation.

1. INTRODUCTION

The classical concept of tunneling with an exponential decrease of the amplitude of the penetrated wave and the appearance of the reflected wave adopted in quantum mechanics is to be completed by the possibility of quantum blooming, providing additional singular quantum wells. It is shown that such an addition to the classical inaccessible barrier allows dampingless traverse through the barrier [1].

The sub-barrier penetration researches inspired by the success achieved in the understanding of the effects of similar tunnel propagation of electromagnetic and acoustic waves in the gradient media without reflection and attenuation [2–4]. In this work we will continue to explore the possibility of passing the classically inaccessible dampingless potential barriers by referring to a more realistic model.

2. FORMULATION OF THE MODEL

A simple example has been formulated and solved in analytical form [1] for the model problem of the potential barrier blooming.

There consider a simplest barrier U_0 and a plane wave propagating along the half-axis $x > 0$. The potential energy is $U(x) = U_0$ when $a > x > 0$ and $U(x) = 0$ for $x > a$. If we measure energy in units of $\hbar^2/2ma^2$ and the distances in units of a , the wave function of the stationary state will satisfy the Schrödinger equation that the energy, length and wave vectors will be dimensionless quantities:

$$\psi''(x) + (\varepsilon - q^2(x)) \psi(x) = 0; \quad (1)$$

where $q^2(x) = (2ma^2/\hbar^2)U(x)$. The dimensionless value $q_0^2 = (2ma^2/\hbar^2)U_0$ characterizes to what extent the barrier or well being big or small.

The solution of Equation (1) is a set of exponential functions

$$\psi(x) = \exp\{i(kx + \varphi)\}, \quad x > 1 \quad (2a)$$

$$\psi(x) = c \exp(\kappa x) + b \exp(-\kappa x), \quad 1 > x > 0. \quad (2b)$$

We do not consider the reflected wave to find the conditions for complete passage through the potential barrier. The values of ε and q_0^2 link to the squares of the wave vectors $k^2 + \kappa^2 = q_0^2$; $q^2 = \varepsilon$. The model consider a solution of Equation (1) whose real and imaginary units will be even and odd respectively. The complex coefficients c and b in (2b) override in the real values A and B such a way

$$\psi(x) = A \operatorname{ch}(\kappa x) + i B \operatorname{sh}(\kappa x), \quad 1 > x > 0. \quad (2c)$$

Of continuity of $\psi(x)$ at $x = 1$ (real and imaginary units separately) write down explicitly the values A and B :

$$A = \cos(k + \varphi)/\operatorname{ch} \kappa; \quad B = \sin(k + \varphi)/\operatorname{sh} \kappa. \quad (3)$$

The value φ is defined by

$$k[\operatorname{tg}(k + \varphi) + \operatorname{ctg}(k + \varphi)] = \kappa[\operatorname{cth} \kappa - \operatorname{th} \kappa]. \quad (4)$$

For to change the shape of the barrier at the boundary at the point $x = 1$ (and, correspondingly, by the symmetry of the problem, at the point $x = -1$). In a certain sense this decision possesses the analogy with optical clearing, as well as dampingless propagation of electromagnetic waves through the barrier with a gradient profile [2–4]. The barrier with a certain conventionality seems qualitatively like the barrier with singular elements at the border).

The wave function kink on the barrier boundary (at $x = 1$) is to be solved by adding a potential additive $C\delta(x - 1)$ [1].

We consider the situations of the small $q_0 \leq 1/2$ and large $q_0 > 1/2$ wells. It is shown that in the first case, the blooming additives exist for all values of the wave vector. In the second case there is a limitation $\kappa \geq D(q_0)$. The function $D(x)$ is the inverse function of $d(\kappa) = (\kappa/2)$ ($\operatorname{th} \kappa + \operatorname{cth} \kappa$); this condition is the reverse of the condition $\kappa \operatorname{cth} 2\kappa \geq q_0$.

The amplitude of the anti-reflective additives is given by:

$$C = -\kappa \operatorname{cth} 2\kappa \pm (\kappa^2 \operatorname{cth}^2 2\kappa - q_0^2)^{1/2} \quad (5)$$

3. DETUNNING

Let's consider the situation where the blooming is inaccurate. Suppose there exists some detuning: we shall set antireflection additives in the form $(C + \alpha)\delta(x - 1)$ and, respectively, $(C + \beta)\delta(x + 1)$.

These values define the gap of derivative of the wave function at the points $x = \pm 1$ and, together with the terms of the continuity of the system will make it possible to write the equations defining all the waves: the incident, reflected passing.

The wave function keeping the conservation laws can be written in the most general form as

$$\psi_{\rightarrow, \leftarrow}(x) = (1 + ei) \exp\{i(kx - \varphi + \delta)\} + e \exp\{-i(kx - \varphi + \nu)\}. \quad (6)$$

There are actual unknown values e , δ and ν , which should appeal to zero without detuning. Indeed, in this case the flux of incident and reflected particles will respectively, $k(1 + e^2)$ and ke^2 .

We introduce an additional parameter $\varepsilon = \nu - \pi/2$. The system of equations can be written as

$$\begin{aligned} \exp[-i(k + \varphi)][(\exp(i\delta) - 1) + 2e \exp\{i(\delta - \varepsilon)/2\} \sin[k + \varphi - (\delta + \varepsilon)/2]] &= (\alpha/\kappa) \operatorname{sh} 2\kappa \exp[i(k + \varphi)] \quad (7) \\ \{\beta + ki[(\exp(i\delta) - 1)] \exp[-i(k + \varphi)] + \alpha \exp[i(k + \varphi)]\}[(C + \beta)/\kappa] \operatorname{sh} 2\kappa + \operatorname{ch} 2\kappa] \\ &= 2ek \exp\{i(\delta - \varepsilon)/2\} \cos[k + \varphi - (\delta + \varepsilon)/2] \quad (8) \end{aligned}$$

4. RESULTS. DETUNING AT THE BEGINNING AND AFTER THE BARRIER, REFLECTION AND TRANSMISSION COEFFICIENTS

The system is asymmetric with respect to the replacement $\alpha \leftrightarrow \beta$. It is not simple enough for analysis, primarily because of the complex phase factors. We restrict ourselves to the simplest case: first we consider the detuning at the beginning of the barrier $\alpha = 0$, and in the second case we consider the detuning after barrier $\beta = 0$.

4.1.

$\alpha = 0$, we multiply the first equation by $-ik$ and add to the second, we shall obtain

$$\beta \exp[-i(k + \phi)] = 2ek \exp[i(k + \varphi - \varepsilon)]. \quad (9)$$

Because the value e is real, we can choose for ε any value satisfactory $\varepsilon = 2(k + \varphi) + \pi n$. The value of e is determined by the simple expression

$$e = \pm \beta/2k, \quad (10)$$

The sign is determined by the parity of n .

4.2.

In the case of $\beta = 0$ having done the same steps we get:

$$\alpha \exp[i(k + \varphi)] [(C + ik) \operatorname{sh} 2\kappa/\kappa + \operatorname{ch} 2\kappa] = 2ek \exp[i(k + \varphi - \varepsilon)] \quad (11)$$

In this case the phase relationship is somewhat more complicated:

$$\varepsilon = -\operatorname{Arctg}[k \operatorname{sh} 2\kappa / (C \operatorname{sh} 2\kappa + \kappa \operatorname{sh} 2\kappa)] \quad (12)$$

and the value e is given by

$$e = \pm(\alpha/2k) [(C^2 + k^2) \operatorname{sh}^2 2\kappa/\kappa^2 + C \operatorname{sh} 4\kappa/\kappa + \operatorname{ch}^2 2\kappa]^{1/2}. \quad (13)$$

5. DISCUSSION

Thus, it was found that detuning of the antireflective quantum wells provides in any considered cases a Lorentz contour transmittance (the transmittance is $1/(1 + e^2)$) and, respectively, the reflection coefficient is e^2 .

The resulting coefficients are significantly different for detuning the well on which the wave falling (the case of $\alpha = 0$), and when the well after barrier is detuned (case $\beta = 0$). The first case is trivial and the expression (10) can be easily obtained from the solution of the problem of quantum scattering on the δ -shaped potential.

The second case is more complicated: after passing the barrier the wave function amplitude is decreases and even a significant perturbation has low value, so detuning antireflective well disposed after the barrier (in the direction of passage of the wave) can have a significantly greater value. As far as — it depends on the very blooming well.

The lack of reciprocity turns out unexpected at first glance: the different detuning of the antireflective wells leads to a substantial change in the conditions of transmission and reflection. However, it is clear that we are dealing with a fundamental symmetry breaking: in one case the well on which the wave falls is detuned, in the second the wells which the wave interacts with after passing. It is easy to see that barrier non-reciprocity weakens under the constriction.

The amount of change of the wave packet is also significantly different. In the first case the width of the transmission is determined by the degree of detuning and own wave vector, in the second case it's determined so by the properties of the barrier.

6. PROBLEMS TO POSE

The first problem to arise: how detuning effect on both “corners” of the barrier additives? What is their additive effects?

We can put a problem that may have a practical solution: do the detuning exist when despite the total ban (complete reflection of waves) in one direction, there will exist a noticeable transmission on the other direction? This problem has not been solved (not found such a statement of the problem in which you can get a meaningful result).

Another important question is how important is the idealization of blooming additives in the form of a δ -shaped potentials. Such a problem is far from a reasonable setting.

ACKNOWLEDGMENT

Author is sincerely grateful to A. B. Shvartsburg for posing the problem and discussion.

REFERENCES

1. Kaklyugin, A. S., “Quantum blooming: The possibility of passing through the classically inaccessible area without attenuation,” *PIERS Proceedings*, 513–519, Moscow, Russia, Aug. 19–23, 2012.
2. Shvartsburg, A. B., V. Kuzmiak, and G. Petite, “Optics of subwavelength gradient nanoflms,” *Phys. Rep.*, Vol. 452, 33–88, 2007.
3. Shvartsburg, A. B. and N. S. Erokhin, “Acoustic gradient barriers (exactly solvable models),” *Phys. Usp.*, Vol. 54, 605–623, 2011.
4. Shvartsburg, A. B., G. Petite, and M. Zuev, “Tunneling of femtosecond pulses through nanophotonic barriers: Superluminal precursors,” *J. Opt. Soc. Am. B*, Vol. 28, No. 9, 2271–2276, 2011.

Magnetolectric Current Sensor

R. V. Petrov, I. N. Solovyev, A. N. Soloviev, and M. I. Bichurin
Novgorod State University, Russian Federation

Abstract— This paper presents the construction and operating principle of ME current sensors and its sensitivity investigation. Sensors presented here were designed for detecting AC and DC leakage currents. It is found that the sensor sensitivity depends on design and material parameters of the ME composite and some conditions such as bias magnetic field and the sensor way. It is defined that the additional bias magnetic field of the ME current sensor should be 10 Oe to obtain high sensitivity and linearity of the sensor. The sensitivity of the sensor is 340 mV/A.

1. INTRODUCTION

This application note is dedicated to current sensing using ME current sensors. Electric current is an important physical quantity and its measurement is required in many applications, be it in industrial, automotive or household fields. Different technical solutions to measure currents are known and are found on the market.

For the safety, performance, and feedback control of the power supplies and electric circuits it is often necessary for the circuit to better understand its load current. Protecting the supply circuits, for instance, against overloading and short-circuit conditions requires the system to self-monitor its current delivery. The tasks can be performed by the ME current sensors.

As we know, when a current flows through a conductor, it creates a magnetic field around the conductor. The magnetic field is direct proportional to the current level.

The ME current sensor uses the ME effect as a basis of its measurements. The ME effect is a polarization response to an applied magnetic field, or conversely a magnetization response to an applied electric field. ME behaviour exists as a composite effect in multiphase systems of piezoelectric and magnetostrictive materials [1]. Magnetostrictive-piezoelectric laminate composites have much higher ME coefficients than that of single-phase ME materials or particulate composites. In a magnetostrictive-piezoelectric layered structure the interaction between magnetic and electric subsystems occurs through mechanical deformation. It means that the ME effect is much stronger at frequencies corresponding to elastic oscillations called resonance frequencies.

In current sensor applications the induced ME voltage coefficient is more important than the induced ME electric field coefficient, as voltage is the physical quantity measured.

The sensor is designed for detecting AC and DC leakage currents in electrical circuits.

The ME current sensors can be applied in many types of sensing devices and perform a function as a magnetic field sensor with some design changes [2, 3]. The sensors can detect currents and magnetic fields according to the purpose. If the quantity (parameter) to be sensed incorporates or can incorporate a magnetic field the ME current sensor will perform the task.

2. ME CURRENT SENSOR

2.1. Construction

The ME current sensor shown in Figure 1 is a system consisting of the ME composite, a generator, a rectifier, a permanent magnet, two coils placed into one another and a body.

The ME composite is a layered structure [2] which includes a thin piezoceramic plate of PZT ($0.9[\text{Pb}(\text{Zr}_{0.52}\text{Ti}_{0.48})\text{O}_3] - 0.1[\text{Pb}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3 + 3 \text{ mol\%MnO}_2]$) placed between two magnetostrictive layers of Metglas (FeBSiC) performing a function of electrodes. The final dimensions of the composite were $6 \times 1 \times 0.62$ mm with the PZT concentration of 0.8. The PZT layer was poled in an electric field in the thickness direction. The ME laminate operates in a transverse mode in which applied magnetic fields (H_{\sim}, H_0) are in the plane direction of the laminate.

The generator system comprises a generator and a coil. The generator having the frequency coincided with a resonance frequency of the ME composite is designed. The coil is wound on the composite and designed to induce an alternative magnetic field.

The control system is intended to provide a bias magnetic field and control an output voltage.

A rectifier converts an alternating current, which periodically reverses direction, to a direct current, which flows in only one direction. The rectifier may contain a diode bridge.

The sensor can detect AC and DC currents having a range of 0.01 A to 100 A depending on the conductor parameters.

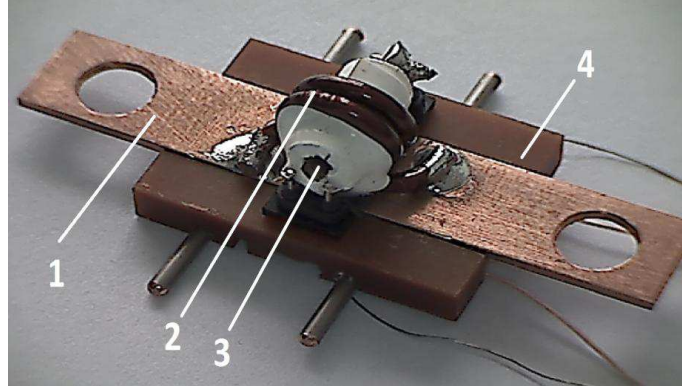


Figure 1: The ME current sensor sample (1 — a conductor with the detected current, 2 — two coil connected to the conductor and a generator, respectively, 3 — the ME composite, 4 — a body).

3. THEORY

An electric current generates a magnetic field around a conductor. The direction can be determined with the “right hand rule”. The field strength H , given in A/m, is directly proportional to the current I and decreases linearly with higher distances r according to

$$H = \frac{I}{2 \cdot \pi \cdot r}. \quad (1)$$

Our design is a long, rectangular-shaped, 3-layer laminate of Metglas/PZT/Metglas. Layers of the laminate are coupled to each other by stress-strain.

We consider small-amplitude oscillations of a layered composite which is formed by magnetostrictive and piezoelectric layers. A bias magnetic field is applied to the magnetostrictive component to obtain a piezomagnetic interaction. The thickness of the sample is assumed to be small compared to other dimensions and its width small compared to its length. In that case, we can consider only one component of strain and stress.

The frequency dependence for transverse ME voltage coefficients of the composite have obtained by solving electrostatic, magnetostatic and elasto-dynamic equations with the corresponding boundary conditions [4]. The ME voltage coefficient equation is defined as

$$\alpha_E = \frac{E}{H} = \frac{A \cdot {}^p d_{31} \cdot {}^p Y^E \cdot {}^m q_{11}}{\Delta \cdot {}^p \varepsilon_{33}} \cdot (2 \cdot r_4 \cdot r_1 + 2 \cdot r_2 \cdot r_3), \quad (2)$$

where ${}^p Y$ and ${}^m Y$ are the piezoelectric and piezomagnetic component modules of elasticity at a constant electric field E and magnetic field H , respectively, ${}^p d_{31}$ and ${}^m q_{11}$ are the piezoelectric and piezomagnetic coefficients at constant stress, ${}^p \varepsilon_{33}$ is permittivity of the piezoelectric component, k is the wave number, L is the length of the ME composite, $r_1 = \cosh(kL)$, $r_2 = \sinh(kL)$, $r_3 = \cos(kL)$ and $r_4 = \sin(kL)$,

$$\Delta = (r_1^2 + 2 \cdot r_1 \cdot r_3 + r_3^2 - r_2^2 + r_4^2) \cdot k \cdot L + (2 \cdot r_4 \cdot r_1 + 2 \cdot r_4 \cdot r_1) \cdot B, \quad (3)$$

$$A = \frac{{}^m Y^H \cdot {}^m t \cdot (2 \cdot z_0 + {}^m t) \cdot (2 \cdot z_0 - {}^p t)}{4 \cdot D \cdot (1 - {}^p K_{31}^2)}, \quad (4)$$

$$B = \frac{{}^p K_{31}^2 \cdot {}^p Y^E \cdot ((z_0 - {}^p t)^3 - z_0^3)}{3 \cdot D \cdot (1 - {}^p K_{31}^2)} + \frac{{}^m K_{31}^2 \cdot {}^m Y^H \cdot ((z_0 + {}^m t)^3 - z_0^3)}{3 \cdot D}, \quad (5)$$

where D is a cylindrical stiffness of the composite, ${}^p t$ and ${}^m t$ are thicknesses of piezoelectric and magnetic layers, z_0 is the distance of middle plane from the composite interface, ${}^m \mu_{33}$ is permeability of the piezomagnetic component.

4. SENSITIVITY

ME current sensors were designed and fabricated. The sensor works on the electromechanical resonance at the resonance frequency of 176 kHz. The sensors sensitivity and ME voltage dependence were researched using the designed measurement setup.

The measurement setup includes two power supplies APS-7315 (Aktakom), multimeter HM 8112-3 (HAMEG instr.), oscilloscope ACIP-4226-3 (Aktakom) and an electromagnet.

The induced ME voltage across the PZT layer was measured as a function of the bias magnetic field (H_0) and the input conductor current (I) at a measurement frequency of 176 kHz. The current flowed through the test conductor and had a range of 0.01 A to 5 A.

Figure 2 shows the output ME voltage of the ME current sensor as a function of detecting input current under various DC magnetic bias.

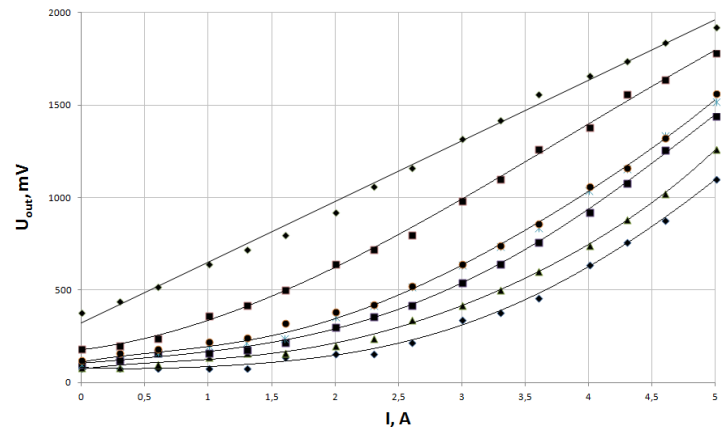


Figure 2: Output ME voltage of the ME current sensor as a function of the conductor current and an additional bias magnetic field from a permanent magnet. With increasing the number of curves the additional bias magnetic field increases.

The results in Figure 2 show that the output ME voltage rate of increase and ME sensor linearity vary in different DC magnetic bias. The maximum linearity is reached at the additional bias field of 10 Oe. Also it is known that the sensitivity is the slope of the linear approximation. According to the data shown in Figure 2 a tangent of the angle ($\text{tg}\alpha$) has been determined. The ME current sensor sensitivity versus a bias magnetic field and the detected current has been obtained.

Figure 3 shows the ME current sensor sensitivity ($\text{tg}\alpha$) dependence on the bias magnetic field.

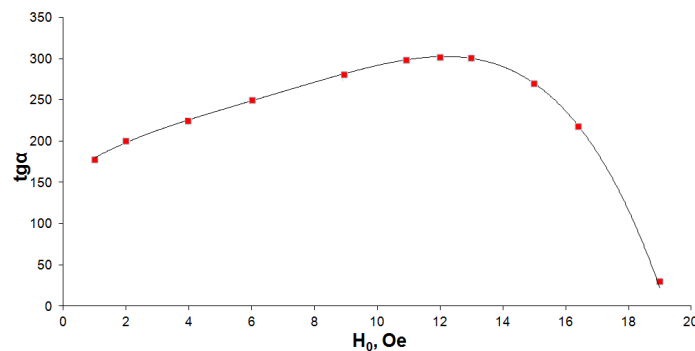


Figure 3: The ME current sensor sensitivity as a function of bias magnetic field.

In accordance with results shown in Figure 3 to obtain the greatest sensitivity and linearity of the ME current sensor it is necessary to select the operating point with the optimal bias magnetic field. It is clear that the additional bias magnetic field of the ME current sensor should be 10 Oe (6th curve in Figure 3). Then a control system should contain a permanent magnet with $H = 10$ Oe. The sensitivity of the sensor is 340 mV/A and close to Hall current sensor (CSA-1V of Sentron) sensitivity.

5. CONCLUSION

A theoretical model has been developed for the resonance ME effect in layered magnetostrictive-piezoelectric composites for the ME current sensor application. The frequency dependence for the

transverse ME voltage coefficient have been obtained using the simultaneous solution of electrostatic, magnetostatic and elastodynamics equations.

An output ME voltage of the ME current sensor as a function of the detecting input current under various DC magnetic bias has been obtained. The sensors sensitivity dependence has been obtained and analyzed. The sensor sensitivity depends on constructive and material parameters of ME composite and a bias magnetic field. The operating point for the ME current sensor has been selected.

It turns out that due to galvanic isolation between the sensed circuit and the measuring circuitry, current sensing using ME effect is a good choice for many applications.

The further improvement of the sensors is using push-pull schemes and the new piezoelectric material named the piezofiber. Other materials with the higher characteristics can be used in ME sensors. For increasing the sensor sensitivity one needs to use the large magnetostrictive and high strain-sensitive piezoelectric ME composite materials.

ACKNOWLEDGMENT

The reported study was partially supported by the grant of the Federal Target Program “Scientific and pedagogical staff of innovative Russia” on 2009-2013 and by RFBR research project #13-02-98801.

REFERENCES

1. Bichurin, M. I. and D. Viehland, Eds., *Magnetoelectricity in Composites*, 273, Pan Stanford Publishing, Singapore, 2012.
2. Soloviev, I. N., M. I. Bichurin, and R. V. Petrov, “Magnetoelectric magnetic field sensors,” *PIERS Proceedings*, 1359–1362, Moscow, Russia, Aug. 19–23, 2012.
3. Bichurin, M. I., V. M. Petrov, R. V. Petrov, Y. V. Kiliba, F. I. Bukashev, Y. V. Smirnov, and D. N. Eliseev, “Magnetoelectric sensor of magnetic field,” *Proceedings of the Fourth International Conference on Magnetoelectric Interaction Phenomena in Crystals (MEIPIC-4), Ferroelectrics*, Vol. 280, 199, 2002.
4. Bichurin, M. I. and V. M. Petrov, “Modeling of magnetoelectric interaction in magnetostrictive-piezoelectric composites,” *Advances in Condensed Matter Physics*, Vol. 2012, 2012.

Microwave Power Absorption in Human Body for Non-invasive Glucose Monitoring

A. Elkady¹, M. El-Hadidy², A. Medhat², A. Khorshid¹, and A. Darwish¹

¹American University in Cairo, Egypt

²CST-Middle East, Egypt

Abstract— The main objective of this contribution is to explore the absorption of electromagnetic microwaves in different human body regions for non-invasive glucose monitoring (NGM). Recently, the use of dielectric measurements for the purpose of NGM has been an area of much interest. However, currently no existing report can explore the power dissipation of electromagnetic waves for the NGM. This work is based on established empirical models of human tissue and standard human body models of the Electromagnetic (EM) simulation tools, such as CST Voxel Model. The Specific Absorption Rate (SAR) is investigated for various frequencies and different regions of the human body to estimate the optimum real-world scenarios for non-invasive glucose monitoring (NGM). Empirical models of dielectric properties for estimating the blood glucose concentrations are applied at frequency range from 1 GHz to 10 GHz and blood glucose concentration of hyperglycemia (300 mg/dl), hypoglycemia (40 mg/dl) and euglycemic (100 mg/dl). A comparative evaluation of SAR measurements for blood in free space and blood in body's vessels is presented in context of investigating the effect of power dissipation for dielectric NGM measurements.

1. INTRODUCTION

Diabetes mellitus is a disorder of the metabolic homeostasis controlled by insulin, resulting abnormalities of carbohydrate and lipid metabolism. According to the International Diabetes Federation, the current diabetic population of 366 million diabetics is estimated to increase up to 552 million by 2030, with the Middle East hosting the fastest growth [1]. Diabetic patients require frequent self-monitoring of blood glucose (SMBG), which is conventionally performed via painful finger-pricking using electrochemical glucose sensors. To overcome the trauma of invasive SMBG, non-invasive glucose monitoring (NGM) devices externally sense glucose without perturbing an external membrane or sampling a body fluid [2].

The feasibility of dielectric measurements for NGM has been demonstrated by Feldman et al., whom demonstrated that the specific cell capacitance of erythrocytes is influenced by changes in the concentration of D-glucose within physiologic range [3, 4]. Building on Feldman's work, Caduff et al. developed the first NGM device based on impedance changes of the skin and underlying tissue [5]. Other attempts have also demonstrated dielectric NGM devices [6–10], however a successful commercial dielectric NGM device is yet to be developed. A recurrent problem in all dielectric NGM attempts was to translate successful results under controlled conditions in the laboratory to real-life clinical scenarios [11].

Despite the established correlation between blood glucose and blood dielectric properties, few models have been proposed for quantification of the correlation [12, 13]. An interesting model has been proposed by Venkataraman et al. [12], whom modified the Cole-Cole model relating blood dielectric properties with frequency to include glucose concentration. The model is important not only because it demonstrates feasibility of dielectric NGM, but it also enables the simulation of various effects in NGM using standard human body models of the Electromagnetic (EM) simulation tools, such as CST Voxel Model. The current contribution explores the specific absorption rate (SAR) of microwaves for major blood vessels of thigh, arm and neck. Furthermore, comparison of blood in free space to that in body's vessels enables the critical analysis of laboratory measurements compared to real-life scenarios. To mimic the conditions investigated by Venkataraman et al., the simulations were conducted utilizing 1–10 GHz frequencies and hyperglycemia (300 mg/dl), hypoglycemia (40 mg/dl) and euglycemic (100 mg/dl) concentrations.

2. MATERIALS AND METHODS

CST Human Body Voxel model (Figure 1(a)) was utilized for simulations, where blood voxels in selected slices in the thigh, arm, and neck were selected for pertinent calculations (Figure 1(b)). These locations were selected because they contain major systemic blood vessels, such as the femoral

artery & vein (thigh), brachial artery & vein (arm), and carotid artery & jugular vein (neck). A near-field solution was obtained for an antenna placed closest to major blood vessels in the inner thigh, outer arm, and neck surface. Permittivity and conductivity values of blood were calculated using the Cole-Cole equation [14]:

$$\varepsilon(\omega) = \varepsilon_{\infty} + \sum_{n=1}^2 \frac{\Delta\varepsilon_n}{1 - (j\omega\tau)^{1-\alpha_n}} + \frac{\sigma_s}{j\omega\varepsilon_o} \quad (1)$$

where $\omega = 2\pi f$ is the angular frequency, $\Delta\varepsilon = \varepsilon_s - \varepsilon_{\infty}$ is the relaxation strength, ε_s and ε_{∞} , τ is the relation time constant, α is an empirical measure of frequency broadening, σ_s is the static (DC) conductivity, and ε_o is the permittivity of free space.

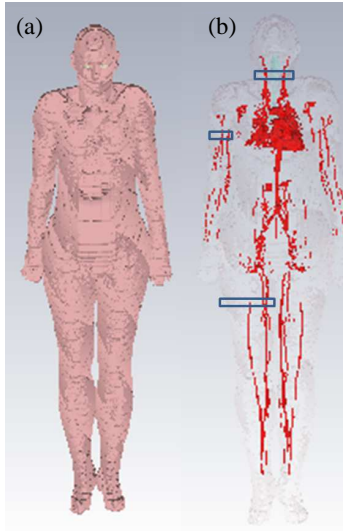


Figure 1: (a) Complete CST human Voxel body model; (b) associated blood vessels of body model, where simulations were conducted on major blood vessels in selected slices in the thigh, arm, and neck.

Model parameters of Equation (1) were used as established by Gabriel et al. [15], and the corresponding permittivity and conductivity of blood for 1–10 GHz were fed into CST. Furthermore, to simulate the effect of variation of blood glucose in NGM, a modified Cole-Cole equation [12] was used to calculate permittivity and conductivity values for use by CST:

$$\varepsilon(\omega) = \Re\left(\varepsilon_{\infty} + \sum_{n=1}^2 \frac{\Delta\varepsilon_n}{1 - (j\omega\tau)^{1-\alpha_n}}\right) \times (-0.001445g + 1.145882) + \Im\left(\varepsilon_{\infty} + \sum_{n=1}^2 \frac{\Delta\varepsilon_n}{1 - (j\omega\tau)^{1-\alpha_n}} + \frac{\sigma_s}{j\omega\varepsilon_o}\right) \quad (2)$$

where g is the blood glucose concentration, and the parametric values utilized in Equations (1) and (2) are displayed in Table 1.

Simulations were carried out for blood in the body's blood vessels and blood in free space to calculate SAR values using permittivity values from Equations (1) and (2).

3. RESULTS AND DISCUSSION

Several body locations have been studied for electromagnetic NGM, including the wrist [5, 8], arm [16], fingertip [7], and ear lobe [17]. However, no comparative evaluation of the optimum body compartment for dielectric NGM has been conducted. SAR values of blood in the thigh, neck and arm demonstrate that the highest power loss is in the thigh (Figure 2(a)). This suggests that the thigh might not be the optimum location for NGM because of high SAR values of thigh major blood vessels. electromagnetic NGM exploits reflected power to calculate the relative permittivities of blood and hence blood glucose concentration. As a result, maximum reflected power and minimum absorbed power would yield optimum Signal to Noise. Although the neck has the lowest SAR value, the arm presents a reasonable choice that balances between patient convenience and moderate power absorption of incident EM wave.

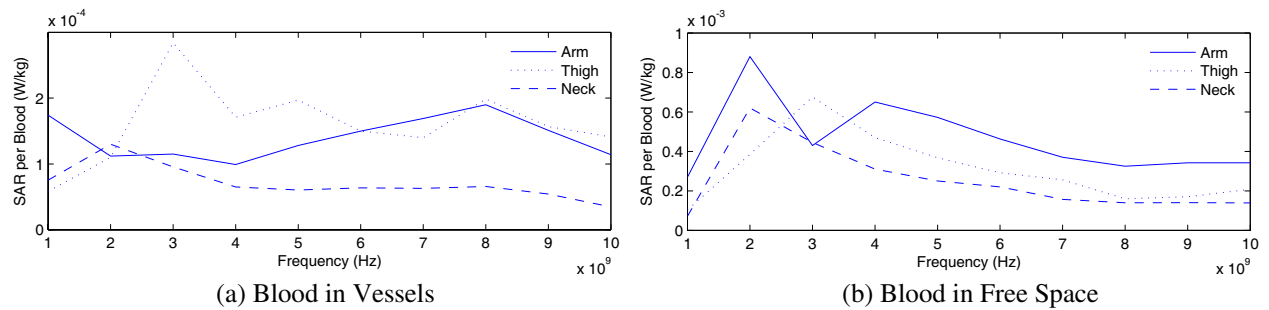


Figure 2: Specific absorption rate (SAR) of blood voxels for selected slices in arm, thigh and neck within (a) the human body, and (b) in free space.

Many dielectric NGM studies perform their tests utilizing blood samples in the laboratory. Therefore, it was important to compare the SAR calculations for blood in the body's vessels relative to free space for the same voxels, which are presented in (Figure 2(b)). The differences between calculated blood SAR in vessels relative to free space can be attributed to the effect of various body components (other than blood) on the electric field patterns, as shown in (Figure 3).

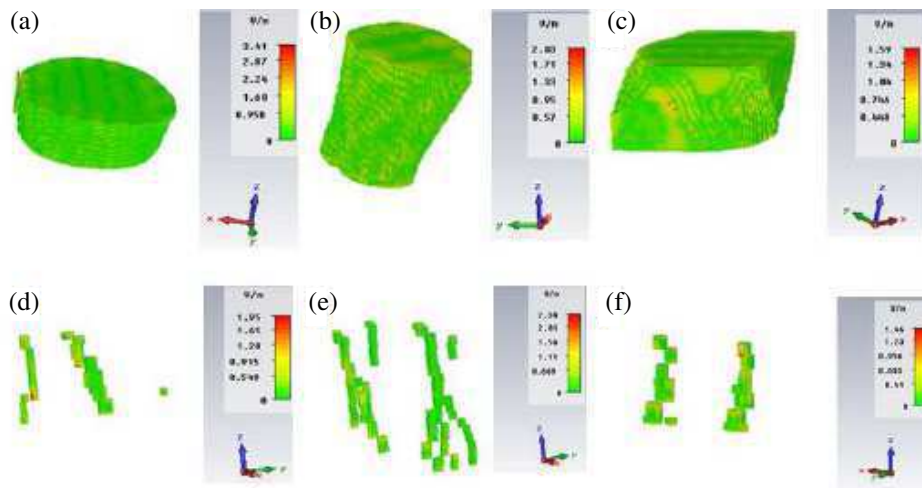


Figure 3: Calculated electric fields of selected slices in (a) thigh, (b) arm, (c) neck, (d) thigh blood in free space, (e) arm blood in free space, and (f) neck blood in free space.

Simulations for the blood SAR values using permittivity and conductivity values calculated using Equations (1) and (2) has shown no significant difference for simulated glucose concentrations, i.e., hyperglycemia (300 mg/dl), euglycemia (100 mg/dl), and hypoglycemia (40 mg/dl) (Figure 4(a)). This may have been expected because SAR is largely a function of conductivity, while the modification of the Cole-Cole model only affected the real component of the complex conductivity and did not include the imaginary component [12]. Other models have modified the Cole-Cole model to account for glucose concentration using both the real (ϵ') and imaginary parts (ϵ'') of the complex permittivity [13]. On the other hand, some differences may be observed for hyperglycemia calculations for blood in free space (Figure 4(b)). In the case that the dimensions of the object becomes small compared to the wavelength (i.e., wavelength inside the material is greater than ten times the dimensions of the object), Electric field becomes inversely proportional to relative permittivity [17]. Under such conditions, SAR becomes inversely proportional to relative permittivity [19]. This was the case for blood in free space, where the object's dimensions were an order of magnitude less than the wavelength at 10 GHz (3 cm). These results are in agreement with previous reports demonstrating the dependence of the localized SAR value of organs in EM Human Body simulations on both the real and imaginary values of complex permittivity [20].

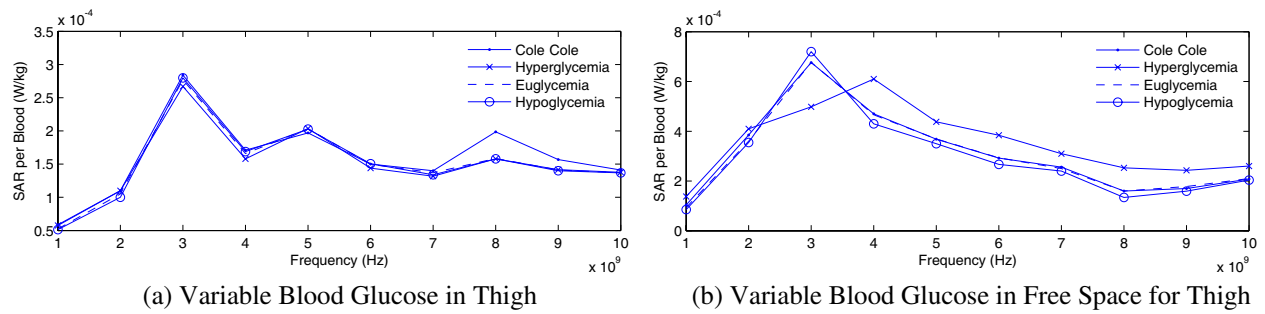


Figure 4: SAR of blood voxels for selected slices in thigh using Cole-Cole model and hyperglycemia, euglycemia, & hypoglycemia in modified Cole-Cole model within (a) the human body, and (b) in free space.

4. CONCLUSION

The results of this contribution suggests that the arm or neck may be better locations for NGM measurements than the thigh, although the arm is naturally more convenient. Differences between blood in vessels and in free space indicate the importance of considering the effect of other body tissues on dielectric measurements of blood glucose. Last but not least, contributions of both real and imaginary components of complex permittivity should be considered when conducting or modelling dielectric NGM measurements.

ACKNOWLEDGMENT

The work was supported by the Youssef-Jameel Science and Technology Research Center.

REFERENCES

- Whiting, D. R., L. Guariguata, C. Weil, and J. Shaw, "IDF diabetes atlas: Global estimates of the prevalence of diabetes for 2011 and 2030," *Diabetes Res. Clin. Pr.*, Vol. 94, No. 3, 311–321, 2011.
- Khalil, O. S., "Spectroscopic and clinical aspects of noninvasive glucose measurements," *Clin. Chem.*, Vol. 45, No. 2, 165–177, 1999.
- Hayashi, Y., L. Livshits, A. Caduff, and Y. Feldman, "Dielectric spectroscopy study of specific glucose in uence on human erythrocyte membranes," *J. Phys. D: Appl. Phys.*, Vol. 36, No. 4, 369–374, 2003.
- Livshits, L., A. Caduff, M. S. Talary, and Y. Feldman, "Dielectric response of biconcave erythrocyte membranes to D-and L-glucose," *J. Phys. D Appl. Phys.*, Vol. 40, No. 1, 15–19, 2007.
- Caduff, L., E. Hirt, Y. Feldman, Z. Ali, and L. Heinemann, "First human experiments with a novel non-invasive, non-optical continuous glucose monitoring system," *Biosens. Bioelectron.*, Vol. 19, No. 3, 209–217, 2003.
- Caduff, A., F. Dewarrat, M. Talary, G. Stalder, L. Heinemann, and Y. Feldman, "Non-invasive glucose monitoring in patients with diabetes: A novel system based on impedance spectroscopy," *Biosens. Bioelectron.*, Vol. 22, No. 5, 598–604, 2006.
- Nikawa, Y., "Blood-sugar monitoring by reflection of millimeter wave," *Asia-Pacific Microwave Conference*, 1–4, Bangkok, Thailand, Dec. 2007.
- Jean, B. R. and E. C. Green, "A microwave frequency sensor for non-invasive blood-glucose measurement," *Sensors Applications Symposium*, 4–7, Atlanta, United States, Feb. 2008.
- Kim, S., H. Melikyan, J. Kim, A. Babajanyan, J. H. Lee, L. Enkhtur, B. Friedman, and K. Lee, "Noninvasive in vitro measurement of pig-blood D-glucose by using a microwave cavity sensor," *Diabetes Res. Clin. Pract.*, Vol. 96, No. 3, 379–384, 2012.
- Mason, A., O. Korostynska, M. Ortoneda, A. Shaw, and A. Al-Shammaá, "A resonant coplanar sensor at microwave frequencies for biomedical applications," *Sens. Actuators A: Phys.*, Corrected Proof, Available online, May 3, 2013.
- Huber, D., M. Talary, F. Dewarrat, and A. Caduff, "The compensation of perturbing temperature fluctuation in glucose monitoring technologies based on impedance spectroscopy," *Med. Biol. Eng. Comput.*, Vol. 45, No. 9, 863–876, 2007.

12. Venkataraman, J. and B. Freer, "Feasibility of non-invasive blood glucose monitoring: In-vitro measurements and phantom models," *2011 IEEE International Symposium on Antennas and Propagation (APSURSI)*, 603–606, Spokane, United States, Jul. 2011.
13. Karacolak, T., E. C. Moreland, and E. Topsakal, "Cole-Cole model for glucose-dependent dielectric properties of blood plasma for continuous glucose monitoring," *Microw. Opt. Technol. Lett.*, Vol. 55, No. 5, 1160–1164, 2013.
14. Cole, K. S. and R. H. Cole, "Dispersion and absorption in dielectrics I. Alternating current characteristics," *J. Chem. Phys.*, Vol. 55, No. 5, 341–351, 1941.
15. Gabriel, S., R. W. Lau, and C. Gabriel, "The dielectric properties of biological tissues: III. Parametric models for the dielectric spectrum of tissues," *Phys. Med. Biol.*, Vol. 41, No. 11, 2271–2293, 1999.
16. Caduff, A., M. Mueller, A. Mergej, F. Dewarrat, R. E. Suri, J. Klisic, M. Donath, P. Zakharov, D. Schaub, D. Schaub, and W. A. Stahel, "Characteristics of a multisensor system for non invasive glucose monitoring with external validation and prospective evaluation," *Biosens. Bioelectron.*, Vol. 26, No. 9, 3794–3800, 2011.
17. Harman-Boehm, I., A. Gal, A. M. Raykhman, J. D. Zahn, E. Naidis, and Y. Mayzel, "Noninvasive glucose monitoring: A novel approach," *J. Diabetes Sci. Technol.*, Vol. 3, No. 2, 253–260, 2009.
18. Jackson, J. D., *Classical Electrodynamics*, Wiley-Interscience, New York, 1975.
19. Hurt, W. D., J. X. Ziriaux, and P. A. Mason, "Variability in EMF permittivity values: Implications for SAR calculations," *IEEE Trans. Biomed. Eng.*, Vol. 47, No. 3, 396–401, 2000.
20. Gajsek, P., W. D. Hurt, J. X. Ziriaux, and P. A. Mason, "Parametric dependence of SAR on permittivity values in a man model," *IEEE Trans. Biomed. Eng.*, Vol. 48, No. 10, 1169–1177, 2001.

Ultra Weak Photon Emission from *Saccharomyces Cerevisiae*

K. Červinková^{1,2}, M. Nerudová^{1,2}, M. Cifra², J. Hašek³, and J. Vrba¹

¹Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic

²Institute of Photonics and Electronics, Academy of Sciences of the Czech Republic, Czech Republic

³Institute of Microbiology, Academy of Sciences of the Czech Republic, Czech Republic

Abstract— This paper introduces the basic concept of measuring spontaneous ultra weak photon emission from yeast *Saccharomyces cerevisiae*. The objective of the work is a comparison of ultra weak photon emission from two mitochondrial mutants which differ in an activity of their energy-generating pathway. Photon emission has been measured in a visible range utilizing sensitive photomultiplier tube.

1. INTRODUCTION

All living organisms exhibit electromagnetic radiation in the optical range of the spectrum. The phenomenon is commonly known as ultra weak photon emission, but the term biophotons is also widely used. These photons are emitted without any external photoexcitation, they have solely endogenous origin resulting from metabolic processes [1]. Ultra weak photon emission occurs in most of the living cells and may play role in regulatory processes of organisms [2, 3].

Generally accepted model of the origin of the ultra weak photon emission considers reactions of free radical and reactive oxygen species (ROS) with biomolecules as lipids and proteins [4] which lead to electron excitation. The main source of ROS derives from the oxidative metabolism of mitochondria [5].

For the experiments described in this paper yeast cells of a strain *Saccharomyces cerevisiae* were used. This model microorganism is applied in many fields of research and contributes significantly towards understanding of eukaryotic cell biology.

S. cerevisiae is a facultative anaerobe microorganism, hence is capable to grow either anaerobically or aerobically. To satisfy its energy needs *S. cerevisiae* can exploit fermentable or nonfermentable carbon sources. When grown in glucose abundant medium *S. cerevisiae* use fermentation exclusively to obtain energy. Glucose is converted to ethanol and carbon dioxide during the process of fermentation and the energy is released in the form of ATP and heat. After switching to nonfermentable carbon source yeast cannot utilize fermentation, hence a different metabolic pathway is used to secure energy supply. Particular reaction depends on an available grow substrate.

As most of other living systems, yeast gains their energy mainly by oxidation of received carbon compounds such as saccharides, lipids or proteins. The key molecule of the carbon metabolism is pyruvate which can be subsequently processed by two possible ways — aerobically or anaerobically through tricarboxylic acid (TCA) cycle or fermentation, respectively. Pyruvate can be converted to Acetyl-CoA by a direct oxidative decarboxylation in the mitochondrial matrix. Acetyl-CoA is an important carbon compound which is essential in the TCA cycle.

Saccharides provide energy in a process of glycolysis where glucose is converted by a sequence of reactions to pyruvate. Energetically, the most effective usage of carbon sources requires coupling of the glycolysis to the TCA cycle which is a crucial metabolic pathway during which cells produce a large amount energy-rich molecules ATP. These two important parts of energy-generating pathway are coupled by pyruvate dehydrogenase complex (PDC) [5].

2. PREMISE

For yeast grown under aerobic conditions the PDC is a typical enzyme for oxidatively de carboxylating pyruvate to acetyl-CoA and carbon dioxide. The activity of PDC is regulated by a reversible phosphorylation of its subunits. Phosphorylation and dephosphorylation are catalysed by a pyruvate dehydrogenase-specific kinase and phosphatase, respectively. Proteins involved in *S. cerevisiae* as the PDC activity regulators are known under the designation Pkp1p and Ptc5p (or Ppp1p, alternatively). The phosphorylation of PDC is ensured by Pkp1p and dephosphorylation is ensured by Ptc5p. Phosphorylation of all three subunits leads to complete inactivation of PDC, therefore yeast strains with deleted gene encoding Ptc5p cannot transfer pyruvate to acetyl-CoA. Thus, the TCA cycle is not exploited, the mitochondrial oxidative metabolism is reduced which lead to decrease of reactive oxygen species production. We assume that *ptc5*Δ mutant exhibits a lower photon emission compared to the *pkp1*Δ mutant.

3. MATERIALS AND METHODS

3.1. Yeast

Two deletion mitochondrial mutants (*ptc5* Δ and *pkp1* Δ) of budding yeast *Saccharomyces cerevisiae* (Euroscarf collection; genetic background BY4741, MATa) were examined.

3.2. Media

Two different types of growth media were prepared. One with a fermentable carbon source — YPD (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) D-glucose in distilled water) and the second one with non-fermentable carbon source — YPG (1% (w/v) yeast extract, 2% (w/v) peptone, 3% (v/v) D-glucose in distilled water).

3.3. Measuring Protocol

Yeast cultures were stored on agar plates (1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) agar, 2% (w/v) D-glucose in distilled water) in a refrigerator at 4°C. Yeast were reinoculated on a new agar plate and left at the room temperature in a dry clean place before beginning of the cultivation in liquid medium. Each strain of yeast was inoculated into two glass Erlenmeyer flasks (50 ml) with 6 ml of YPD medium and cultivated for 16 hours at 30°C on an orbital shaker (Biocer) at 300 rpm. After this period the yeast suspense was centrifugated at 3000 rpm for 5 min. The pellet was resuspended with 1 ml of 2% sucrose solution. Optical absorbance was measured by spectrophotometer (Varian 4000) at the wavelength 600 nm in order to determine cell density. Cell quantity was adjusted to $5 \cdot 10^9$ to ensure the same number of yeast cells in each individual experiment. Required number of cell was transferred from sucrose solution into 5 ml of YPG or YPD medium and the photon emission for each sample was measured in a glass vessel for 50 min.

Cultivation and sample preparation was carried out in conditions of darkness. Media were stored in darkened containers. Following this procedure probability of delayed luminescence from samples was reduced.

3.4. Measurement Equipment

Photomultiplier tube R4220P, selected type (Hamamatsu Photonics K. K.) with a spectral sensitivity in the range 185–710 nm with highest sensitivity at 410 nm was employed to detect photon emission. This photomultiplier has a very sensitive cathode and anode and a low dark current, making it suitable for measuring ultra weak photon emission. It is a side-on type detector with an opaque reflection photocathode and the annular cage of dynodes. The input window is made from a quartz glass that transmits a wider range of radiation compared to conventional glass. The photomultiplier was powered by –1150 V from a high voltage power supply PS350 (Stanford Research Systems). Cooling unit C9144 (Hamamatsu Photonics K. K.) operating at –23°C was used to reduce thermoemission of the photomultiplier to the lowest achievable value. Thus, the average dark count of the system persisted around 0.8 counts per second. Measurement of the sample took place in a light-tight chamber made of dark anodized duralumin plates and specially designed for the purposes of ultra weakphoton emission measurements. The photomultiplier was placed in the bottom of the chamber targeting into a sample which was placed 2.5 cm from the photomultiplier tube. Temperature inside the chamber was regulated by A2A thermocontrol unit (UWE Electronic) to the value 28°C.

4. RESULTS AND DISCUSSION

Data of photon emission measurements of two different mitochondrial mutants are analyzed for the mean value of photon emission and signal to noise ratio (SNR). Results of photon emission measurements are presented in the following figures. Both yeast strains grown in glucose containing medium use the same energy-gaining pathway (fermentation), thus the level of photon emission shows comparable values (Figure 1). The average value of the measured dark count was 0.764 counts per second. The average values of the photon emission from *ptc5* Δ (MUT 90) and *pkp1* Δ (MUT 42) mitochondrial mutants were 0.960 and 0.958 counts per second, respectively. Since the measured signals originating from living cells are extremely low, signal to noise ratio (SNR) is an important factor to be considered. The calculated value of SNR for signals from *ptc5* Δ (MUT 90) and *pkp1* Δ (MUT 42) was 1.247 and 1.245, respectively.

Figure 2 represents data measured from yeast grown in a medium with glycerol (YPG). The average value of the dark count was 0.8 counts per second. The average values of the photon emission from *ptc5* Δ (MUT 90) and *pkp1* Δ (MUT 42) mitochondrial mutants were 1.25 and 1.37 counts per second, respectively. The value of SNR for signals from *ptc5* Δ (MUT 90) and *pkp1* Δ

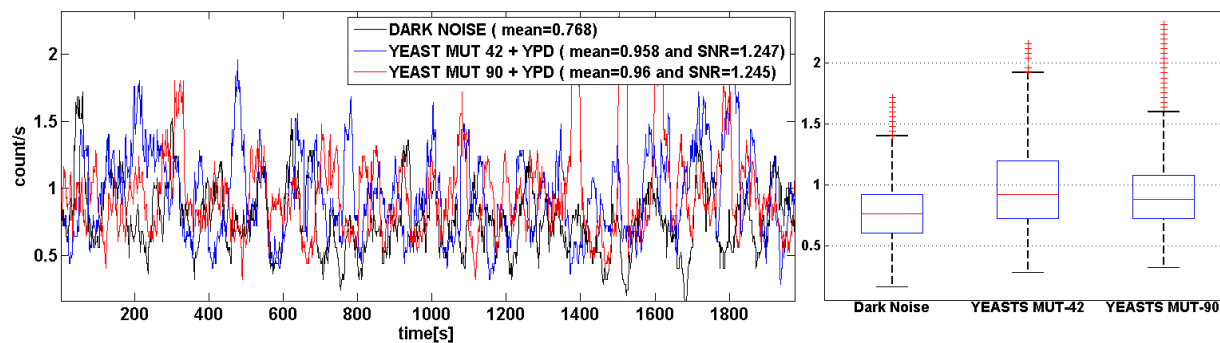


Figure 1: Photon emission from *ptc5* Δ (MUT 90) and *pkp1* Δ (MUT 42) mitochondrial mutants measured in YPD.

(MUT 42) was 1.72 and 1.57, respectively. Used yeast strains were expected to metabolize glycerol with different efficiency and thus to vary in the ROS production in mitochondria. There is not a statistically significant difference, but only slight tendency of higher photon emission from the yeast strain *pkp1* Δ (MUT 42), which was expected to produce a higher amount of ROS.

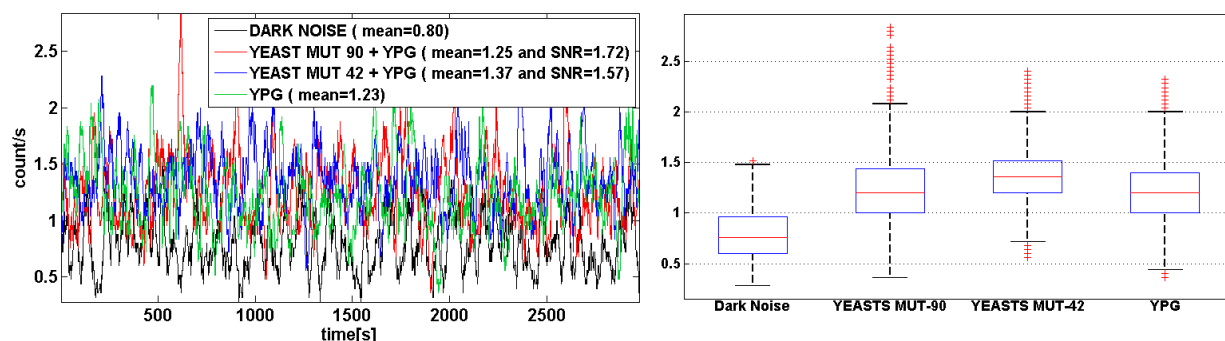


Figure 2: Photon emission from *ptc5* Δ (MUT 90) and *pkp1* Δ (MUT 42) mitochondrial mutants measured in YPG. YPG — emission from medium only, Dark Noise — noise of the detector.

Presented data are not completely satisfying due to low SNR. In comparison to the measured photon emission the value of the noise was relatively high. The noise originates mainly from the thermoemission of the photomultiplier. Several other factors also contribute to the noise level, e.g., cosmic radiation, leak current, scintillation of the glass envelope, etc.. Light emission from medium has to be taken into account as it may contribute to the background noise. The YPG medium contains glycerol from which a several groups of lipids are derived. Since lipid peroxidation is considered as a relevant source of photon emission, photon emission from glycerol due to autooxidation by ambient air can be also an important factor influencing the total noise value.

From the biological point of view it is also very important to perceive that a conversion of pyruvate to acetyl/CoA in *S. cerevisiae* does not rely on just one protein. Genome of *S. cerevisiae* encodes an additional protein included in the pyruvate — acetyl-CoA transfer, so an inactivation of the PTC5 gene may necessarily eliminate the whole carbon metabolism [6]. The missing part of the metabolic pathway can be bypassed by other ways. Thus, the loss of PDC function could be insignificant for the creation of ROS and the consecutive formation of electron excited states and emission of photons. That may be one of the reasons why the differences in emissions between the strains are less apparent.

5. CONCLUSION

We presented here first attempt to detect differences of ultra weak photon emission between mitochondrial yeast mutants. Our preliminary results seem to indicate possible differences of the ROS formation in mitochondria between two deletion strains of *S. cerevisiae* leading to a different intensity of ultra weak photon emission but the differences are not of statistical significance. To

state any conclusion more repetitions of experiments and modifications of the protocol are necessary because the statistical evaluation did not yield significant differences in studied groups of cells. The possible improvements of the experiment is founded in better noise reduction of the measuring system, noise elimination from the medium and utilization of bigger quantity of cells.

ACKNOWLEDGMENT

The research presented in this paper has been financially supported by the grant No. SGS13/077/OHK3/1T/13 (Study of Electromagnetic Processes in biomolecules, cells and tissue), by the Czech Science Foundation, GA ĀR grant No. P102/11/0649 and by the CzechScience Foundation, GA ĀR grant No. GP13-29294S.

REFERENCES

1. Cifra, M., "Electromagnetic cellular interactions," *Progress in Biophysics and Molecular Biology*, Vol. 105, No. 3, 223–246, 2011.
2. Fels, D., "Cellular communication through light," *PLoS ONE*, Vol. 4, No. 4, e5086, 2009.
3. Farhadi, A., "Evidence for non-chemical, non-electrical intercellular signaling in intestinal epithelial cells," *Bioelectrochemistry*, Vol. 71, No. 2, 142–148, 2007.
4. Prasad, A., "Linoleic acid-induced ultra-weak photon emission from *chlamydomonas reinhardtii* as a tool for monitoring of lipid peroxidation in the cell membrane," *PLoS ONE*, Vol. 6, No. 7, e22345, 2011.
5. Scheffler, I. E., *Mitochondria*, Wiley-Liss, Hoboken, 2008.
6. Krause-Buchholz, U., "YIL042c and YOR090c encode the kinase and phosphatase of the *Saccharomyces cerevisiae* pyruvate dehydrogenase complex," *FEBS Letters*, Vol. 580, No. 11, 2553–2560, 2006.

Changes in DNA Methylation Wheat Seedlings in First and Second Generations under Influence of EHF EMI

L. A. Minasbekyan and M. S. Abovyan

Department of Biophysics, Yerevan State University, 1 Alek Manukyan St, Yerevan 0025, Armenia

Abstract— In the cells *in vitro* can be occurred the changes in the gene expression on replay to the different stresses such as high temperature, hypoxia and other injurious factors. DNA methylation in critical regulatory regions is involved in regulating gene expression. There is an overall inverse correlation between DNA methylation in regulatory regions of genes and gene expression, and it was confirmed by whole-genome approaches. Nevertheless the role of modified 5-methylcytosine in controlling gene expression is unclear yet.

We study the level of DNA methylation of 4-day seedlings of hexaploid wheat in two generations. In our investigations we separate DNA from seedlings of control and treated by EHF EMI (Extremely High Frequencies of Electromagnetic Irradiation) of soaked seeds on 3-rd day after treated. Part from each samples of treated seedlings growing in soil to get a harvest (obtained a second generation). The treatment of seeds we carried out in the range of 45–53 GHz frequencies of EMI. As we suggested EHF EMI can bring to some changes in aqua environment and DNA conformation, which can lead to changes in the level of DNA methylation. The data indicate that level of DNA 5m-C changes in the first generation depend of EHF EMI frequencies and exposition. After that we have investigate the level of DNA methylation from 4-th day seedlings of seeds of new harvest (second generation), which was growing from the previously treated seedlings. Obtained in our study data shows that the changes in DNA methylation in first generation of seeds during the plant ontogeny partially conserved and pass to the seedlings of obtained seeds in second generation. So we have shown on the plant model that the changes in biological systems have rather epigenetic character and partially pass to the next generation.

1. INTRODUCTION

The increase in the use of mobile phones (MPs), as well as wide using mm-waves-therapy in medicine recent years has raised the problem of health risk connected with high-frequency electromagnetic fields. There are reports of headache, dizziness, numbness in the thigh, and heaviness in the chest among MP users. This paper deals effect of mm-waves on DNA methylation in such model system as the growing wheat seeds. Methylation/demethylation of DNA are one of the mechanisms used by cells to control gene expression. As it is known, DNA methylation helps to maintain and stabilize the global inactive state of chromatin, thus playing role in turning off transcription from a large number of genes that are not required in a particular differentiated cell [1]. DNA methylation patterns are susceptible to change in response to environmental stimuli, whereby the epigenome seems to be most vulnerable during early *in utero* development. Aberrant DNA methylation changes have been detected in several diseases, particularly cancer where genome-wide hypomethylation coincides with gene-specific hypermethylation [2]. An analysis of the mutation frequency in potentially methylated sites revealed a three-fold increase in transposable elements mutation frequency relative to intron and untranslated genic regions. This increase is consistent with wheat transposable elements being preferentially methylated, likely by sRNA targeting. The repetitive, non-genic regions of wheat, as in many plants genomes, primarily consist of transposable elements [3].

Taking into account that DNA 5mC is an important parameter for evaluated the state of cell we investigate influence of mm-waves on the wheat seeds during germination. The importance changed in the DNA methylation level has been revealed on the seedlings growth to the 3-d day after alone exposition by mm-waves of soaked seed. Thus, we have shown a significant effect of MM-waves on the level of DNA methylation, which as proposed by us has epigenetic nature and allows the biological system to adapt to new environmental conditions.

2. MATERIALS AND METHODS

Seed germination. In our experiments have been used seeds of hexaploid wheat (*Triticum Aestivum*, cvs. Voskehask) soaked overnight in tap water in darkness at 27°C, and then treated by nonthermal coherent waves in the range of 45–53 GHz for 40 min. After that seeds growing in the Petri dishes at 26 for 72 h yet. From each samples taken a part of seedlings and continued to grow in the ground.

DNA isolation. DNA extracted according to the method detail described in our article [9] and after washing 2–3 times by ether: ethanol mix and dried in thermostat at 170°C. Dry DNA preparations were hydrolyzed with formic acid in closed ampules. Bases were separated by twice-repeated one dimensional ascending paper (Filtrak, Czech Republic) chromatography in the system of n-butanol: water: 25%NH₄OH in proportion 60:10:0.1. The contents of individual nucleotides were estimated spectrophotometric and the quantity was calculated as described in [cit. 9 in 4].

3. RESULTS AND DISCUSSION

In the cells *in vitro* can be occurred a protein degradation, and in more cases can lead to the changes in the gene expression on replay to the influence of different stresses such as high temperature, hypoxia and other injurious factors. Potential biological effects of low-power millimeter waves on endoplasmic reticulum [5], on cell proliferation and cell cycle of melanoma cell lines [6] and some others have been shown by many researchers on last decade. We have been chosen very well investigated by us model of wheat seeds where dry embryo of wheat seeds are the fully passive state of genome and 4-day seedlings are an activated and differentiated state of cells [4, 7–9].

Investigation carried out in range of 45–51.8 GHz frequencies of EMI. As we suggested EHF EMI can bring to some changes in DNA conformation and aqua environment of cells, which can lead to DNA methylation level changes as well. The data indicate that 5-mC changes depend of EHF EMI frequencies and exposition.

Our investigation we carry out *in vivo*, which take us advance to study differences in the control and irradiated cells, comprised in the range of frequency effects of low power millimeter waves (MMWs) at 45 GHz and 50.3 GHz used in millimeter wave therapy in the order to revealing impact of this factor on a cell genome. Our study has shown that under influence of EHF EMI have been occurred changes in DNA methylation of 4-day wheat seedlings. Data presented on the Fig. 1 evidence about increasing in DNA methylation level of total DNA sequence under treatment during 40 min at the 45 GHz and 46 GHz. Beside it, has been obtained decreasing in DNA methylation level under same time exposition after impact by 49 GHz, 50 GHz and 50.3 GHz frequency of mm-waves. In this work we study level of DNA methylation of total DNA wheat seedlings, not a special sequence of any gene. Nevertheless from literature it is known, that absence of methylation at the ABCB1 gene promoter of human cells correlated with progressive disease during doxorubicin treatment, and DNA methylation status at the promoter of GSTP1, FOXC 1 and ABCB 1 correlated with survival [10].

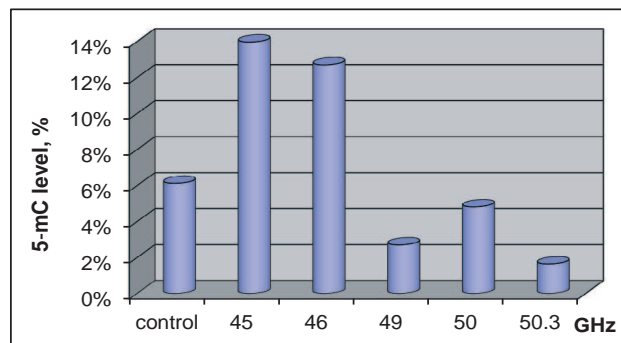


Figure 1: Correlation between level of DNA methylation and frequency EMI in the first generation of wheat seeds growth.

Although the genome often depicted as a static structure upon which protein-aqueous factors bind to control turn expression, the genome is actually highly mobile and capable of exploring the complex domain architecture of the nucleus, which in turn controls genome maintenance and gene expression. Epigenetics can be defined as the study of heritable changes of a phenotype such as the gene expression patterns of specific cell type that are not caused by changes in the nucleotide sequences of genetic code itself. Epigenetic phenomena are mediated by a variety of molecular mechanisms including post-transcriptional histone modifications, histone variants, ATP-dependent chromatin remodeling complexes, polycomb/trithorax protein complexes, small and other non-coding RNAs including siRNA and miRNAs, and DNA methylation [2]. Some activated genes are recruited to the nuclear periphery through interactions with nuclear pore complexes (NPCs), and NPC components are capable of preventing the spread of silent chromatin into adjacent regions

of active chromatin, leading to the speculation that NPCs may facilitate the transition of chromatin between transcriptional states [10, 11]. As such, it is perhaps not surprising that many disease states and moreover stress situations in the cellular life are frequently associated with alterations in NE. So alteration in many direction of nuclear envelope and adjacent to it peripheral chromatin may alter and as we show in this article chromatin conformation may changes through DNA methylation of transposable elements of genome under EHF EMI influence. But conformational changes it is not a change in the DNA sequences. Quantitative DNA methylation profiling is a powerful tool to identify molecular changes associated with specific phenotypes. We analyzed the DNA methylation of seedlings, obtained from the seed of second generation, which before growth was preliminary treated by mm-waves 50 GHz and 50.3 GHz for evidence of epigenetic character of such alterations.

Our experiments drawing by scheme, presented on Fig. 2. DNA separated from control and treated by EHF EMI seedlings on 3-rd day. Part of treated seedlings was growing in soil with aim to get a harvest of treated seeds (as a second generation).

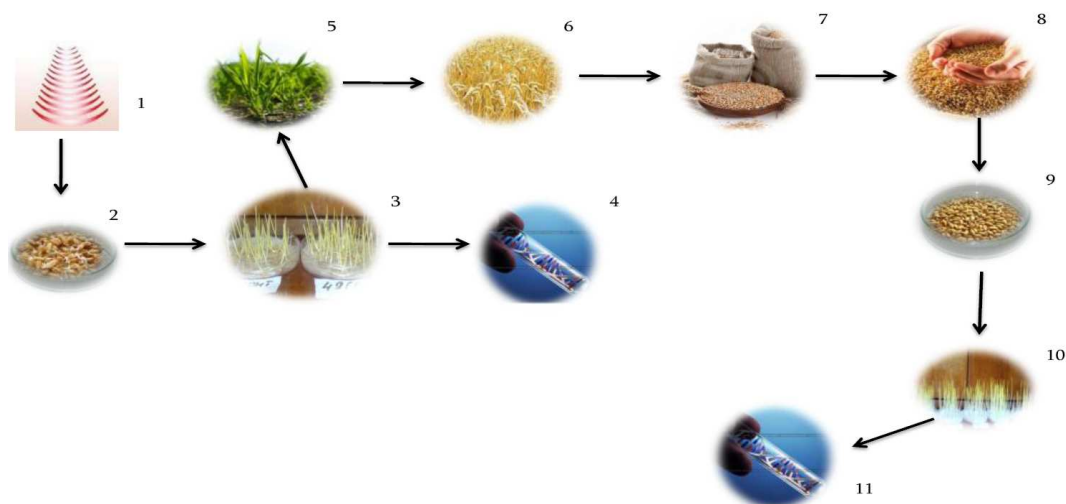


Figure 2: Scheme of our experiments: 1 — impact by EHF of EMI of soaked seeds (2), 3 — four-day seedlings, 4 — DNA extraction and estimation of 5-mC residues in DNA, 5 — growing of treated seedlings in soil, 6, 7, 8 — obtained harvest of treated seedlings, 9 — soaked seeds, without impact, 10 — four-day seedlings of second generation, 11 — DNA extraction and estimation of 5-mC residues.

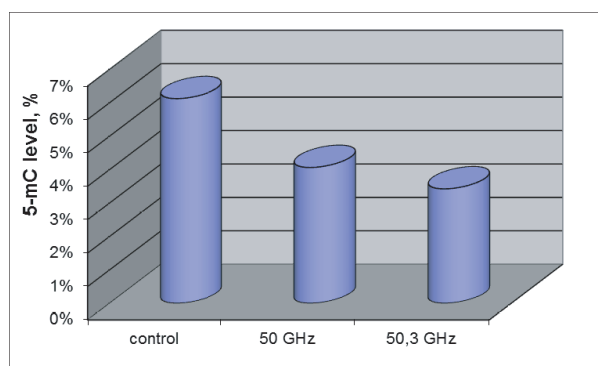


Figure 3: The level of DNA methylation of 4-day seedlings from the second generation.

After that we have investigate change in the level of DNA methylation from 4-th day seedlings of seeds from new harvest (second generation). But at this time we investigate DNA methylation without treated of seeds. Obtained in our study data shows that the decreasing in level of DNA methylation in first generation of seeds treated by 50 GHz and 50.3 GHz during the plant ontogeny partially conserved and pass to the seedlings of obtained seeds of the second generation, presented on the diagram (Fig. 3). So we have shown that changes in DNA methylation level have epigenetic character and can conserve the whole life of plant but they have shown partially pass to the next generation.

As it is above mentioned in the second generation we test only seeds, treated by 50 GHz and 50.3 GHz, as soon at this frequency we have been obtained sharp changes and this frequencies at same time correspondence to the preferentially absorption by water.

Data presented on the Fig. 3 evidence that treatment by EHF EMI at 50 GHz in the first generation decreased content of 5 mC from 6.12% to 4.81% and such rather low contents 5 mC remained in the 2nd generation approximately 4.06%. On the contrary influence of EHF EMI at 50.3 GHz led to the sharp down of 5 mC content in first generation, but in the second generation seeds though low level of a methylation rather control, but in the second generation it higher than in the first remained that testifies to gradual restoration.

4. CONCLUSION

As it is known DNA methylation is a mechanism for diversification of DNA identity by providing within the same chemical entity two layers of information: the ancestral identity encoded on the level of sequence and the cellular identity encoded on the level of the DNA methylation pattern [12]. Our experiments evidence the hypothesis that similar to alteration in DNA methylation that occur in response to innate signals during development, external signals triggered by the environment can modulate the DNA methylation pattern to generate differential “environmental history” DNA methylation identities. We hope, that our research to enhance investigations in this field for revealing really path and nature of mm-waves impact on the biological systems.

REFERENCES

1. Bird, A., “Gene number, noise reduction and biological complexity,” *Trend. Genet.*, Vol. 11, 94–100, 1995.
2. Tost, J., “DNA methylation: An introducing to the biology and diseases-associated changes of a promising biomarker,” *Mol. Biotechnol.*, Vol. 44, 71–81, 2010.
3. Cantu, D., L. S. Vanzetti, A. Sumner, M. Dubcovsky, M. Matvienko, A. Distelfeld, R. W. Michelmore, and J. Dubcovsky, “Small RNAs, DNA methylation and transposable elements in wheat,” *BMC Genomics*, Vol. 11, 408–417, 2010.
4. Minasbekyan, L. A., M. A. Parsadanyan, G. A. Panosyan, and P. O. Vardevanyan, “Changes in the nucleotide composition and pattern of DNA methylation during germination of cereal seeds,” *Russ. J. of Plant Physiology*, Vol. 48, No. 2, 256–259, 2000.
5. Nicolaz, C. N., M. Zhadobov, F. Desmots, A. Ansart, R. Sauleau, D. Thouroude, D. Michel, and Y. Le Drean, “Study of narrow band millimeter-wave potential interactions with endoplasmic reticulum stress sensor genes,” *Bioelectromagnetics*, Vol. 30, 365–373, 2009.
6. Beneduci, A., “Evaluation of the potential in vitro antiproliferative effects of millimeter waves at some therapeutic frequencies on RPMI 7932 human skin malignant melanoma cells,” *Cell. Biochem. Biophys.*, Vol. 55, 25–32, 2009.
7. Minasbekyan, L. A., S. A. Gonyan, M. A. Parsadanyan, and P. O. Vardevanyan, “The RNA – Export and electrokinetic potential of nucleus surface during germination of cereal embryos,” *Russ. J. of Plant Physiology*, Vol. 49, No. 2, 250–254, 2002.
8. Minasbekyan, L. A., Z. V. Yavroyan, M. R. Darbinyan, and P. O. Vardevanyan, “Comparative characteristics of phospholipid content in nucleic subfractions during embryos germination,” *Russ. J. of Plant Physiology*, Vol. 51, No. 5, 784–789, 2004.
9. Minasbekyan, L. A., Z. V. Yavroyan, M. R. Darbinyan, and P. O. Vardevanyan, “Changes in the content of phospholipids in nuclear subfractions of wheat seedlings under influence of Hibberellic acid,” *Russ. J. of Plant Physiology*, Vol. 55, No. 3, 372–377, 2008.
10. Van de Vosse, D. W., Y. Wan, R. W. Wozniak, and J. D. Aitchinson, “Role of the nuclear envelope in genome organization and gene expression,” *Willey Interdisciplinary Reviews: System Biology and Medicine*, Vol. 3, No. 2, 147–166, Apr. 2011.
11. Minasbekyan, L. A., V. P. Kalantaryan, and P. O. Vardevanyan, “Influence of lowintensity electromagnetic irradiation on the wheat seedlings chromatin fractions,” *Brilliant Light in Life and Material Sciences, Book Series: NATO Security Through Science Series B: Physics and Biophysics*, 199–203, 2007.
12. Szif, M., “The early-life social environment and DNA methylation,” *Clin. Genet.*, Vol. 81, 341–349, 2012.

Fractal Labyrinths and Planar Nanostructures

V. I. Grachev, A. A. Potapov, and V. A. German

Kotel'nikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Russia

Abstract— The paper presents the definition of a fractal labyrinth and a model of the 4-node elementary fractal labyrinth. Through an iterative procedure derived sequence of sets, the limit of which is a fractal labyrinth. In the work presented the basic equations and fractional operators to explore fractal labyrinths and phenomenon in them: anomalous diffusion and Levy flights, etc. Is spent a fractal analysis of images of labyrinth-similar planar nanosystems. It is shown that in the problems of the analysis of complex stochastic objects necessary and sufficient condition for an unequivocal result is to obtain a fractal signature and fractal cepstrum.

1. INTRODUCTION

Currently are widely used fractal-topological methods of processing and analysis of nanostructures [1]. The paper deals with the application of fractal formalism to the tasks of reconstruction complicated images of microstructures of nano-modified materials based on fractal labyrinths [2, 3], in particular, at equal fractal signatures of their parameters.

2. MODEL OF THE FRACTAL LABYRINTH

Definition: *Fractal labyrinth or labyrinth fractal is a topological connected dendritic structure with fractal dimension $d_f > 1$ and the scaling nature of the conductive pathway [2, 3].* For the study of such structures — labyrinthine paths, routes to them, their fractal properties, topology of exits, etc. — are commonly used model representations [4–6]. We will construct a labyrinth fractal on the square $m \times m$ grid with a characteristic size L . Here generating element, fractal generator is the transformed generating element of L. Kristi model [6], built on a 4-node grid ($n = 2$) in the unit square (Fig. 1(a)), here $n = 1$. On Fig. 1 the pre-fractal of generation n ($n = 2$) shows the development of the generating element by the recursive branching and scaling.

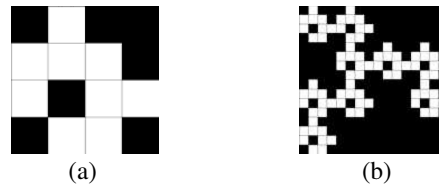


Figure 1: 4×4 -labyrinth fractals: (a) L_1 , origin, (b) L_2 , evolution.

Construction a fractal labyrinth (a fractal set) is performed using the following iterative procedure. Consider an arbitrary subset of points in the unit square. Let x, y, q — the point of the unit segment $x, y, q \in [0, 1]$ and the square $Q = [x, x + q] \times [y, y + q]$ is a subset of the unit square $Q \subseteq [0, 1] \times [0, 1]$. A subset of the points with $m \geq 1$ is given by $S_{i,j,m} = \left\{ (x, y) \mid \frac{i}{m} \leq x \leq \frac{i+1}{m}, \frac{j}{m} \leq y \leq \frac{j+1}{m} \right\}$ and $S_m = \{S_{i,j,m} \mid 0 \leq i \leq m-1, 0 \leq j \leq m-1\}$, i.e., for $m = 1$ we have the entire unit square $S_1 = (x, y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1$; for $m = 2$ — half of its $S_2 = (x, y) \mid 0 \leq x \leq 1/2, 0 \leq y \leq 1/2 + (x, y) \mid 1/2 \leq x \leq 1, 1/2 \leq y \leq 1 = S_{21} + S_{22}$; for $m = 3$ — one third of its $S_3 = (x, y) \mid 0 \leq x \leq 1/3, 0 \leq y \leq 1/3 + (x, y) \mid 1/3 \leq x \leq 2/3, 1/3 \leq y \leq 2/3 + (x, y) \mid 2/3 \leq x \leq 1, 2/3 \leq y \leq 1 = S_{31} + S_{32} + S_{33}$, etc. These sets are compact, their sequence is monotonically decreasing for $1 \leq m \leq 5$ is shown in Fig. 2.

For $n = 1$ let $W_1 \subset S_m$ and call it a *set of white squares of the first order* (Fig. 1(a)). Then the *set of the black squares of the first order* is defined as $B_1 = S_m \setminus W_1$. For any point of the unit square $(z_x, z_y) \in [0, 1] \times [0, 1]$ define the function $P_Q(z_x, z_y) = (qz_x + x, qz_y + y)$. For $n \geq 2$ we have $S_{m^n} = \{S_{i,j,m^n} \mid 0 \leq i \leq m^n - 1, 0 \leq j \leq m^n - 1\}$. Then the *set of white squares of order n* is defined by $W_n = \bigcup_{W_1 \in W_1, W_{n-1} \in W_{n-1}} \{P_{W_{n-1}}(W_1)\}$. Thus $W_n \subset S_{m^n}$. A *set of the black squares of order n* defined by $B_n = S_{m^n} \setminus W_n$.

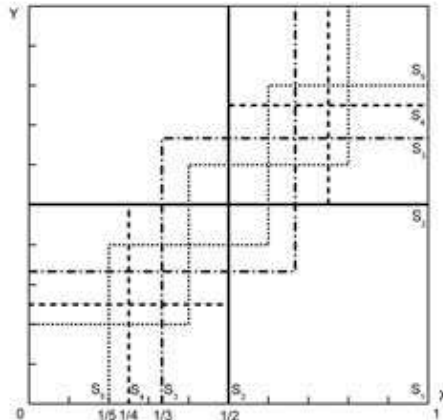


Figure 2: The sets of S_m on the unit square, with $1 \leq m \leq 5$.

If on the set W_n there is an associated graph $G(W_n)$, the sequence vertices of which is the path, if this path is the only and it contains no cycle, then such graph is a tree, the *dendrite*. At that the top row of the set W_n on the dendrite border is the set of all white squares in $\{S_{i,m^n-1,m^n} | 0 \leq i \leq m^n - 1\}$. The bottom row, left and right columns on the boundaries are defined analogously. And then *the top exit* T_n in W_n is the one and only one white square in the top row of order n , such that there is a white square in the same column in the bottom row of order n . The *bottom, left and right* exits are defined analogously. Such dendritic set W_n with exits is called an $m \times m$ -*labyrinthine set*.

We select on the labyrinthine set W_n ($n \geq 1$) a sequence of compact sets $L_n = \bigcup_{W \in W_n} W$. Then $\{L_n\}_{n=1}^{\infty}$ is her monotonically decreasing sequence, whose limit is $L_{\infty} = \bigcap_{n=1}^{\infty} L_n$, i.e., the *limit of set* W_1 . The limit set L_{∞} of labyrinthine set W_1 is called the *labyrinth fractal*. Its top exit is the limit $\bigcap_{n=1}^{\infty} T_n$. Another exit is defined analogously. And $(x, 1)$, $(x, 0)$ L_{∞} if and only if $(x, 1)$ is the point of the top exit L_{∞} and $(x, 0)$ — the lower exit L_{∞} . To the left and right exits an analogous statement is fairly.

In the connected sets the expansion intersection of any two subsets is an arc between their elements [7], the only arc between any pair of points in a limited labyrinth set. Length of this arc is infinite, and the set of all her points, in which there is no tangent, is dense. Hausdorff or fractal dimension $d_H(L_{\infty}) = \log |W_1| / \log m$ of the labyrinth fractal, as for any self-similar set W_1 , such, whose $|W_1| > m$, is $d_H > 1$. For six possible pairs of outlets (Fig. 1) — from top to bottom, left to right, top-right, right-down, bottom-left, left-up — there is a non-negative *matrix path of the labyrinth set*, whose element in row x and column y is the number of y -squares in the x -path. Matrix multiplication reflects the replacement of paths.

3. FRACTALS AND FRACTIONAL OPERATORS

The description of the fractal systems does not fit into the traditional frameworks of differential equations of integer order. The only and necessary mathematical apparatus for the study of fractal labyrinths — a fractional calculus [8–10]. Fractal labyrinth is a static prototype, the picture of the dynamics of diffusion in fractal space, the apparatus description which — fractional calculus. Diffusion in a fractal medium — anomalous diffusion — is described by fractional diffusion equation (FDE) [8, 9]

$$\frac{\partial W}{\partial t} = {}_0D_t^{1-\alpha} K_{\alpha} \frac{\partial^2 W}{\partial x^2}, \quad (1)$$

where $W(x, t)$ is the probability density function (PDF) to be at x at time t , depends, in general, on the particular geometry of the interaction; K_{α} — the generalized diffusion coefficient. The operator of the fractional derivative ${}_0D_t^{1-\alpha} = \frac{\partial}{\partial t} {}_0D_t^{-\alpha}$, $0 < \alpha < 1$ — Riemann-Liouville operator, defined by the integral relation [8, 9].

$${}_0D_t^{1-\alpha} W(x, t) = \frac{1}{\Gamma(\alpha)} \frac{\partial}{\partial t} \int_0^t dt' \frac{W(x, t')}{(t-t')^{1-\alpha}} \quad (2)$$

— direct continuation of the multiple integral Cauchy for an arbitrary complex α with $Re(\alpha) > 0$. Fractional derivative is set with fractional integration and then – the usual differentiation according to the formula ${}_t D_t^\beta f(t) = \frac{d^n}{dt^n} {}_t D_t^{\beta-n} f(t)$ with $Re(\beta) > 0$, the natural number n satisfies the inequality $n \geq Re(\beta) > n - 1$.

A generalized diffusion equation of fractional order is constructed using a continuous time random walk model, when the events of the transfer are subject to a wide statistics [1, 11, 12]. Probability density function $W(x, t)$ obeys the algebraic relations in Fourier-Laplace space $W(k, u) = \frac{1-w(u)}{u} \frac{W_0(k)}{1-\psi(k, u)}$, where $W_0(k)$ denotes the Fourier transform of the initial condition $W_0(x)$. The case of finite characteristic waiting time T and a diverging jump length variance Σ^2 can be modeled by a Levy distribution for the jump length, i.e., $\lambda(k) = \exp(-\sigma^\mu |k|^\mu) \approx 1 - \sigma^\mu |k|^\mu$ for $1 < \mu < 2$; corresponding to the asymptotic behavior $\lambda(x) \approx A_\mu \sigma^{-\mu} |x|^{-1-\mu}$ for $-x \gg \sigma$. Due to the finiteness of T , this process is of Markovian nature. Substituting the asymptotic expansion of the jump length in the ratio for $W(x, t)$, one obtains $W(k, u) = \frac{1}{u+K^\mu |k|^\mu}$, from which, upon Fourier and Laplace inversion, a generalized fractional diffusion equation FDE [11, 12]

$$\frac{\partial W}{\partial t} = K^\mu {}_{-\infty} D_x^\mu W(x, t) \quad (3)$$

is inferred. Here the operator of fractional derivative in the degree of μ is ${}_{-\infty} D_x^\mu$ – a fractional Weyl operator. The generalized diffusion constant is $K^\mu \equiv \sigma^\mu / \tau$ and carries the dimension $[K^\mu] = cm^\mu / c$.

Thus, both the case of particle dynamics in fractal labyrinths are essential – these being the Riemann-Liouville operator ${}_0 D_t^\alpha$ for $t_0 = 0$ and the Weyl operator ${}_{-\infty} D_x^\mu$ for $t_0 = -\infty$. Mathematically, the expression of a work ${}_0 D_t^\alpha f(t) = \frac{1}{\Gamma(\alpha)} t^{\alpha-1} * f(t)$ is a Laplace convolution, whereas the same expression for the Weyl operator ${}_{-\infty} D_x^\mu f(x) = \frac{1}{\Gamma(\mu)} x^{\mu-1} * f(x)$ is a Fourier convolution. Therefore, Laplace and Fourier transformations are an essential tool for solving fractional order differential and integro-differential equations.

4. FRACTAL ANALYSIS OF PLANAR NANOSTRUCTURES

Multi-component nanomodified planar structures synthesized by the Langmuir monolayer, their images obtained by transmission electron microscopy [13]. On Fig. 3(a) — image palladium nanostructures grown under ambient conditions and placed on a copper grid. Bar size 150 nm. Here is clearly visible labyrinth-similar structure of the synthesized nanosystem. Its physical nature is the result of the dynamic nonequilibrium processes. On Fig. 3(d) — planar palladium net-like nanostructure synthesized on the surface of the water phase. Bar size 200 nm.

On Figs. 3(b), (c), (e), (f) — the results of fractal analysis of these images: measured the fractal dimension D and fractal signatures (fractal kepstrums) of surface of microsamples. These measurements were made using software developed at the IRE RAS [1]. Showed the mean value of

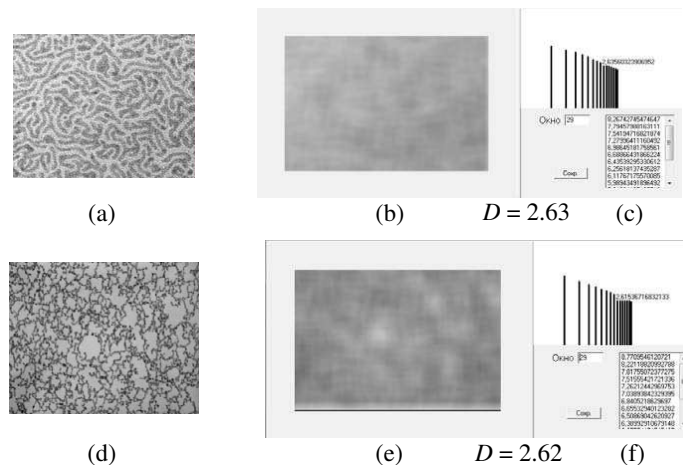


Figure 3: (a), (d) TEM micrograph of Pd nanostructures, (b), (e) and (c) the results of the fractal analysis, (f) — fractal signatures (fractal kepstrums).

the fractal dimension D and its root-mean-square value σ . In the mode of clustering, specifying the desired range of local fractal dimensions ΔD of labyrinth structure, the operator on the computer monitor receives field D , for which all the values $D_i \in \Delta D$ are displayed in one color. In those cases, when the image is composed of fragments with different values of the estimates of the fractal dimension, it makes sense to produce clustering of images in order to highlight areas with approximately constant estimates of D .

The data obtained show one of the main conclusions reached by us earlier, namely, the point estimate of the fractal dimension leads to absurd results. In the problems of the analysis of complex stochastic objects necessary and sufficient condition for an unequivocal result is to obtain a fractal signature and fractal keprum.

5. CONCLUSION

The problem of expanding entirely new class of fractal-scaling approaches under equal fractal signatures of the object and the background considered by the authors at an angle of mathematics of fractal structures and labyrinths. The concept of fractal labyrinths allows you to identify the parameters of complex stochastic nanostructures.

ACKNOWLEDGMENT

The authors are grateful to Prof. G. B. Khomutov (Moscow State University, Physics Department) for providing the initial images of microstructures of modern nano-modified materials and to Ph.D. S. V. Krupenin (Kotelnikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences) for technical assistance.

REFERENCES

1. Potapov, A. A., Y. V. Gulyaev, S. A. Nikitov, A. A. Pakhomov, and V. A. German, "Noveyshie metody obrabotki izobrazheniy," *Newest Images Processing Methods*, 496, Fizmatlit Publication, Moscow, 2008.
2. Grachev, V. I., A. A. Potapov, and V. A. Potapov, "Fractal Labyrinths," *RENSIT*, Vol. 3, No. 2, 103–108, 2011.
3. Potapov, A. A. and V. I. Grachev, "Fractal labyrinths: Physics and fractional operators," *PIERS Proceedings*, 1584–1587, Moscow, Russia, August 19–23, 2012.
4. Moran, P. A. P., "Additive functions of intervals and hausdorff measure," *Proc. Cambridge Philis. Soc.*, Vol. 42, 15–23, 1946.
5. Pesin, Y. B. and H. Weiss, "A multifractal analysis of equilibrium measures for conformal expanding maps and Markov Moran geometric constructions," *J. Stat. Phys.*, Vol. 86, 233, 1997.
6. Cristea, L. L. and B. Steinsky, "Curves of infinite lenght in 4×4 -labyrinth fractals," *Geom. Dedicata.*, Vol. 141, 1–17, 2009.
7. Kuratovsky, K., *Topologiya (Topology)*, Vol. 2, Mir Publication, Moscow, 1969.
8. Hilfer, R. (Editor), *Application of Fractional Calculation in Physics*, 472, World Scientific, Singapore, 1999.
9. Uchaykin, V. V., "Metod drobnykh proizvodnykh," *Method of Fractional Derivatives*, 512, ArtiShok Publication, Ul'yanovsk, 2008.
10. Lizorkin, P. I., *Fractional Integration and Differentiation*, in Hazewinkel, Michiel, Encyclopaedia of Mathematics, Springer, 2001.
11. Zaslavsky, G. M., "Gamiltonov khaos i fraktal'naya dinamika," *Hamiltonian Chaos and Fractal Dynamics*, 472, NIZ RKhD Publication, Moscow-Izhevsk, 2010.
12. Metzler, R. and J. Klafter, "The Random Walk's guide to anomalous diffusion: A fractional dynamics approach," *Phys. Rep.*, Vol. 339, 1–77, 2000.
13. Khomutov, G. B. and S. P. Gubin, "Interfacial synthesis of noble metal nanoparticles," *Mater. Sci. Eng. C*, Vol. 22, No. 2, 141–146, 2002.

Passive Super-Low Frequency Remote Sensing System

Nan Wang and Qiming Qin

Institute of Remote Sensing and Geographical Information System
Peking University, Beijing 100871, China

Abstract— Passive Super-Low Frequency (SLF) remote sensing system is presented here as a formidable geophysical tool for obtaining high-resolution information that can be related to resources and geological processes. As an intensively complementary method in the regular earth observation, the non-imaging passive SLF remote sensing system, including the hardware and software system, has been developed to acquire and process the SLF signal. And we propose a principle to interpret geo-object anomalies which can hardly be detected by other remote sensors. Results of field experiments indicate that the system can be successfully applied in the deep earth observation. We demonstrate that the passive non-imaging SLF remote sensing is a rapidly advancing method in logistically challenging areas with hundreds of meters to kilometers.

1. INTRODUCTION

Regular remote sensing techniques with respect to visible, infrared and microwave bands utilize the reflective and radiating electromagnetic information. These electromagnetic waves, however, are often reluctant to penetrate into the ground within tens of meters due to the strong attenuation in the near surface. Recently, many advanced sensors or remotely sensed systems are being designed and developed to enlarge the remote sensing band ranges, such as Super-Low Frequency (30Hz-300Hz) in Figure 1. Over the past decade, the Super-Low Frequency remote sensing has undergone a revolution driven by its increasing use in geological applications and the resource exploration[1]. The SLF electromagnetic wave, originating from natural source (passive) or controlled-source (active) within a general range of 3–3000 Hz, propagates through the surface of the earth a few kilometers deep, and then the correspondingly reflective, inductive or radiating wave carries abundant underground information. We mainly focused on the passive remote sensing by developing a SLF remote sensing system including the hardware and software system to collect, process and analyze the SLF signal with the purpose of the underground information extraction.

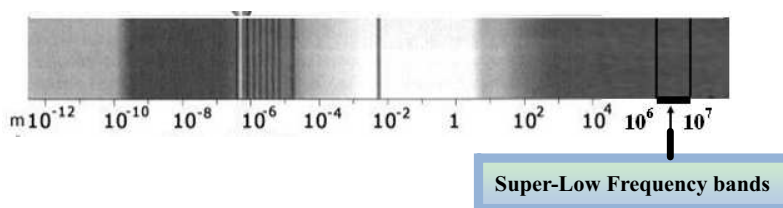


Figure 1: The location of the SLF electromagnetic spectrum.

2. PASSIVE SLF REMOTE SENSING SYSTEM

2.1. Hardware System

The success of the non-imaging SLF remote sensing is largely determined by the quality of data acquisition. Therefore, the hardware system can be divided into three parts: a high-resolution SLF remote sensor, the host computer and the power supply in Figures 2(a), 2(b), 2(c).

The SLF electromagnetic signal detector is a high-resolution sensor used to detect the passive SLF signal, and then the signal would be enlarged and filtered by multi-scale amplifying circuits and the low-pass filter circuit. After that, the analog signal in time domain is in progress with A/D transformation. A frequency sweep circuit under the control of the host computer can be utilized to digitally record high-quality SLF frequency spectrum data sets at every frequency point. In the third part, the digital signal should be processed by FFT. By means of the Frequency-Depth model [1, 2], we can calculate the Amplitude-Depth curves of SLF signal to be ultimately interpreted. The SLF curve will be in the real-time display on the computer screen. The horizontal axis is the Depth/m as well as the vertical axis for SLF signal magnitude. The characteristics of anomalies we are interested in can be interpreted by analyzing the SLF amplitude changes with respect to targets at their depths. The diagram is in Figure 3.



Figure 2: The hardware part of SLF remote sensing system.

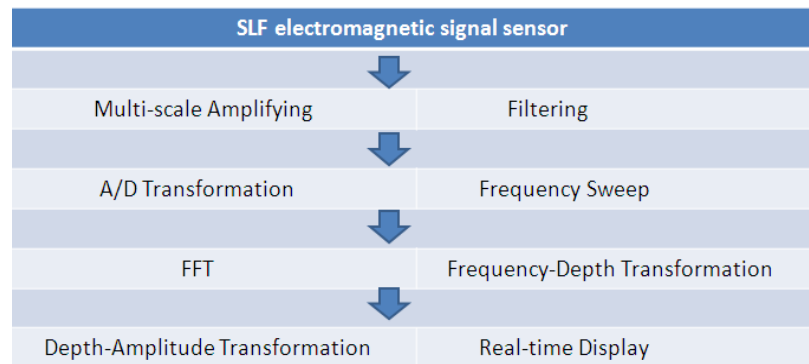


Figure 3: The workflow chart of the hardware system.

2.2. Software System

Passive SLF electromagnetic signal is non-stationary signal including the random noise, the electromagnetic interference and effective underground information. We have carried out a series of research on data processing in order to extract convincing SLF characteristics and parameters corresponding to targeted anomalies based on which we put forward a principle of SLF remote sensing interpretation.

We have designed the SLF processing and interpretation modules in charge of the pre-processing, data evaluation, de-noising and interpretation in Figure 4. The pre-processing module contains data format transformation, stacking, filtering and normalization; the data evaluation module accounts for assessing the SLF signal repeatability and stability judged by the correlating coefficients exceeding 0.95; the de-noising module aims at filtering the high-frequency random noise and industrial electromagnetic interference in SLF signal using targeted processing sub-module; the interpretation module builds the maps or profiles of underground anomalies by using the typical geo-object databases combined with geological materials. All the above modules help the SLF remote sensing contribute to identify and extract representative underground anomalies so as to realize the enormous expansion of the earth observation.

3. PRINCIPLE OF THE PASSIVE SLF REMOTE SENSING TECHNIQUE

We have summarized a SLF remote sensing principle in practical application illustrated in Figure 5. It covers three parts as follows:

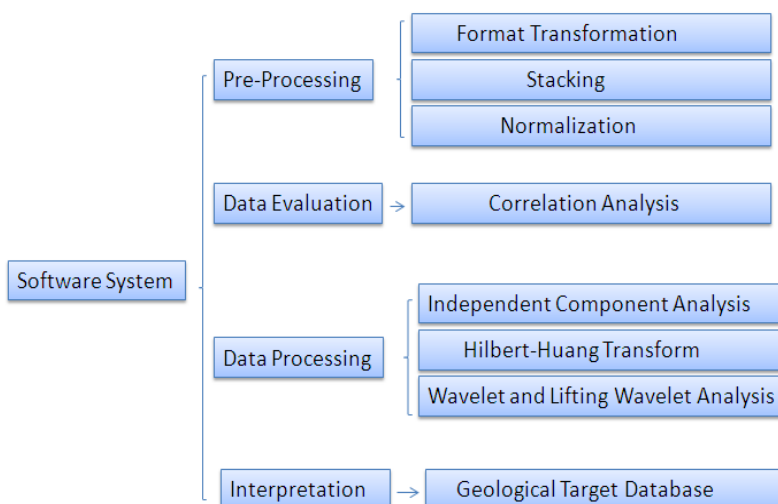


Figure 4: The SLF processing and interpretation software structure.

3.1. Data Acquisition

In the field experiments, we should logistically arrange the SLF hardware system sites or measuring lines according to GPS and geological materials. And then other parameters should be set such as the detector direction, step interval, magnifying parameter and depth range.

3.2. Data Processing

The raw materials should be transformed into readable files. For every high quality SLF curve, we should take measures to stacking, normalization and de-noising. The Independent Component Analysis (ICA), Hilbert-Huang Transform (HHT), Wavelet and Lifting Wavelet Analysis would be integrated to remove random noise and electromagnetic interference.

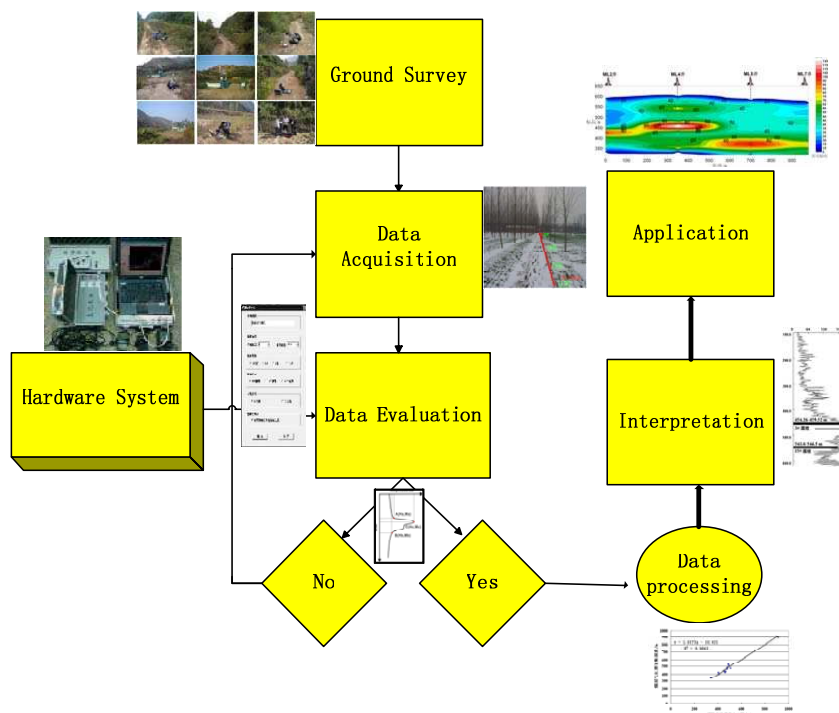


Figure 5: The principle of the passive SLF remote sensing technique.

3.3. Geological Interpretation

The SLF signal is in the close connection with stratum physical properties and target identification. The attenuation is commonly little when the SLF electromagnetic wave penetrates deep in magneto-tellurics theory. Especially when the SLF signal meets high resistance layers, it will be intensively reflective and vice versa. Therefore, we demonstrate that the high SLF amplitude corresponds to high resistance layers (high density rocks, dense stratum and oil-gas reservoirs). In contrast, the lower SLF amplitude is regarded as the characteristics of high porosity and loose lithology (sandstone, clay and underground mineralized water). After the Frequency-Depth transform, we can show a passive SLF monitoring curve to identify underground anomalies. Up until now, the passive Super-Low Frequency (SLF) remote sensing system have successfully been applied to geothermal well surveys, petroleum-gas reservoir prospecting, fault identification and mining industry [3].

4. CONCLUSION

A fast and powerful non-imaging remote sensing technique using the passive SLF detecting system is developed to resolve deeper geological structures or geo-objects rather than other regular remote sensing methods. We have developed the hardware system to obtain high quality and repeatable data, and then effective underground information is deprived by the software. A practical principle is successfully proved applicable in resource exploration and other geological fields. Further numerical and theoretical simulations will be studied.

REFERENCES

1. Burrell, G. A. and L. Peters, Jr., "Pulse propagation in lossy media using the low-frequency window for video pulse radar application," *Proceedings of the IEEE*, Vol. 67, 981–990, 1979.
2. Qin, Q. M., B. S. Li, R. B. Cui, H. B. Jiang, Q. P. Wang, Z. X. Zhang, and N. Zhang, "Analysis of factors affecting natural source SLF electromagnetic exploration at geothermal wells," *Chinese Journal of Geophysics-Chinese Edition*, Vol. 53, 685–694, 2010.
3. Hou, G., X. Qian, S. Wang, and J. Sun, "Possibility of the application of the ultra-long electromagnetic wave remote sensor to marine geological exploration," *Acta Geologica Sinica-English Edition*, Vol. 74, 391–394, 2000.

Multi-band Circular Patch Antenna for Wideband Application

Ebrahim Sailan Aabidi, M. R. Kamarudin, T. A. Rahman, and Hashimu Uledi Iddi

Wireless Communication Centre (WCC), Universiti Teknologi Malaysia

UTM Skudai, Johor 81310, Malaysia

Abstract— This paper presents a multi-band circular patch antenna for wideband applications. The antenna operates at 1.5, 2.8, 4.0, 4.8, 7.5 and 8.8 GHz frequency bands. The antenna has been designed, simulated and fabricated on an FR4 substrate with dielectric constant (ϵ_r) of 4.4 and thickness of 1.6 mm. A dual feed approach has been used in the design of the proposed wideband antenna. The multi-band circular patch antenna is smaller than the conventional antenna for wideband application. The final optimized design is $50 \times 50 \text{ mm}^2$. The measured and simulated return loss results and simulated radiation pattern results of proposed antenna has been presented. The dual feed improves the gain of the proposed antenna. The measured and simulated results show that the proposed antenna provides good performance in term of return loss and radiation pattern for wideband applications.

1. INTRODUCTION

A lot of researches have been done for the different applications of portable wireless communication devices, such as notebook computers and mobile phone handset antennas with low manufacturing cost and low-profile planar structures. One of the advantages of a multi-frequency band operation includes accommodating multifunctional services, such as instance video, voice and data transmissions for mobile communications.

It has been reported in [1] the suitability of multi-band wideband antenna for many applications because it requires less space for composition and less cost compared to the use of different antenna for each frequency band [2]. In recent years, a lot of low-profile multi-band wideband antennas have been developed for wireless applications [1–6], however, most of the previous antenna have large size and few bands for wideband application.

This paper presents the multi-band circular patch antenna cover (1.27–1.51 GHz), (2.66–2.83 GHz), (3.8–4.1 GHz), (4.74–4.86 GHz), (7.5–7.65 GHz) and (8.8–8.92 GHz) in the frequency bands for Radio location application. The antenna design and simulation has been done using CST Microwave Studio software 2012 software package based on finite integral method.

2. ANTENNA DESIGN

The multi-band circular patch antenna is designed to operate at 1.5, 2.8, 4.0, 4.8, 7.5 and 8.8 GHz frequency bands. The geometry of the proposed multi-band circular patch antenna is shown in Figure 1. The antenna consists of circular patch with ring slot fed by double fed. A dual feed approach used to improve the gain of the proposed antenna. The antenna is printed on the FR4 substrate with dielectric constant of 4.4, loss tangent of 0.02, thickness of 1.6 mm. The optimized dimension of the proposed antenna is $50 \times 50 \times 1.6 \text{ mm}^3$. The parameters of circular patch are as follows: $A = 50 \text{ mm}$, $B = 50 \text{ mm}$, $a = 10.5 \text{ mm}$, $b = 13 \text{ mm}$, $W = 50 \text{ mm}$, $L = 50 \text{ mm}$, radius = 18 mm, and feed line width = 1.5 mm. The prototype of the circular patch antenna is shown in Figure 2.

3. RESULTS AND DISCUSSION

The comparison of the simulated and measured return loss results of the proposed antenna has been shown in Figure 3. It can be seen that the resonances of the proposed antenna are 1.5, 2.8, 4, 4.8, 7.5 and 8.8 GHz. The measured frequency bandwidths obtained of the proposed antenna were 240 MHz (1.27 to 1.51 GHz), 170 MHz (2.66 to 2.83 GHz), 30 MHz (3.8 to 4.1 GHz), 120 MHz (4.74 to 4.86 GHz), 150 MHz (7.5 to 7.65 GHz) and 120 MHz (8.8 to 8.92 GHz), while the simulated bandwidth obtained were 240 MHz, 170 MHz, 27 MHz, 90 MHz, 92 MHz and 95 MHz. It can be seen that the measured and simulated results have a similar result. The experimental results indicate wide impedance bandwidths (VSWR < 2) of 18% (1.27–1.51), 6.2% (2.66–2.83), 7.6% (3.8–4.1), 2.55% (4.74–4.86), 1.93% (7.5–7.65) and 1.35% (8.8–8.92) in the specified frequency bands.

Figure 4 shows the E -plane and H -plane radiation patterns at 1.5, 2.8, 4, 4.8, 7.5 and 8.8 GHz of the proposed antenna. The simulation gain obtained of the circular patch antenna were 4 dB at

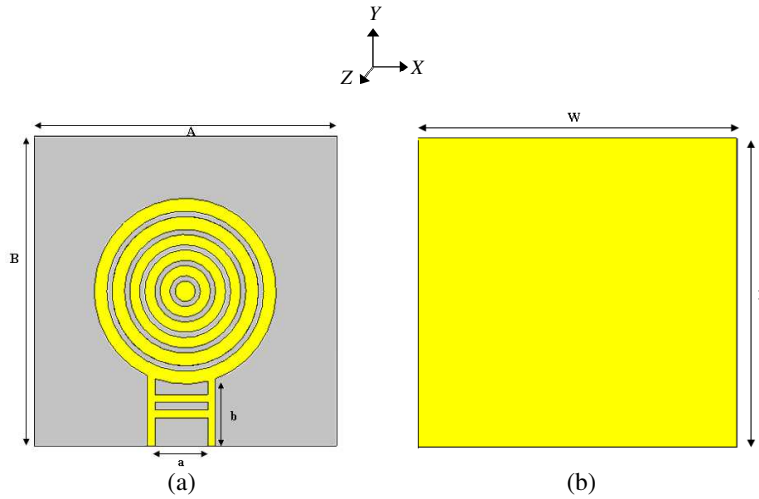


Figure 1: Geometry of the proposed P-shaped monopole antenna. (a) Front view. (b) Back view.

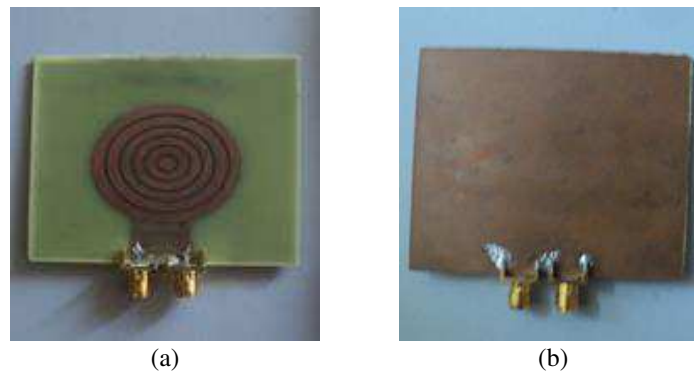


Figure 2: Prototype of the proposed antenna. (a) Front view. (b) Back view.

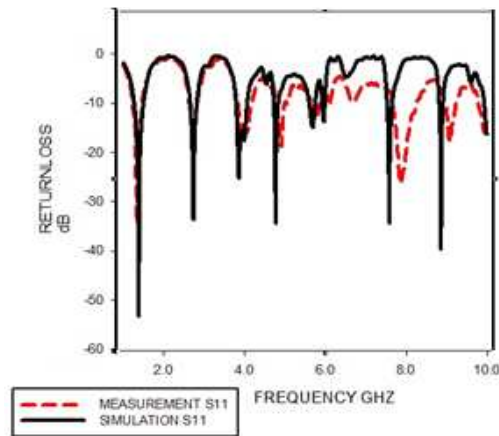


Figure 3: Measured and simulated return loss of the proposed antenna.

1.5 GHz, 4.5 dB at 2.8 GHz, 3.2 dB at 4 GHz, 2.1 dB at 4.8 GHz, 6.8 dB at 7.5 GHz and 1.5 dB at 8.8 GHz.

(4.74–4.86 GHz) 0.15 GHz (7.5 to 7.65 GHz) and 0.12 GHz (8.8 to 8.92 GHz). The dual feed improves the gain of the circular patch antenna which gives 4, 4.5, 3.2, 2.1, 6.8 and 1.5 dB at 1.5, 2.8, 4, 4.8, 7.5 and 8.8 GHz, respectively.

ACKNOWLEDGMENT

The authors would like to thank to Ministry of Higher Education (MOHE) and UTM GUP (vote 01H00 and vote 00G36) for sponsoring this work. Besides, thanks to the members of Wireless Communication Center (WCC), UTM for helping and providing resources to enable this work to be completed.

REFERENCES

1. Li, R. L., B. Pan, J. Laskar, and M. M. Tentzeris, "A novel low-profile broadband dual-frequency planar antenna for wireless handsets," *IEEE Transactions on Antennas and Propagation*, Vol. 56, No. 4, 1155–1162, April 2008.
2. Pan, C. Y., T. S. Horng, W. S. Chen, and C. H. Huang, "Dual wideband printed monopole antenna for WLAN/WiMAX applications," *IEEE Antennas and Wireless Propagation Letters*, Vol. 6, 149–151, 2007.
3. Porcino, D. and W. Hirt, "Ultra-wideband radio technology: Potential and challenges ahead," *IEEE Communications Magazine*, Vol. 41, No. 7, 66–74, July 2003.
4. Yazdandoost, K. Y. and R. Kohno, "Ultra wideband antenna," *IEEE Communications Magazine*, Vol. 42, No. 6, S29–S32, June 2004.
5. Qu, F., J. Li, L. Yang, and T. Talty, "Measured channel capacity of SIMO-UWB for intra-vehicle communications," *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, 3056–3060, April 11–15, 2011.
6. Yang, L. and G. B. Giannakis, "Analog space-time coding for multiantenna ultra-wideband transmissions," *IEEE Transactions on Communications*, Vol. 52, No. 3, 507–517, March 2004.

The Comparative Analysis of UWB Antennas with Complementary Characteristics: A Functionality in FS and Applicability for the Usage Close to Tissues

Tommi Tuovinen, Markus Berg, and Erkki Salonen

Centre for Wireless Communications, University of Oulu, Finland

Abstract— The present paper considers planar ultra wideband (UWB) antennas with complementary characteristics and their operation close to tissues. The dipole and slot antennas are first examined by forming they appear as complementary structures according to the Babinet's principle (BP). The influence of a substrate on the complementarity is evaluated and the addition of a substrate is noticed to affect radiators to diminish the capability to appear as ideal complements. A dipole-slot pair is re-designed to have corresponding S_{11} bandwidth to cover the lower UWB band of 3.24–4.74 GHz and the complementary operation is acquired according to the pattern duality. The antennas are concluded to have complementary characteristics based on their radiation properties and current distributions, even though the BP is not ideally satisfied. The performance close to tissues is analysed and discrepancies shown in terms of mismatch and body loss. Although free space performance is observed to appear almost equal (excluding a pattern duality), the performance is found to differ close to tissues.

1. INTRODUCTION

The research work for wireless body area networks (WBANs) has been comprehensive during recent years. The publication of an international WBAN standard 802.15.6 [1] by the Institute of Electrical and Electronics Engineers (IEEE) in 2012 was a step toward coordinated guidelines in the field. As an introduced WBAN technology, Ultra wideband (UWB) [2] has alignment for 3.2448–4.7424 GHz (low band) and 6.24–10.2336 GHz (high band). For the WBAN usage, antenna properties are proposed to investigate in the proximity of a body. Analysis of various broadband (or UWB) antennas close to tissues are of interest. A comprehensive overview to present UWB antenna characteristics and suitability for applications is presented in [3]. There are also hundreds of papers in the open literature showing various UWB antennas or describing the antenna performance for both the part of UWB band or covering the entire Federal Communications Commission (FCC) [2] UWB band of 3.1–10.6 GHz. Further, multiple UWB impulse radar (IR) or broadband antenna structures, but also the requirements for designing them can be found in many books, e.g., in [4–6].

In this paper, the operation of two complementary antennas (dipole and slot) based on the Babinet's principle (BP) for the antenna input impedance and pattern duality is considered: we examine the possibility to implement planar UWB antenna structures with complementary characteristics. Also their applicability and discrepancies for the UWB WBAN on-body usage is evaluated. The well-known dipole is traditionally used as a resonant-type antenna on narrow band applications. However, many articles are published considering the properties of dipoles for the implementation of wideband usage, e.g., in [7, 8] in recent years. Respectively, various wideband slot antennas are studied such as in [9, 10]. Self- and quasi-complementary structures for UWB are earlier analysed [11, 12] and some comparing antennainvestigations are already reported, e.g., [13, 14]. However, the comparative study if these dipole and slot antennas can appear as planar complementary UWB structures for each other and on the other hand how are they performed close to tissues is not carried out according to the best knowledge of authors. Simulations are carried out by Computer Simulation Technology (CST) software.

2. ANTENNAS: THEORY, COMPLEMENTARITY AND EVALUATED DESIGNS

2.1. Antennas Based on Babinet's Principle

The Babinet's principle (BP) can be used to clarify the impedance for a complementary antenna [4, 9]. The antenna input impedance Z_{in} that includes the input resistance R_{in} showing dissipation in terms of power, and ohmic losses associated with heating on the antenna structure can be defined as $Z_{in} = R_{in} + jX_{in} = (R_r + R_L) + jX_{in}$ [5] where R_r and R_L are a radiation and loss resistance of the antenna and X_{in} is an antenna reactance. According to the BP, if one has the antenna made of a metal that holds $Z_{in} = Z_{metal}$, the dual structure can be achieved by replacing the metal with air, and vice versa: this results to the complementary antenna which holds

$Z_{in} = Z_{air}$ [4]. The product of complementary impedances can be given as constant $Z_0^2/4$ (Z_0 is the wave impedance in free space) [5, 9]. Fig. 1(a) shows the example of complementary dipole-slot antenna pair, which is given in [4, 5]. The antennas hold that the slot impedance can be derived by $Z_{slot} = Z_0^2/4Z_{dipole}$ [4, 9] where $Z_0 = \sqrt{\mu_r\mu_0}/\sqrt{\epsilon_r\epsilon_0}$ as μ and ϵ are the permeability and permittivity of medium (sub-indices: r refers to relative value and zero to that of free space).

The dipole shown in Fig. 1(b) (above) is resonated at 3 GHz (the total length of 43 mm and the width of 4 mm) for which -10 dB S_{11} bandwidth is 2.80–3.26 GHz, and the real part of $Z_{dipole} = 61.3 \Omega$ at 3 GHz is observed. The complementary slot shown in Fig. 1(b) (below) is realized with the size of slot length = 43 mm the width = 4 mm. The impedance of slot appears in the range of $Z_{slot} = 450\text{--}600 \Omega$ at 3 GHz, which depends of the size of the extension of metal area in xy -plane while the slot area is kept constant. By substituting the theoretical values ($Z_A = 73 + j42.5 \Omega$ for $\lambda/2$ dipole and $Z_0 = 376.7 \Omega$ in free space) to (2), $Z_{slot} = 486.0 \Omega$ is attained. It seems that these are complementary structures according to the BP.

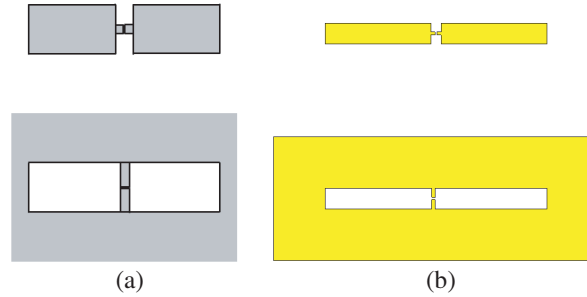


Figure 1: (a) Theoretical complementary antennas based on the BP: dipole (above) and slot (below) [4], and (b) designed antennas.

A wearable capability is realized by using the well-known 1.6 mm thick FR4 substrate with ϵ_r of 4.3. The addition of substrate causes the resonant of dipole (i.e., centre frequency f_c) introduced in Fig. 1(b) to appear at 2.4 GHz such that it covers the band of 2.2–2.6 GHz, and $Z_{dipole} = 49.3 \Omega$ at 2.4 GHz is attained. By executing the same substrate addition for the slot, the impedance range of $Z_{slot} = 340\text{--}500 \Omega$ at 2.4 GHz is noticed depending on the size of metal over the edges of the slot. For the calculation of the theoretical Z_{in} for the complement, the effective dielectric constant $\epsilon_{eff} = (\epsilon_r + 1)/2$ [15] for the antenna substrate must be considered due to the fact that there are the substrate below and air above the antenna radiator. By using $\epsilon_{eff} = 2.65$ and $Z_0 = 231.8 \Omega$ ($\mu_r = 1$, $\mu_0 = 1.26 \cdot 10^{-6}$, $\epsilon_0 = 8.85 \cdot 10^{-12}$), $Z_{slot} = 1090 \Omega$ is attained. Without the inclusion of ϵ_{eff} (using ϵ_r of substrate), $Z_{slot} = 671.6 \Omega$ is observed. The BP is concluded to be applicable and match for the formulation of complementary antennas that are not using the substrate, but even though the structures are complementary according to the theory they do not cover the same S_{11} bandwidth (the complement slot over-tune the dipole). In the case of inclusion of the substrate in the design, determination of the correct wave impedance is observed to be slightly confusing; whether one should use ϵ_r or ϵ_{eff} substrate for defining the wave impedance for the medium. The better impedance match for the complementary structure is derived with ϵ_r however it is not absolutely realistic situation. Apparently, complementary structures based on the BP can be achieved with the inclusion of a substrate but they do not operate in the same S_{11} bandwidth.

2.2. Complementary Wearable Antennas

Based on the considerations in Section 2.1, the antennas were re-designed to form planar UWB antenna structures that appear the most equal each other by generating equal -10 dB S_{11} bandwidth for the lower UWB band. The result of evaluation is shown in Fig. 2(a) where the antenna simulation models are depicted. Both ribbons have the size of $8 \times 12.4 \text{ mm}^2$ and the feeding lines of $0.7 \times 0.6 \text{ mm}^2$ are brought to a 1 mm-long feeding gap where a 50- Ω discrete port is used for an excitation. In comparison with the wavelength λ , free space values of $\lambda_{3.2 \text{ GHz}} = 9.4 \text{ cm}$, $\lambda_{4 \text{ GHz}} = 7.5 \text{ cm}$ and $\lambda_{4.7 \text{ GHz}} = 6.4 \text{ cm}$ are found. The total height of dipole is noticed to appear roughly 28% at 3.2 GHz, 36% at 4 GHz, and 42% at 4.7 GHz of the wavelength. For evaluating the UWB slot antenna with a FR4 substrate with the equal S_{11} to dipole, a slot within the coppering of the size of $43 \times 4 \text{ mm}^2$ is required. It is observed that the designed slot antenna is exceptional sensitive for the change of width at the both ends of slot area (i.e., the edges of the slot, i.e., slit area in the direction of $+/-x$ axis).

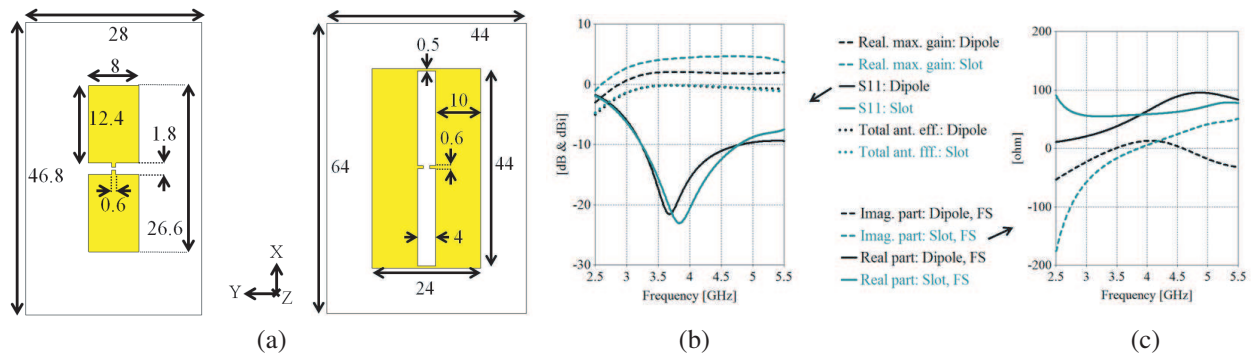


Figure 2: Designed antennas in xy -plane (a) dipole (left) and slot (right) antennas. Comparing (b) S_{11} , gain and efficiency in free space (-10 dB S_{11} bandwidth is 3.23–4.77 GHz), and (c) the real and imaginary parts of Z_{in} for the dipole and slot in free space.

3. PERFORMANCE RESULTS

3.1. Free Space Performance

The free space bandwidth in terms S_{11} is targeted to appear equal with each other. Fig. 2(b) shows simulated S_{11} with the realized gain G_{max} and total antenna efficiency e_0 while Fig. 2(c) presents Z_{in} behaviour. Exceptional equal S_{11} bandwidth and promising almost identical e_0 is found. The real-valued Z_{in} presents the travelling wave nature of a wideband antenna. G_{max} of 2.06 dBi (at 3.80 GHz) and 4.64 dBi (at 4.51 GHz) are observed for the dipole and slot. Rather smooth gain is good property for an antenna used in IR systems. The higher gain for slot can be explained due to the directivity of slot; the small maximum peaks of patterns are attained in $+/-z$ directions. In theory, directivity for both $\lambda/2$ and ideal dipole are 2.15 dBi and 1.76 dBi [4]. The proper agreement from the patterns in Fig. 3(a) can be observed with respect to the assumptions based on the duality: the interchange of fields. The dipole generates E -field that is vertically polarized while the slot appear as horizontally polarized. The currents of the dipole are found to appear linear and equal in both wires in magnitude as they should appear.

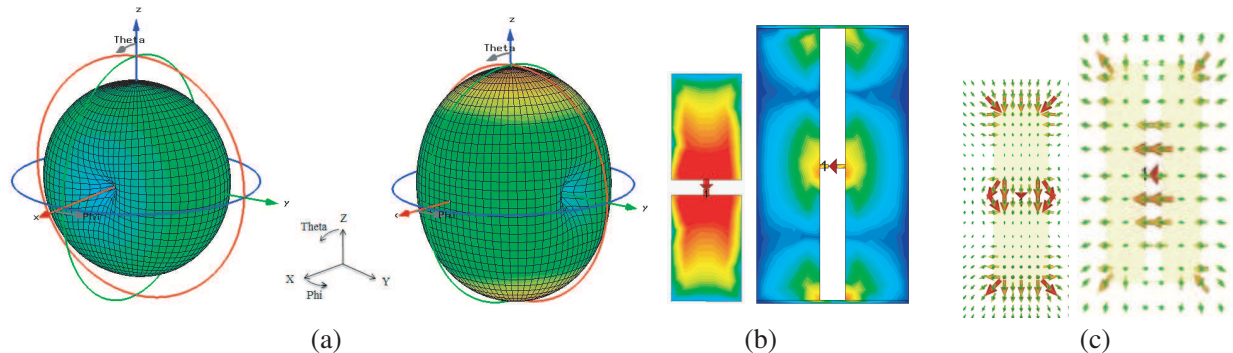


Figure 3: Comparing (a) free space patterns at 4 GHz (realized gain): dipole (left) and slot (right), (b) instantaneous current distributions [A/m], and (c) E -field vectors at 4 GHz at 0° phase. In (a) scaling is from -35 dBi (blue) to 5 dBi (red) and in (b) the colour range is a logarithmic from 0 A/m (blue) to 5 A/m (red) Coordinates in (b)–(c) are the same in Fig. 2.

3.2. Performance Close to Tissues

Operation through the range of $0 \text{ mm} \leq \text{antenna-body distance } D \leq 20 \text{ mm}$ in steps of 1 mm above a tissue surface is examined. The layered model including skin, fat, muscle and bone tissues described, e.g., in [8] with the size of size of 300 mm (X) \times 300 (Y) \times 35 mm (Z) is decided to use. The antenna is positioned in the middle on the top of the model. The dispersion for the tissues in UWB is determined based on the 2nd order Debye model, which principle is described, e.g., in [16]. The used values for tissues can be found in [8, 16].

Figures 4(a)–(b) show the matching close to the modelled tissues at various distances, while body and mismatch losses are shown in Fig. 4(c) (derived as in [18]). Differences in matchings' are observed on contact: the dipole shows a smooth wide matching varying the lower -10 dB band edge

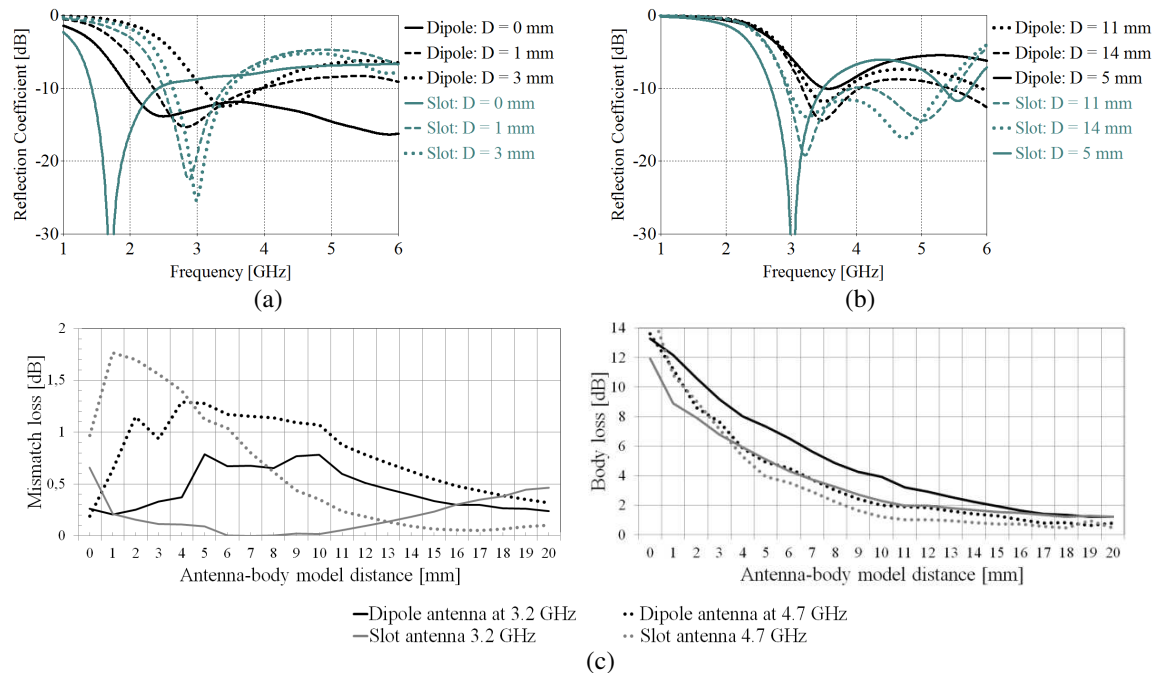


Figure 4: On-body matching for the dipole and slot through (a) $0 \text{ mm} \leq D \leq 3 \text{ mm}$ and (b) $5 \text{ mm} \leq D \leq 14 \text{ mm}$. Fig. (c) shows the proportion of mismatch and body loss at 3.2 GHz and 4.7 GHz as a function of antenna-model distance through $0 \text{ mm} \leq D \leq 20$.

1.25 GHz downwards to 1.98 GHz, while the slot appears like a downward shifted wideband resonant covering -10 dB band from 1.44–2.42 GHz. The difference on contact might be at least because of two reasons: 1) the pattern differences, since the power of slot propagates more strongly to $+/-z$ directions (the doughnut-shaped pattern for the dipole) therefore the stronger effect for the slot could be a rational explanation. 2) E -field of the slot is over the slot area as shown in Fig. 3(c). Hence, because of tissues in the proximity of the slit, the stronger change for Z_{in} of the slot than for the dipole is understandable. By considering higher distances, the rapid change for the slot is observed at $D = 1$ mm; lower band edge shifts upwards 1.06 GHz up to 2.5 GHz and after it the matching stays stable for $1 \text{ mm} \leq D \leq 5 \text{ mm}$ as values of $2.50\text{--}2.70 \text{ GHz} \leq \text{band} \leq 3.40\text{--}3.49 \text{ GHz}$ are acquired. After $D = 5$ mm the higher resonance of the slot start smoothly shift downwards while the lower band edge stays rather constant. The dipoles' matching is varied exceptionally smoothly upwards; the lower band edge shifts roughly 0.2 GHz per 1 mm through $D = 1\text{--}3$ mm such that the poorest matching is noticed at $D = 5$ mm. After $D = 5$ mm, the matching start to smoothly widen and the dipole covers -10 dB lower UWB band at $D = 17$ mm again. The behaviour of the dipole is found to appear that of expected. Corresponding findings are reported earlier, e.g., in [17]. The stable nature of the matching of slot seems to retain relatively low mismatch losses in comparison with the dipole, especially in lower frequencies. If the dipole is operated either very close or clearly farther (here: referring to the studied range of 0–20 mm) to tissues, mismatch losses remain low but there is still the challenging area in the middle from the matching point of view where losses increase. A performance close to free space after exceeding 15 mm distance to tissues was observed. The exact distance where an antenna is not affected from the perspective of tissues is challenging to give as being the sum of many factors; the tissue combination depending upon a body under test, the effect of a substrate, the size of antennas' reactive near-field, and so forth. The results can be contrasted to the field regions; the distance over the size of antennas' reactive near-field is good indicator for attaining the satisfactory antenna performance and operation, such as concluded in [17].

4. CONCLUSIONS

Complementary UWB antennas based on the Babinet's principle (BP) were examined in this paper. The complement structures were observed to be able to form but they do not operate in the same bandwidth. The addition of a substrate was noticed to affect radiators to reduce the capability to appear as ideal complements. Planar UWB antenna structures with complementary characteristics based on the pattern duality were observed to be possible to generate but they do not satisfy the

BP. The operation between the antennas was observed to vary in the close proximity of tissues: based on matching, strong variations for the slot are noticed in the studied range however body losses of both antennas are relatively close each other.

ACKNOWLEDGMENT

This work was supported in part by the Finnish Funding Agency for Technology and Innovations (Tekes) through Enabling Wireless Healthcare Systems (EWiHS) project and Infotech Oulu Doctoral Program. Also TES, HPY, Walter Ahlström, Emil Aaltonen and Tauno Tönning foundations are appreciated.

REFERENCES

1. IEEE standard for local and metropolitan area networks, IEEE 802.15.6-2012 — Part 15.6: wireless body area networks, 2012.
2. “First Report and Order in the matter of revision of Part 15 of the Commission’s rules regarding ultrawideband transmission systems,” Federal Communications Commission, ET-Docket 98-153, FCC 02-48, 2002.
3. Adamiuk, G., T. Zwick, and W. Wiesbeck, “UWB antennas for communication systems,” *Proc. of IEEE*, Vol. 100, No. 7, 2308–2321, Jul. 2012.
4. Stutzman, W. L. and G. A. Thiele, *Antenna Theory and Design*, 2nd Edition, 10–12, 37–43, 59–62, 250–252, John Wiley & Sons, The United States of America, 1998.
5. Balanis, C. A., *Antenna Theory Analysis and Design*, 3rd Edition, 64–69, 80–86, 699–701, John Wiley & Sons, Canada, 2005.
6. Hall, P. S. and Y. Hao, *Antennas and Propagation for Body-centric Wireless Communications*, 2nd Edition, 139–160, Artech House, Norwood, 2012.
7. Low, X. N., Z. N. Chen, and T. S. P. See, “A UWB dipole antenna with enhanced impedance and gain performance,” *IEEE Trans. Antennas Propag.*, Vol. 57, No. 10, 2959–2966, Oct. 2009.
8. Tuovinen, T., M. Berg, K. Yekeh Yazdandoost, E. Salonen, and J. Iinatti, “Human body effect on the polarization properties of the new UWB dipole antenna in UWB WBAN applications,” *Proc. 7th Int. Conf. BANs (BodyNets)*, 1–7, Norway, Sep. 2012.
9. Zhou, Y., “A new efficient method of design on slot antenna,” *Proc. Int. Conf. on Microw. Millimeter Wave Tech. (ICMMT)*, 1–4, May 2012.
10. Clementi, G., N. Fortino, J.-Y. Dauvignac, and G. Kossiavas, “Frequency and time domain analysis of different approached to the backing of an UWB slot antenna,” *IEEE Trans. Antennas Propag.*, Vol. 60, No. 7, 3495–3498, Jul. 2012.
11. Sonkki, M., M. Ferrando-Bataller, E. Antonino-Daviu, and E. Salonen, “Optimized dimensions of ultra wideband quasi-complementary antenna with switching capability,” *Proc. 4th Europ. Conf. Antennas Propag. (EuCAP)*, 1–4, Barcelona, Spain, Apr. 2010.
12. Mushiake, Y., “Self-complementary antennas,” *IEEE Trans. Antennas Propag. Magazine*, Vol. 34, No. 6, 23–29, Dec. 1992.
13. See, T. S. P. and Z. N. Chen, “Experimental characterization of UWB antennas for on-body communications,” *IEEE Trans. Antennas Propag.*, Vol. 57, No. 4, 866–874, Apr. 2004.
14. Wui, X. H. and Z. N. Chen, “Comparison of planar dipoles in UWB applications,” *IEEE Trans. Antennas Propag.*, Vol. 53, No. 6, 1973–1983, Jun. 2005.
15. Berg, M., M. Komulainen, E. Salonen, and H. Jantunen, “Frequency reconfigurable microstip-fed annular slot antenna,” *Proc. 1st Europ. Conf. Antennas Propag. (EuCAP)*, 1–6, Nizza, France, Nov. 2006.
16. Barnes, F. S. and B. Greenebaum, *Bioengineering and Biophysical Aspects of Electromagnetic Fields*, 58–59, 316–317, 340–343, Taylor & Francis Group, Boca Raton, 2007.
17. Tuovinen, T., T. Kumpuniemi, K. Yekeh Yazdandoost, M. Hämäläinen, and J. Iinatti, “Effect of the antenna-human body distance on the antenna matching in UWB WBAN applications,” *7th Int. Symp. Med. Inform. Commun. Technol. (ISMICT2013)*, Japan, Mar. 2013.
18. Berg, M., M. Sonkki, and E. Salonen, “Absorption loss reduction in a mobile terminal with switchable monopole antennas,” *IEEE Trans. Antennas Propag.*, Vol. 59, No. 11, 4379–4384, Nov. 2011.

Optimized Reflector Position for Vlasov Antennas

H. M. El Misilmani^{1,2}, M. Al-Husseini², K. Y. Kabalan¹, and A. El-Hajj¹

¹American University of Beirut, Beirut 1107 2020, Lebanon

²Lebanese Center for Studies and Research, Beirut 2030 8303, Lebanon

Abstract— This paper presents a Vlasov antenna with optimized reflector position and angle suitable for high power microwave applications. With the proposed configuration, the reflector is directly attached to the waveguide, which is an advantage and makes it simpler to radiate in the direction of the axis of the waveguide. Bevel-cut and Step-cut Vlasov antennas, designed for operation at 3 GHz, are used to validate the effect of the reflector. In addition to proper radiation of the direction of maximum radiation, the optimized reflector results in increased antenna gain and reduced half-power beamwidth.

1. INTRODUCTION

High Power Microwave (HPM) sources, such as the Backward-Wave Oscillator (BWO), the gyrotron, and the vircator (virtual cathode oscillator), generate power in cylindrically symmetric transverse electric TE_{0n} or transverse magnetic TM_{0n} modes. The side lobe generation, gain reduction, and inefficient power loading on the antenna aperture, make these modes unsuitable for driving conventional antennas. This gave the idea of using mode converters at the output of these sources to convert these modes into a plane-parallel linearly polarized beam. Vlasov antenna is one of the most known mode converters used. The well-known Vlasov types are the Step Cut and the Bevel Cut antennas [1]. The step cut, originally suggested by Vlasov, has sharp edges and therefore may suffer from electrical breakdown when radiating HPM. The bevel cut, which was later suggested by Nakajima, avoids the sharp points of step cut, and as a result has a more suitable shape for HPM applications. However, a Vlasov antenna with either the bevel or step cut has its maximum radiation shifted by some angle with respect to the axis of the waveguide.

In [2], a comparison of the performance of bevel-cut and step-cut Vlasov antennas in HPM is conducted, concluding that the bevel cut has better performance in such applications. Other studies focused on increasing the gain of bevel-cut and step-cut Vlasov antennas. In [3], a reflector is added to a bevel-cut Vlasov antenna to increase its gain and to obtain more directive radiation. In [4], two methods are proposed for increasing the gain of a bevel-cut antenna, one using a parabolic cylinder reflector, and the second using a horn. In [5], the step cut is studied in the presence of a parabolic reflector. However, none of these studies considered bringing back the maximum radiation along the axis of the waveguide.

In this paper, an optimized reflector position for Vlasov antennas is presented, which will help, with the proper rotation angle, to orient the generated waves along the $+Z$ direction, which is the axis of the waveguide in our case. In addition, with our proposed configuration, the reflector is directly attached to the waveguide structure, decreasing the size of the usual Vlasov antennas with reflectors, and eliminating the need of extra components to hold the waveguide and reflector together. The proposed reflector is applied to a bevel-cut and a step-cut Vlasov antennas to evaluate its performance. It could also be applied to the cut proposed in [6], where the same results will hold.

2. BEVEL-CUT VLASOV ANTENNA

Both step- and bevel-cut Vlasov antennas are the result of shaping the end part of a circular waveguide. For operation at 3 GHz, the used circular waveguide has a radius of 45 mm and a length of 300 mm.

2.1. Bevel-cut Vlasov Antenna without Reflector

A Vlasov antenna with a beveled cut is shown in Figure 1(a). The cut angle α is the single parameter available for optimization, and it has the main effect on the gain and radiation patterns of the antenna. The angle that maximizes the gain of the antenna is given by [1]:

$$\alpha = \sin^{-1}((\rho_{0n}\lambda)/(2\pi a)) \quad (1)$$

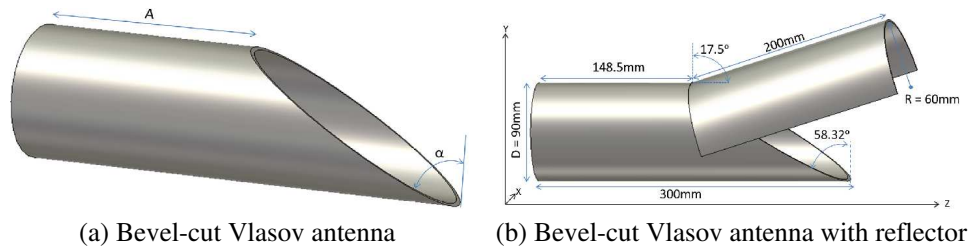


Figure 1: Bevel-cut Vlasov antenna: (a) without reflector and (b) with reflector.

where ρ_{0n} is the n -th root of the equation $J_0(\rho_{0n}) = 0$, λ is the wavelength, a is the inner radius of the waveguide, and J_0 is the Bessel function of the first kind and zeroth order.

For the TM_{01} circular waveguide designed for 3 GHz, $a = 4.5$ cm and $\lambda = 10$ cm. Also, $\rho_{01} = 2.405$, so the bevel cut will be calculated as follows:

$$\alpha = \sin^{-1}((2.405 \times 10)/(2\pi \times 4.5)) = 58.32^\circ \quad (2)$$

The highest gain according to the equation is obtained at a cut angle of 58.32° . This result has been verified by simulations using ANSYS HFSS [7]. For this angle, the resulting peak gain is 10.9 dB. The gain patterns, computed in CST MWS [8], are shown in Figure 3. Maximum radiation is obtained in the shifted direction corresponding to $\theta = \theta_m = 28^\circ$ and $\phi = 90^\circ$, computed using HFSS.

2.2. Bevel-cut Vlasov Antenna with Reflector

A reflector having the shape of a half hollow cylinder is attached to the bevel-cut Vlasov antenna as shown in Figure 1(b). The added reflector has the optimized values of 60 mm for the cylinder radius and a length of 200 mm (height of the cut cylinder). Upon rotating the reflector by a specific angle, it is seen that as the angle increases the shift angle approaches to origin. The initial bevel-cut Vlasov antenna gives a maximum computed gain of 10.3 dB, with the maximum radiation along the $\theta = 28^\circ$ and $\phi = 90^\circ$ direction. The optimized reflector angle for perfect direction along the $+Z$ axis is seen at angle of 17.5° . For this angle, the maximum radiation is back along the axis of the waveguide, i.e., $\theta = 0^\circ$ and $\phi = 90^\circ$.

This bevel-cut antenna operates at 3 GHz as shown in the reflection coefficient plot (S_{11}) shown in Figure 2. It has a gain of 10.9 dB and a reduced HPBW, as indicated in Figures 3(a) and 3(b). It is shown that, for the case with the reflector, the maximum radiation is redirected along the axis of the waveguide.

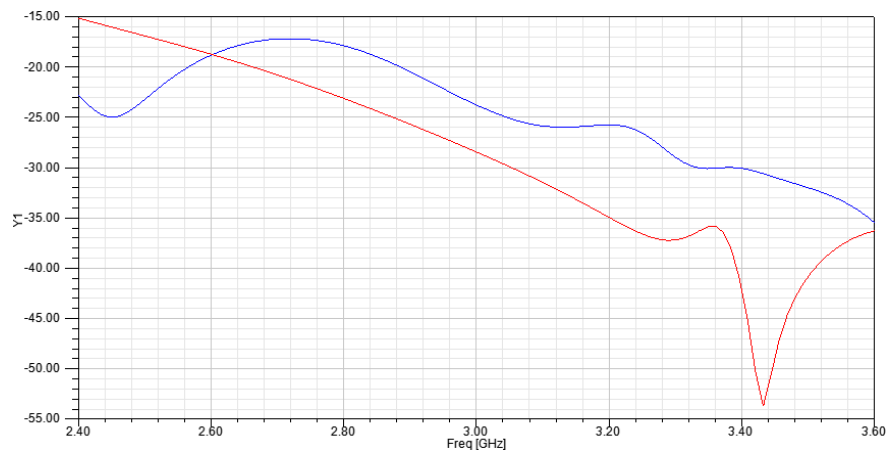


Figure 2: Reflection coefficient computed using HFSS, with no-reflector case shown in red, and with reflector in blue.

2.3. Verification of the Results Using CST

The results in Section 2.2 have been verified using CST. The gain patterns in the two cases (without and with reflector) are shown in Figures 4 and 5. As can be seen, the maximum gain of the antenna

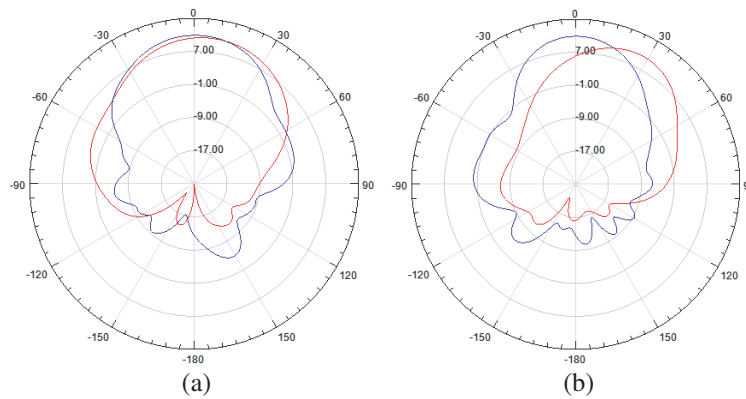


Figure 3: Simulated gain patterns computed using HFSS, with initial bevel-cut results shown in red and proposed design results in blue. (a) Red in the plane formed by the X -axis and the point of maximum radiation ($\theta = \theta_m$, $\phi = 90^\circ$), blue in $\phi = 0^\circ$ plane. (b) $\phi = 90^\circ$ plane.

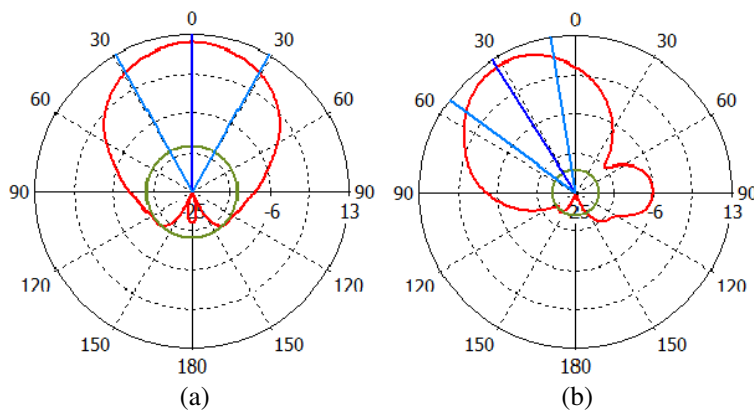


Figure 4: Bevel cut simulated gain patterns using CST without adding the reflector. (a) In the plane formed by the X -axis and the point of maximum radiation ($\theta = \theta_m$, $\phi = 90^\circ$). (b) $\phi = 90^\circ$ plane.

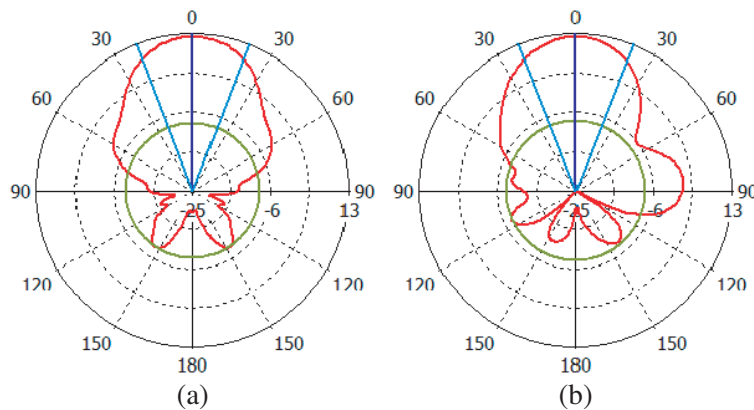


Figure 5: Bevel cut simulated gain patterns using CST after adding the reflector. (a) $\phi = 0^\circ$ plane. (b) $\phi = 90^\circ$ plane.

is redirected along the axis of the waveguide. The 3D gain patterns comparing the two cases, are shown in Figure 6. Furthermore, the peak gain has increased and HPBW has decreased, as listed in Table 1.

3. STEP-CUT VLASOV ANTENNA

The step cut is determined by two parameters, A and B , as indicated in Figure 7(a). The value of A is fixed at 148.5 mm, which is the same value obtained with the bevel cut after finding the

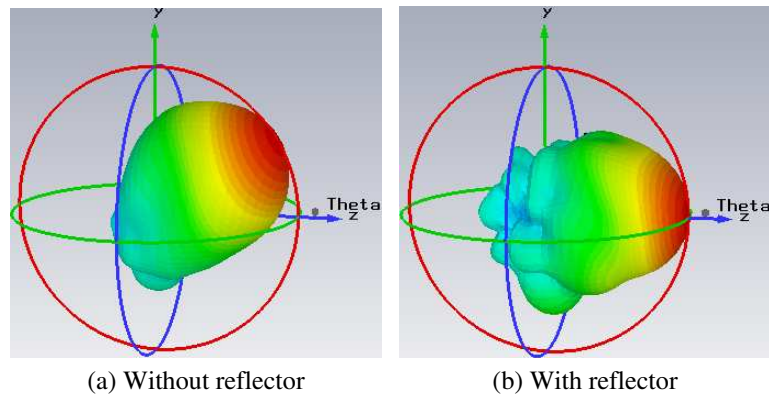


Figure 6: Bevel cut 3D gain patterns using CST: (a) without reflector and (b) with reflector.

Table 1: Comparison of the radiation characteristics of the bevel-cut antenna having and without having an added reflector computed using CST.

With/Without Reflector	HPBW ^o at $\phi = 0^\circ$ plane	HPBW ^o at $\phi = 90^\circ$ plane	3D Gain (dB)
Without Reflector	58.4	44.5	10.88
With Reflector	41.7	42.9	12

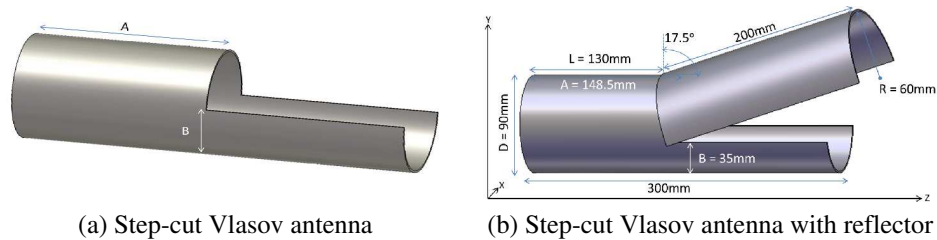


Figure 7: Step-cut Vlasov antenna: (a) without reflector and (b) with reflector.

angle α . For comparison purposes, the step-cut Vlasov is designed so that it has the same angle of maximum radiation obtained with the bevel cut, which is 28° . For this purpose, B is found to be 35 mm. The gain patterns of the step-cut Vlasov antenna, computed using HFSS in the $\theta = \theta_m = 28^\circ$ and $\phi = 90^\circ$ planes, are shown in Figures 8(a) and 8(b) respectively.

The same reflector used in Section 2.2 is then attached to the step-cut antenna as shown in Figure 7(b). Here, L is the distance between the waveguide port and the start of the reflector. By inspecting Figure 8(b), the concept of rotating the reflector is validated and maximum radiation

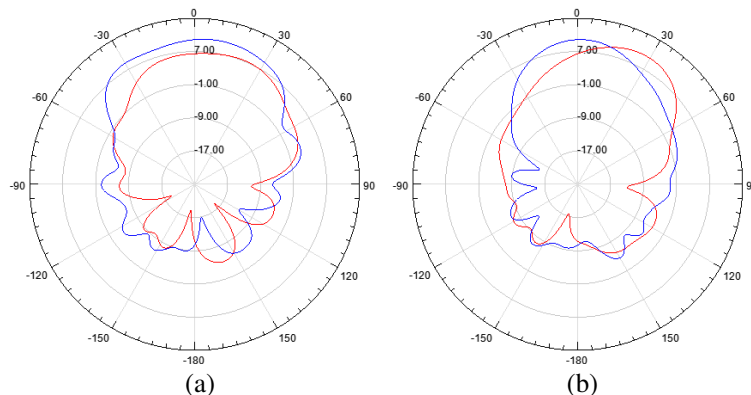


Figure 8: Step cut simulated gain patterns, with initial step-cut results shown in red and proposed design results in blue. (a) Red in the plane formed by the X -axis and the point of maximum radiation ($\theta = \theta_m$, $\phi = 90^\circ$), blue in $\phi = 0^\circ$ plane. (b) $\phi = 90^\circ$ plane.

is obtained along the waveguide axis for a rotation angle of 17.5° similar to the one used for the bevel-cut case.

4. CONCLUSION

In Vlasov antennas with step cuts or bevel cuts, the maximum radiation is shifted by some angle with respect to the axis of the waveguide. In previous work, reflectors have been used to focus the beam in some direction. In this paper, a reflector attached to the waveguide structure was proposed, and its rotation angle was optimized to obtain maximum radiation along the axis of the waveguide. The advantages of this reflector structure are the smaller overall size of the antenna, its simpler design in terms of attaching the reflector to the waveguide, an increased gain and a decreased HPBW. The proposed reflector was tested with both bevel-cut and step-cut Vlasov antennas.

REFERENCES

1. Ling, G. S. and C. W. Yuan, "Design of a Vlasov antenna with reflector," *International Journal of Electronics*, Vol. 91, No. 4, 253–258, Apr. 2004.
2. Ruth, B. G., R. K. Dahlstrom, C. D. Schlesiger, and L. F. Libelo, "Design and low-power testing of a microwave Vlasov mode converter," *IEEE MTT-S International Microwave Symposium Digest*, Vol. 3, 1277–1280, 1989.
3. Dahlstrom, R. K., L. J. Hadwin, B. G. Ruth, and L. F. Libelo, "Reflector design for an X-band Vlasov antenna," *Antennas and Propagation Society International Symposium*, Vol. 2, 968–971, 1990.
4. Fazaelifar, M. and M. R. Fatorehchy, "Design, fabrication and test of parabolic cylinder reflector and horn for increasing the gain of Vlasov antenna," *Progress In Electromagnetics Research Letters*, Vol. 4, 191–203, 2008.
5. Zhang, X., Q. Wang, Y. Cheng, and S. Wen, "Design of a 220 GHz Vlasov antenna mode converter," *International Worksho on Microwave and Millimeter Wave Circuits and System Technology (MMWCST)*, 1–2, 2012.
6. El Misilmani, H. M., M. Al-Husseini, K. Y. Kabalan, and A. El-Hajj, "Improved Vlasov antenna with curved cuts for high power microwaves," *High Performance Computing and Simulation Conference (HPCS 2013)*, Helsinki, Finland, 2013.
7. ANSYS HFSS, [Online], Available: <http://www.ansys.com/Products/Simulation+Technology/Electromagnetics/High-Performance+Electronic+Design/ANSYS+HFSS>.
8. CST Computer Simulation Technology, [Online], Available: <http://www.cst.com/>.

The Imbedding Method in the Internal Electrodynamics Problem of Parabolic Reflector Antennas

S. V. Boyarkin and V. L. Kuznetsov

Moscow State Technical University of Civil Aviation, Russia

Abstract— A new method for compute of electrodynamics' characteristic of parabolic reflector, examine as section of irregular waveguide, is proposed. The value of matrix reflection coefficient is calculated by invariant imbedding method.

1. INTRODUCTION

Modern hybrid methods of computing parabolic reflector antennas include two steps. First step consists of evaluating surface currents on antenna reflector and a second one is computing far wavefields, generated by this currents. There is a natural question appears — is first step indeed necessary in reflector antennas theory.

The aim of this work is to examine another approach to analysis of direct focal parabolic reflector antenna which is based on representation of its reflector as a non-planar face of irregular waveguide with radius equals to aperture radius (Fig. 1).

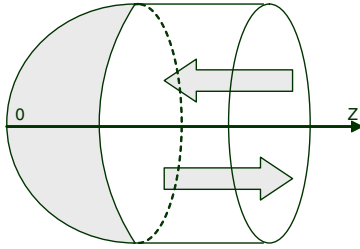


Figure 1. Mirror reflector antenna as a non-planer face of circular waveguide.

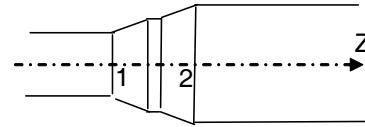


Figure 2. Irregularities in irregular waveguide.

The advantages of such representation are obvious because it is in good agreement with rigorous solution to open face waveguide radiation problem which was obtained by L. A. Veinshtein [1]. We note that corresponding convenient for practical calculate formulas are used in [2] for horn antennas.

Then we need to find the reflection coefficient of parabolic mirror. We will compute it with invariant imbedding method, which has proven itself in the description of irregular waveguides [3, 4]. Simple calculations which lead us to embedding equation for considering case will be given below.

2. IMBEDDING EQUATION FOR REFLECTION COEFFICIENT OF PARABOLIC REFLECTOR

In irregular waveguide theory has long been used the method of generalized scattering matrix [5, 6], which can be interpreted as some kind as “addition rule” for reflection and transition coefficients for sequence of two abutting each other irregularities (Fig. 2).

For the wave incident on the right side of this irregularities reflection coefficient R_{Σ}^{\pm} (upper sign indicates the propagation direction of reflected wave — “+” sign corresponds to propagation in the Z axial direction) is expressed through reflection — R_i^{\pm} , $i = 1, 2$ and transition — T_i^{\pm} , $i = 1, 2$ coefficients of each irregularity as follows [7]

$$R_{\Sigma}^{\pm} = R_2^{\pm} + T_2^{\pm} (R_1^{\pm} + R_1^{\pm} R_2^{\mp} R_1^{\pm} + \dots) T_2^{\mp}. \quad (1)$$

Members in brackets in (1) describing multiple scattering of wave from semitransparent irregularities. Now assume that the second irregularity is small in the sense that Δz — the area it occupies goes to zero. Given that in our case the profile of irregular waveguide changes smoothly along z for second irregularity we could write:

$$R_2^{\pm}(z, \Delta z) = \rho^{\pm}(z) \Delta z + o(\Delta z) \quad (2)$$

$$T_2^{\pm}(z, \Delta z) = I + \tau_2^{\pm}(z) \Delta z + o(\Delta z) \quad (3)$$

Here I — the identity operator. Substituting (2) and (3) in (1), keeping only members proportional to Δz and introducing the notation $R_{\Sigma}^{\pm} = R^{\pm}(z + \Delta z)$, $R_1^{\pm} = R^{\pm}(z)$, we obtain the following finite difference equation

$$R^+(z+\Delta z) = R^+(z) + [\rho^+(z) + R^+(z)\tau^-(z) + \tau^+(z)R^+(z) + R^+(z)\rho^-(z)R^+(z)] \cdot \Delta z + o(\Delta z) \quad (4)$$

In limit $\Delta z \rightarrow 0$ we obtain differential Riccati equation for reflection coefficient of irregular waveguide, known as imbedding equation

$$\frac{dR^+}{dz} = \rho^+ + R^+\tau^- + \tau^+R^+ + R^+\rho^-R^+ \quad (5)$$

Here we limit only this equation. Imbedding equation for generalized scattering matrix was obtained in [8].

3. REPRESENTATION OF FIELD AND CALCULATION OF IMBEDDING EQUATION COEFFICIENTS

So far we have not specified the form of reflection and transition operators. Introducing the field in the cross section z of the reflector in basis of modes in the waveguide of radius $a = a(z)$, we obtain that reflection coefficient could be represented in 4-index matrix form ${}_{\alpha\beta}R_{nm}$, where m — mode number and β — mode type (e or h) of initial field, n and α — corresponding parameters of the reflected field. The imbedding equation in such notation obtains following form

$$\frac{d}{dz} ({}_{\alpha\beta}R_{nm}^+) = {}_{\alpha\beta}\rho_{nm}^+ + {}_{\alpha\gamma}R_{np}^+ \cdot {}_{\gamma\beta}\tau_{pm}^- + {}_{\alpha\gamma}\tau_{np}^+ \cdot {}_{\gamma\beta}R_{pm}^+ + {}_{\alpha\gamma}R_{np}^+ \cdot {}_{\gamma\theta}\rho_{pk}^- \cdot {}_{\theta\beta}R_{km}^+ \quad (6)$$

Here all repeated indexes are summed over.

Now lets compute matrix coefficients of Equation (6). Corresponding to (2) and (3) this coefficients could be represented as following limits:

$${}_{\alpha\beta}\rho_{n,m}^{\pm}(z) = \lim_{\Delta z \rightarrow 0} \frac{{}_{\alpha\beta}R_{n,m}^{\pm}(z, \Delta z)}{\Delta z}, \quad {}_{\alpha\beta}\tau_{n,m}^{\pm}(z) = \lim_{\Delta z \rightarrow 0} \frac{{}_{\alpha\beta}T_{n,m}^{\pm}(z, \Delta z)}{\Delta z} \quad (7)$$

This means that reflection ${}_{\alpha\beta}R_{n,m}^{\pm}(z, \Delta z)$ and transition ${}_{\alpha\beta}T_{n,m}^{\pm}(z, \Delta z)$ coefficients of smaller section of irregular waveguide should be calculated only up to terms of order Δz . It allows to simplify the task and obtain analytical expressions for ρ^{\pm} and τ^{\pm} .

In conventional methods of computing irregular waveguide using various approximations of waveguide section profile: piecewise (in mode matching method), continuous in cross-sections method and method of integral equation.

In [4] on the example of flat horn transition shown that all this approaches lead to the same analytical expressions for coefficients of imbedding equation. So here we choose the most simple method — mode matching method for the stepwise approximation of irregular waveguide (Fig. 3).

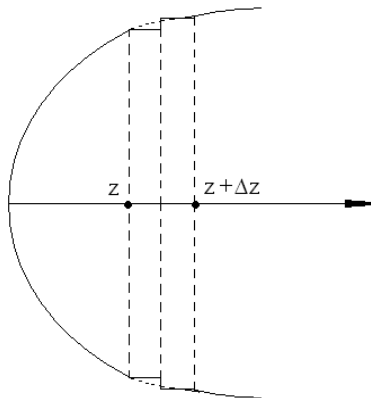


Figure 3. Stepwise approximation of mirror element.

Method of calculation of such irregularity is well-known [6]. As an input field is considered one of the natural modes of waveguide TE- or TH-type:

$$\begin{aligned} {}_e\Phi_n(\rho, \varphi; a) &= -\sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{\mu_n^2 - 1}} \frac{J_1(\mu_n \cdot \frac{\rho}{a})}{\rho \cdot J_1(\mu_n)} \cdot \cos(\varphi) \cdot \vec{e}_\rho + \sqrt{\frac{2}{\pi}} \frac{\mu_n}{\sqrt{\mu_n^2 - 1}} \frac{J_1'(\mu_n \cdot \frac{\rho}{a})}{a \cdot J_1(\mu_n)} \cdot \sin(\varphi) \cdot \vec{e}_\varphi, \\ {}_h\Phi_n(\rho, \varphi; a) &= -\sqrt{\frac{2}{\pi}} \frac{J_1'(\chi_n \cdot \frac{\rho}{a})}{a \cdot J_2(\chi_n)} \cdot \cos(\varphi) \cdot \vec{e}_\rho + \sqrt{\frac{2}{\pi}} \frac{J_1(\chi_n \cdot \frac{\rho}{a})}{\chi_n \cdot \rho \cdot J_2(\chi_n)} \cdot \sin(\varphi) \cdot \vec{e}_\varphi, \end{aligned}$$

Here a is the section radius of reflector in z point on the left border of elementary irregularity. Further $b = a(z + \Delta z)$ — the same, but on the right border of irregularity.

The interaction of wave with stepwise section of waveguide with length equals Δz includes the propagation on irregular sections and transformation of field in another modes in the plane of interconnection. In case of small Δz the coupling coefficients of natural modes in interconnected waveguides ${}_{\alpha\beta}C_{mn}(a, b) = \int_0^a \int_0^{2\pi} {}_\alpha\Phi_m(\rho, \varphi; a) \cdot {}_\beta\Phi_n(\rho, \varphi; b) \cdot \rho \cdot d\rho \cdot d\varphi$ obtains a particularly simple linear by Δz form:

$$\begin{aligned} {}_{ee}C_{mn}(a, b) &= \frac{2f}{a} \cdot ({}_{ee}\beta_{mn}) \cdot \Delta z; & {}_{he}C_{mn}(a, b) &= \frac{2f}{a} \cdot ({}_{he}\beta_{mn}) \cdot \Delta z; \\ {}_{hh}C_{mn}(a, b) &= \frac{2f}{a} \cdot ({}_{hh}\beta_{m \cdot n}) \cdot \Delta z \end{aligned} \quad (8)$$

Here μ and χ — the roots of $J_1'(\dots) = 0$ and $J_1(\dots) = 0$ and

$$\begin{aligned} {}_{ee}\beta_{nm} &= \begin{cases} \frac{2 \cdot \mu_n^2 \cdot (1 - \mu_m^2)}{a \cdot \sqrt{\mu_n^2 - 1} \cdot \sqrt{\mu_m^2 - 1} \cdot (\mu_n^2 - \mu_m^2)}, & n \neq m \\ 1/a, & n = m \end{cases}, & {}_{hh}\beta_{nm} &= \begin{cases} \frac{2 \cdot \chi_m^2}{a \cdot (\chi_n^2 - \chi_m^2)}, & n \neq m \\ -1/a, & n = m \end{cases}, \\ {}_{he}\beta_{nm} &= \frac{2}{a \cdot \sqrt{\mu_n^2 - 1}} \end{aligned}$$

writing the amplitudes of reflected and transited fields for initial mode with number m and type β (TE or TM) in the form $\sum_\alpha \sum_n {}_\alpha\Phi_n^\pm(\rho, \varphi; a) \cdot {}_{\alpha\beta}\rho_{n,m}^\pm(z) \cdot \Delta z$ and $\sum_\alpha \sum_n {}_\alpha\Phi_n^\pm(\rho, \varphi; a) \cdot {}_{\alpha\beta}\tau_{n,m}^\pm(z) \cdot \Delta z$, and using the smoothness conditions in plane of interconnection only with first-order terms by Δz , we obtain analytical form for coefficients of imbedding Equation (5):

Table 1.

TE waves	TM waves
${}_{ee}\tau_{lk}^+ = i\kappa_k^a \delta_{kl} + {}_{ee}\rho_{lk}^- + \frac{2f}{a} \cdot {}_{ee}\beta_{lk}$	${}_{hh}\tau_{lk}^+ = i\kappa_k^a \delta_{kl} + {}_{hh}\rho_{lk}^- + \frac{2f}{a} \cdot {}_{hh}\beta_{lk}$
${}_{he}\tau_{lk}^+ = \frac{f}{a} \cdot {}_{he}\beta_{lk} \cdot \left[1 + \frac{\kappa_k^a}{\kappa_l^a} \right]$	${}_{eh}\tau_{lk}^+ = 0$
${}_{ee}\rho_{lk}^- = \frac{f}{a} \cdot {}_{ee}\beta_{lk} \cdot \left[\frac{\kappa_k^a}{\kappa_l^a} - 1 \right] - \frac{f}{a} \cdot \frac{\kappa_k^a}{\kappa_k^a} \cdot \delta_{kl}$	${}_{hh}\rho_{lk}^- = \frac{f}{a} \cdot {}_{hh}\beta_{lk} \cdot \left[\frac{\kappa_k^a}{\kappa_l^a} - 1 \right] - \frac{f}{a} \cdot \frac{\kappa_k^a}{\kappa_k^a} \cdot \delta_{kl}$
${}_{he}\rho_{lk}^- = \frac{f}{a} \cdot {}_{he}\beta_{lk} \cdot \left[\frac{\kappa_k^a}{\kappa_l^a} - 1 \right]$	${}_{eh}\rho_{lk}^- = 0$
${}_{ee}\tau_{lk}^- = 2i\kappa_k^a \delta_{kl} - {}_{ee}\tau_{lk}^+$	${}_{hh}\tau_{lk}^- = 2i\kappa_k^a \delta_{kl} - {}_{hh}\tau_{lk}^+$
${}_{ee}\rho_{lk}^- = -{}_{ee}\rho_{lk}^+$	${}_{hh}\rho_{lk}^- = -{}_{hh}\rho_{lk}^+$
	$\kappa_k^a = \sqrt{k^2 - \left(\frac{\chi_k}{a}\right)^2}$

Here κ_k^a — longitudinal wave number for mode with number k in waveguide with radius a .

4. NUMERICAL RESULTS

And so, we have reduced boundary problem for field to Cauchy problem for reflection coefficient. As initial conditions we take the following: ${}_{\alpha\beta}R_{nm}(z_0) = -\delta_{\alpha\beta} \cdot \delta_{nm}$. Here δ_{ij} — Kronecker delta. This condition corresponds to that fraction of a paraboloid at its apex, we replaced by the plane. In

numerical computations we assumed the distance from the top of paraboloid to plane — $z'_0 = 0.02$. Hereinafter linear geometrical parameters of problem given in dimensionless variables: $a' = a \cdot k$ where k — the wave number in free space.

Dimensionless focus of mirror is equal to 33 and number of modes taken into account for each type is 23. This parameters corresponds that in antenna aperture we consider 20 propagating and 3 evanescent TM modes. In TE it would be 20 and 2 respectively.

The Fig. 4(a) shows the graphs of the reflection and transition coefficient modules from first mode in first for TE and TM. The Fig. 4(b) shows the same graphs zoomed to critical sections $z_1 = 0.02568$ for TE and $z_2 = 0.11123$ for TM.

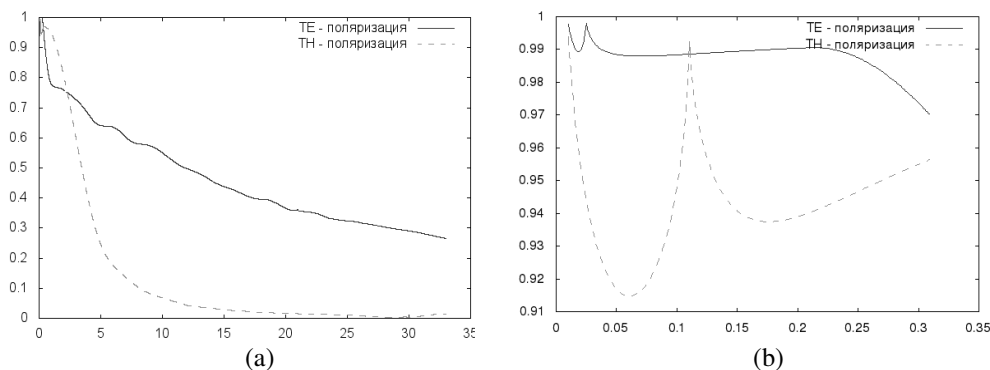


Figure 4.

The Figs. 5(a) and 5(b) show modules of $\alpha_\alpha R_{3,1}$ and $\alpha_\alpha R_{3,14}$ which define the pumping of field from first and fourteenth modes to third mode for TE and TM fields. We see that in the critical sections of the third mode this coefficients goes to zero. The same folds for other modes. This means that section of paraboloid with depth $z_k = \frac{\mu_k^2}{4 \cdot f}$ for TE field (or $z_k = \frac{\chi_k^2}{4 \cdot f}$ for TM fields) does not convert energy other bombarding modes of this type in mode with number k .

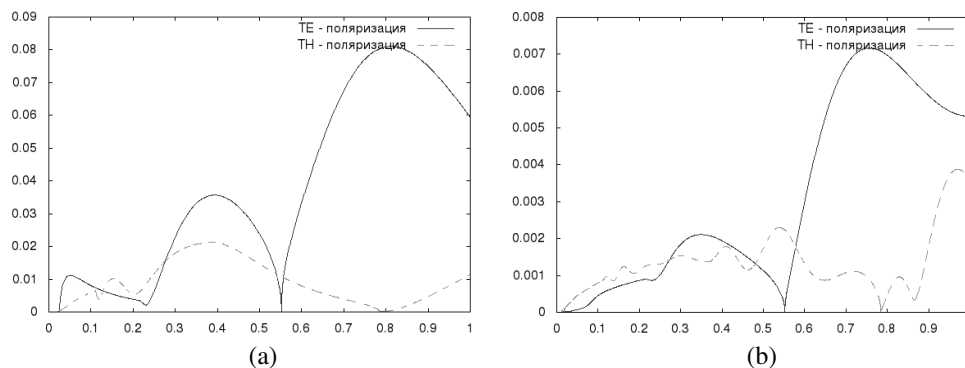


Figure 5.

5. CONCLUSION

This paper proposes a new approach to solution the inner electrodynamics problem of symmetrical mirror antennas based on representation parabolic mirror as a section of irregular waveguide. For calculating the reflection coefficients of mirror the imbedding method is proposed. This method allows to reduce such problem to Cauchy for reflection coefficients.

Proposed approach is in a good agreement with well-known solution of open waveguide radiation problem, and thus allows to hope for a rigorous solution of radiation parabolic reflector antenna problem.

REFERENCES

1. Veinshtein, L. A., *The Theory of Diffraction and the Factorization Method (Generalized Wiener-Hopf Technique)*, Golem Press, 1969.

2. Skobelev, S. P., I. L. Vilenko, Y. A. Suserov, et al., “Combined approach to analysis of axisymmetric horn antennas,” *Radiotechnics*, Vol. 4, 82, 2007.
3. Kuznetsov, V. L., S. P. Skobelev, and P. V. Filonov, “The modification of imbedding method for analysis of horn array antenna excited by TE-waves,” *Radiotechnics*, Vol. 4, 30–38, 2010.
4. Kuznetsov, V. L. and P. V. Filonov, “Imbedding equation and a small parameter in the problem of an irregular waveguide,” *Journal of Communications Technology and Electronics*, Vol. 56, 1086–1092, 2011.
5. Amitay, N., V. Galindo, and C. P. Wu, *Theory and Analysis of Phased Array Antennas*, Wiley, New York, 1972.
6. Mittra, R., *Computer Techniques for Electromagnetics*, Pergamon Press, New York, 1973.
7. Mittra, R. and S. Li, *Analytical Methods in Waveguide Theory*, M. Mir, 1974.
8. Barabanenkov, Y. N. and M. Y. Barabanenkov, “Energy invariants to composition rules for scattering and transfer matrices of propagating and evanescent waves in dielectric structures,” *PIERS Proceedings*, 10–12, Cambridge, USA, Mar. 26–29, 2006.

Feasibility of Using an Absorptive Cover for Ice, Snow or Water Removal from Radome Surface

Aleksey Solovey

L-3 Communications ESSCO, 90 Nemco Way, Ayer, MA 01432, USA

Abstract— The feasibility of using a thin layer of absorptive and/or conductive dielectric cover for the ice, snow or water removal from the radome outer surface was investigated. Constraints on the design parameters and usability of such de-icing system were derived. The new de-icing approach that employs the effect of total internal reflection of radome outer skin is introduced.

1. INTRODUCTION

To prevent or remove the ice, snow or water build-up from the radome surface (referred further as a de-icing) a chemical, mechanical, and thermal de-icing approaches were introduced at the early stage of radomes development [1]. This paper investigates the feasibility of the particular type of thermal de-icing approach when the radome outer surface is covered by a heating layer of absorptive and/or conductive dielectric. The similar method that uses the heating wire mesh embedded into the radome outer skin was used for the reference.

The infrared [2] and hot air blowing [3] de-icing approaches were not considered because those methods are suitable only for the metal space frame or thin solid laminate radomes whose wall has no thermo insulating core. For sandwich radomes with a low density core that works as a thermo insulator those approaches can be used only if the infrared radiators or hot air blowers are installed outside of the radome, which is impractical for most radome applications.

2. EVALUATION OF SURFACE POWER DENSITY REQUIRED FOR RADOME DE-ICING

Extensive experimentation with various environmental conditions [3] indicates that 0.16–0.48 W/cm² of heating power density applied to the radome surface can prevent up to 38 mm/hr ice build-up. The same power density evaluated from the ice melting latent heat is 0.34 W/cm², which results in the conservative estimate of the radome surface heating efficiency $\eta = 70\%$. Thus, for melting the ice at 0°C, i.e., just for the ice build-up prevention, the ice melting rate is:

$$r_{\text{ice melt}} = 12.7 \frac{\text{mW}}{\text{cm}^2} / \frac{\text{mm}}{\text{hr}} \quad (1)$$

To remove an existed ice build-up at a temperature below 0°C, the ice should be warmed before melting. Based on the same heating efficiency and the specific heat of ice, the ice warming rate can be evaluated:

$$r_{\text{ice warm}} = 0.078 \frac{\text{mW}}{\text{cm}^2\text{K}} / \frac{\text{mm}}{\text{hr}} \quad (2)$$

Similar considerations for the water warming and evaporation (if required) give:

$$r_{\text{water warm}} = 0.166 \frac{\text{mW}}{\text{cm}^2\text{K}} / \frac{\text{mm}}{\text{hr}} \quad (3)$$

$$r_{\text{water evapo}} = 90.1 \frac{\text{mW}}{\text{cm}^2} / \frac{\text{mm}}{\text{hr}} \quad (4)$$

The snow is a blend of an ice, air and water (if wet) thus, depending on snow density and wetness, its melting and warming rates should be less than those for an ice or water.

From (1)–(4) it can be concluded that (13–17) mW/cm² for the ice removal and an additional (90–107) mW/cm² for the water evaporation (if needed) should be applied to the radome surface per 1 mm/hr precipitation. So, from the overall power consumption standpoint (up to 1.3 kW/m² per 1 mm/hr precipitation) the thermal de-icing approach is feasible for the radome de-icing areas up to a few tens square meters.

3. RADOME DE-ICING USING HEATING WIRE MESH

As is well known from the car window de-icing experience, for the effective de-icing the distance between horizontal heating wires should be close to 3 cm. To propagate through such conductive mesh having just an optical blockage limit of the transmission loss, the antenna polarization should be either a) vertical (VP) or b) horizontal or circular (HP, CP) with the wavelength that is much lower than 3 cm.

This is illustrated in Figure 1 for the wire diameter 0.7 mm. The HP transmission loss, being unacceptably big at low frequencies, approaches the optical blockage limit when its wavelength becomes much less than the distance between wires. On the contrary, the VP transmission loss, being negligibly small at low frequencies, approaches that limit when its wavelength is a much smaller than the wire diameter.

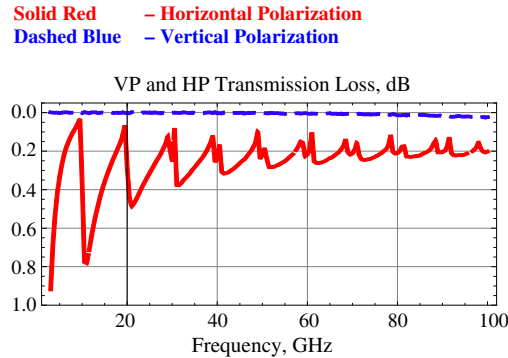


Figure 1: VP and HP transmission loss through heating wire mesh with 3 cm distance between horizontal wires at normal incidence.

If σ is the specific conductivity of the wire material (copper or aluminium), $h \approx 3$ cm is a distance between wires, and q_s is the surface power density that is needed for the proper radome de-icing (see Section 2), then the power balance between powers emitted by wires and needed for the radome de-icing defines the product of the wire diameter d and the effective value of vector of electric strength E :

$$dE = \sqrt{4q_s h / \pi \sigma} \quad (5)$$

Because value of d is determined by an acceptable level of optical blockage, the value of vector E can be found from (5). This value of E is shown in Table 1 against the wire diameter and the corresponding optical blockage limit of the transmission loss for 1 mm/hr precipitation. Since according to (1)–(5), the value of E is proportional to the square root of the precipitation level, values of E shown in Table 1 can be easily recomputed for any precipitation level of interest. The minimum value of E shown in each cell of Table 1 corresponds with an ice melting at 0°C , while the maximum value of E corresponds with melting of a cold ice at -50°C with subsequent water evaporation.

Table 1: Wire diameter and applied vector of electric strength needed for proper de-icing using heating wire mesh approach.

Optical Blockage Limit of Wire Mesh	Vector of Electric Strength and Diameter of Heating Wires	
	d, mm	E, V/cm
0.02	0.069	0.043–0.14
0.05	0.17	0.017–0.054
0.1	0.35	0.0086–0.027
0.2	0.70	0.0043–0.013

As is seen from Table 1 even for the biggest value of E , the maximum distance between electrodes that apply voltage to the heating wires and therefore, the wire length might be up to a few meters, depending on maximum voltage allowable and the precipitation level. These few meters long wire

mesh sections have no restriction in width and can be replicated many times to cover the entire radome surface. Thus, apart from the overall power consumption, the heating wire mesh de-icing approach has only one limitation concerning the HP and CP antennas: to have an acceptable level of transmission loss the antenna frequency must be significantly more than 10 GHz. Fine tuning of a heating mesh below the 10–20 GHz for the HP and CP radiations using minima of the transmission loss curve shown in Figure 1 is possible for rare cases of single, narrow band antennas that illuminate the radome surface with a very limited range of incident angles.

4. RADOME DE-ICING USING DC/50/60 Hz CONDUCTIVITY CURRENT THROUGH ABSORBER LAYER

To overcome the limitations of the heating wire mesh de-icing approach, the use of a thin layer of conductive absorber that covers the radome outer surface was investigated. To introduce just a small transmission loss at the antenna frequency band, such cover should be highly absorptive either well below or far above of this band. At very low, DC/50/60 Hz frequencies it might be achieved through the conductivity of the absorber, while at high frequencies the polarization of the absorber cover should be employed.

4.1. Surface Power Density Generated by DC/50/60 Hz Conductivity Current through Absorber

When the DC/50/60 Hz voltage is applied to a thin layer of the conductive absorber, the polarization portion of the transmission loss must be negligible. Otherwise, the absorber will not be transparent at the antenna frequency band, where the polarization losses are much greater than at the DC/50/60 Hz. Thus, the surface power density q_s generated within the absorber layer by the DC/50/60 Hz conductivity current is:

$$q_s = \sigma_a t_l E^2 \quad (6)$$

where, σ_a is the absorber conductivity, t_l is the thickness of the absorber layer, and E is an effective value of vector of electric strength.

4.2. Transmission Loss of Layer of Conductive Absorber

The analytical solution for the transmission coefficient S_{21} through the flat layer of a conductive dielectric is given in [1]. Although general expressions for the S_{21} are fairly complicated, for the absorber layer in vacuum (air) whose thickness is much less than the wavelength it can be reduced:

$$|S_{21 \text{ TE}}| = 1/(1 + 0.5\sigma_a t_l Z_c / \cos \theta) \quad |S_{21 \text{ TM}}| = 1/(1 + 0.5\sigma_a t_l Z_c \cos \theta) \quad (7)$$

where, $Z_c = 120\pi \Omega$ is the characteristic impedance of vacuum (air), and θ is the angle of incidence. From the practical standpoint the desirable thickness of the absorber layer is better to be a tiny fraction of mm and thus, the expressions (7) are valid at least up to 40 GHz, depending on particular dielectric constant and thickness of the absorber layer.

4.3. Constrains on DC/50/60 Hz Conductivity Current De-icing System Design and Absorber Material Parameters

According to (7), the absorber transmission loss defines the value of $\sigma_a t$. Then from (6), where q_s is defined in Section 2, the effective value of the vector of electric strength E can be determined. An example of values $\sigma_a t$ and E that correspond with the transmission loss of an absorber layer in vacuum (air) at normal incidence is shown in Table 2. In reality, such an absorber layer either covers the radome outer surface or might be an integral part of the latter (when the radome outer skin is intentionally made conductive, for instance). However, as it was investigated, this as well as a possible oblique incidence of the EM wave to the radome surface, makes no qualitative or even significant quantitative difference in values shown in Table 2.

The absorber transmission loss is shown in Table 2 against the corresponding absorber material parameter $\sigma_a t$. Since the value of E depends on the precipitation level, the range of its values is shown in Table 2 for 1 mm/hr precipitation. Based on (6) and (1)–(4), values of E can be easily recomputed for any precipitation level of interest. The smallest value of E in each cell of Table 2 corresponds with the ice melting at 0°C. The biggest value of E corresponds with melting of a cold ice at –50°C with subsequent water evaporation.

Table 2: Absorber material parameters and vector of electric strength for radome de-icing using DC/50/60 Hz conductivity current.

Acceptable Level of Absorber Transmission Loss, dB	Vector of Electric Strength and Absorber Material Parameters	
	$\sigma_{at1}, 10^{-6} 1/\Omega$	$E, V/cm$
0.02	12.2	32.6–99.1
0.05	30.6	20.6–62.6
0.1	61.4	14.5–44.2
0.2	123.6	10.3–31.2

4.4. Feasibility of Radome De-icing Using DC/50/60 Hz Conductivity Current through Thin Absorber Cover

As is seen from Table 2, for the 0.025–0.25 mm paint or spray like absorber (preferable from the cost standpoint), the range of absorber conductivity is (0.05–5) Si/m depending on thickness of the absorber layer and the acceptable level of the transmission loss. These values of absorber conductivity do not look unachievable [4].

However, as is also seen from Table 2, even the smallest magnitude of vector E is equal to 10 V/cm, though it corresponds with noticeable transmission loss through the absorber layer, minimal thermal de-icing requirements and just 1 mm/hr of precipitations. This means that the maximum distance between electrodes that apply voltage onto absorber layer should be no more than 1–10 cm depending on voltage allowable and precipitation level. Thus, for the proper de-icing the radome surface must be covered by 1–10 cm wide strips of absorber powered by two parallel conductive electrodes attached to each side of those absorber strips.

Because unlike the heating wire mesh case, the heat released by the absorber layer spreads uniformly across the radome surface, the orientation of the absorber strips does not affect its de-icing capability. Thus, by proper orientation of the absorber strips, the DC/50/60 Hz conductivity current de-icing approach can be used at any frequency for linearly polarized radiation (VP, HP or slant). That gives this de-icing approach an edge over the heating wire mesh one since the latter can be used at any frequency only for the VP radiation. With respect to CP radiation both of these approaches are the same (the frequency has to be significantly more than 10 GHz and the length of heating wires or the absorber strips can be up to few meters).

5. RADOME DE-ICING USING HIGH FREQUENCY POLARIZATION CURRENT THROUGH ABSORBER LAYER

For further enhancing the applicability of thermal de-icing approach, the use of a high frequency polarization current was investigated. In that case the absorber conductivity should be minimized to avoid an additional unwanted transmission loss at the radome frequency band of operation.

5.1. Polarization Current Induced by Quasi-stationary Electric Fields Between Two Parallel Electrodes

The surface power density released by the polarization current flowing through the non-conductive dielectric absorber layer can be calculated as follows:

$$q_s = \text{Re} \left[(\varepsilon - 1) \varepsilon_0 \frac{\partial E}{\partial t} E \right] = 2\pi f t_l \varepsilon_r \varepsilon_0 E^2 \tan \delta \quad (8)$$

where, f is the frequency of applied electric fields, ε_r is the real part of dielectric constant and $\tan \delta$ is the loss tangent of the absorber. The quasi-stationary electric fields between two parallel thin conductive strip or wire electrodes embedded into the absorber layer can be approximately evaluated as:

$$E = \frac{V}{2 \ln [h/r_w - 1]} \left(\frac{1}{r} + \frac{1}{h-r} \right) \quad (9)$$

where, h is the distance between electrodes, r_w is the radius of electrode wire (or the radius of the curvature of a strip edge), V is the voltage applied between electrodes and r is a distance from either one of two electrodes.

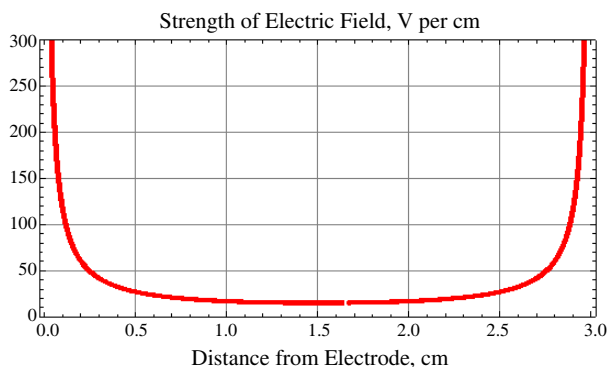


Figure 2: Case of two, 0.7 mm diameter electrodes embedded into typical radome outer skin at 3 cm distance apart. For proper de-icing frequency of applied voltage should be at least 2.2 GHz.

Based on (8) and (9) it was found that for any realistic values of radius of electrode wire (or strip edge) and absorber material parameters, the electric fields and therefore, the heat emission are highly concentrated in close proximity to the electrodes (see Figure 2). Thus, from a surface heating prospective the quasi-stationary polarization current de-icing approach has no advantages over the heating wire mesh one. Moreover, to produce the amount of power that is needed for proper de-icing, the frequency of applied voltage should be at least at hundreds of MHz range making the quasi-stationary polarization current de-icing approach uncompetitive with the heating wire mesh one from the cost standpoint.

5.2. Polarization Current Induced by Propagating Fields

Unlike the quasi-stationary fields, the propagating fields in a media with a small dissipation factor could spread heat emission more evenly. If the radiating antenna (printed strip or wire dipole, for example) would be embedded into the radome outer skin whose dielectric constant is bigger than the dielectric constant of the surrounding media (air, foam, ice or snow), its radiating power will be partly captured within the skin due to the effect of total internal reflection.

For a typical radome skin with the dielectric constant close to 4.0 the critical angle is about 30° and thus, approximately from one to two thirds of the radiated power will be captured within the skin and moved around by the surface wave. If the skin outer surface is covered by an ice or water, the fraction of power carried by the surface wave would be closer to one third; otherwise it would be rather close to two thirds. For a typical radome skin the dissipation factor is relatively low ($\tan \delta < 0.02$) and therefore, before its complete fading, the surface wave could travel many times around the radome surface (for spherical or cylindrical radomes) or create the standing wave within the skin (for flat radome panels with the reflective boundaries). That would result in almost uniform heat distribution across the radome outer surface. For big radomes several strip or wire type of dipole antennas (or linear arrays of antennas, depending on the radome size) should be embedded into radome outer skin at certain distance to insure the proper radome de-icing.

To quantify this idea let's estimate the area of the radome surface that can be adequately heated by just one radiating dipole. According the classic theory of dipole antennas [5], the maximum power radiated by a thin half-wave dipole is proportional to the square of applied voltage and for 100 V to be applied the power radiated by one dipole is close to 150 W. Being conservative let assume that just one third, i.e., 50 W would be carried by the surface wave and eventually absorbed by the radome outer skin. Then, the power balance between the absorbed power and the power needed for the radome de-icing gives the following relation between frequency f of the applied voltage and distance h between the dipoles (or linear arrays of dipoles):

$$h [\text{cm}] = (25 - 250)f [\text{GHz}] / r_p [\text{mm/hr}] \quad (10)$$

where, r_p is the precipitation rate and range of h corresponds with the range of surface power density needed for the radome de-icing under different de-icing scenarios (see Section 2).

As is seen from (10) for the de-icing approach that is based on the polarization current induced by the radiating fields at 1 GHz the distance between the conductive strips or wires that constitute the dipoles and its feeder lines is higher by at least one order of magnitude than for the heating wire mesh or DC/50/60 Hz conductivity current approaches. Because the heat released by the absorber layer spreads almost uniformly across the radome surface, the orientation of the radiating elements

does not affect its de-icing capability. Thus, the de-icing approach that uses the polarization current induced by the radiating fields has no frequency limitation for linearly polarized antennas and for CP antennas it can work above 1 GHz range depending on the particular de-icing scenario (Section 2). That gives this de-icing approach an edge over both the heating wire mesh and the DC/50/60 Hz conductivity current approaches.

6. CONCLUSIONS

1. The feasibility of a particular thermal ice, snow and/or water removal approach when the radome outer surface covered by thin absorptive and/or conductive dielectric heated by either low or high frequency current was investigated.

2. From the overall power consumption standpoint any thermal de-icing approaches required up to 1.3 kW/m^2 per 1 mm/hr precipitation and thus, are feasible for radomes with the de-icing area up to a few tens square meters.

3. The well-known heating wire mesh de-icing approach can be used for the VP radiation at any frequencies and for the HP and CP radiations only at frequencies that are much higher than 10 GHz. The distance between heating wires should be about 3 cm with the wire length up to few meters, depending on voltage allowable and precipitation level.

4. The de-icing approach that uses the polarization current induced by the quasi-stationary fields is uncompetitive with the heating wire mesh approach.

5. The DC/50/60 Hz conductivity current de-icing approach can be used for linearly polarized radiation (VP, HP or slant) at any frequency. This gives it an edge over the heating wire mesh de-icing approach. With respect to the CP radiation, the DC/50/60 Hz conductivity current and the heating wire mesh de-icing approaches are the same.

6. The new radome thermal de-icing approach is introduced. It uses the strip or wire type radiating antennas embedded into radome outer skin and employs the effect of total internal reflection. Although the most expensive, this approach can be used for linearly polarized radiation (VP, HP or slant) at any frequency and, depending on particular de-icing scenario, above 1 GHz for the CP radiation. That gives it an edge over both the heating wire mesh and the DC/50/60 Hz conductivity current de-icing approaches.

ACKNOWLEDGMENT

Cleared by DoD/OSR for public release under 13-S-1267 on March 14, 2013.

REFERENCES

1. Ridenour, L. N., Editor-in-Chief, *Radar Scanners and Radomes*, MIT Radiation Laboratory Series, McGraw-Hill, 1948.
2. Bowman, D. F., "Radome de-icing by infra-red focused via parabolic reflector through waveguide onto window," US Patent 3,173,141, March 9, 1965.
3. Goulding, M. K., "Study of the de-icing properties of the ASDE-3 rotodome," Final Report, DOT-TSC-FAA-81-24, September 1980.
4. Tour, J. M., "Conductive coatings based on graphene nanoribbon thin films for use in de-icing applications," Rice University Research Project, August 11, 2011.
5. Drabkin, A. L. and V. L. Zuzenko, "Antenna and feeder devices," *Soviet Radio*, Moscow, 1961 (in Russian).

Windowing of the Discrete Green's Function for Accurate FDTD Computations

T. P. Stefański

Department of Microwave and Antenna Engineering, Gdansk University of Technology
Narutowicza 11/12, Gdansk 80-233, Poland

Abstract— The paper presents systematic evaluation of the applicability of parametric and nonparametric window functions for truncation of the discrete Green's function (DGF). This function is directly derived from the FDTD update equations, thus the FDTD method and its integral discrete formulation can be perfectly coupled using DGF. Unfortunately, the DGF computations require processor time, hence DGF has to be truncated with appropriate window function.

The presented results extend previously published report which evaluates the accuracy of the DGF truncation for the most frequently applied window functions, i.e., Hann, Hamming, Gaussian and exponential windows. In this contribution, the accuracy of the DGF windowing is demonstrated for other window functions useful in the digital signal processing, i.e., Barlett, Blackman, Bohman, flat top and Kaiser windows. The study concludes that abrupt truncation of DGF waveforms results in an increase in the error of the electromagnetic field computations in comparison to the solution obtained for the DGF waveform truncated using a suitable window function. Truncation errors were compared for a wide range of window functions demonstrating the best performance for the Hann window.

1. INTRODUCTION

Recently, the discrete Green's function (DGF) [1] has become a popular tool facilitating the finite-difference time-domain (FDTD) method [2]. The DGF applications include electromagnetic simulations on disjoint domains [3, 4] and implementation of nonlocal absorbing boundary conditions [5, 6]. Consequently, free-space cells between scatterers and perfectly matched layers absorbing the electromagnetic wave at the grid boundaries can be avoided in FDTD simulations. The above-mentioned advantages of the DGF implementation in FDTD solvers have resulted in several applications of this function in antenna simulations [7, 8].

Unfortunately, numerical implementations of DGF are cumbersome. A three-dimensional (3-D) dyadic DGF formula in [1, 7] was derived with the use of Jacobi orthogonal polynomials by applying operators to the impulse response of the scalar wave equation. On the other hand, Jeng's formula [9] has a very compact form but employs powers of Courant numbers and factorials that cause numerical problems for large time indices. In fact, feasibility of the DGF generation for large time indices is limited by current computing technology, i.e., size of available memory and computing power. Therefore, only short DGF waveforms can be generated with the use of existing computing resources. To overcome this limitation, the DGF waveform has to be truncated with the use of a window function for accuracy of computations. However, although windowing has been reported as an efficient method of the DGF truncation and a remedy for stability issues [6], the accuracy and usability of this technique have not been deeply investigated yet.

The evaluation of accuracy of the DGF truncation for the most frequently applied window functions, i.e., Hann, Hamming, Gaussian and exponential windows, has already been presented [10]. In this contribution, the accuracy of the DGF windowing is demonstrated for other window functions useful in the digital signal processing, i.e., Barlett, Blackman, Bohman, flat top and Kaiser windows. In Section 2, the idea behind the DGF formulation of the FDTD method is presented. In Section 3, the evaluation of accuracy of the DGF windowing is presented in numerical benchmarks. Finally, the conclusions are given in Section 4.

2. THE 3-D DISCRETE GREEN'S FUNCTION FORMULATION OF THE FDTD METHOD

Linear and invariant systems such as FDTD equations can be represented by means of the convolution of the input sequence and the response of the system to the Kronecker delta-pulse source [1]:

$$\begin{bmatrix} \mathbf{E} \\ \eta \mathbf{H} \end{bmatrix}_{ijk}^n = \sum_{n'i'j'k'} \begin{bmatrix} \mathbf{G}_{ee} & \mathbf{G}_{eh} \\ \mathbf{G}_{he} & \mathbf{G}_{hh} \end{bmatrix}_{i-i'j-j'k-k'}^{n-n'} \begin{bmatrix} c\Delta t \eta \mathbf{J} \\ c\Delta t \mathbf{M} \end{bmatrix}_{i'j'k'}^{n'} \quad (1)$$

where c denotes the speed of light, η is the intrinsic impedance of free space, n is the time index, and Δt is the time-step size. Equation (1) is referred to as the convolution formulation of the FDTD method [7]. Computation of the electromagnetic field using the DGF formulation of the FDTD method requires to generate DGF waveforms. In the presented below research results, the method [11] was employed to generate the DGF waveforms. If the length of the DGF waveforms is equal to the number of iterations in the time-marching procedure of the FDTD method, the DGF formulation returns the same results as the direct FDTD method (assuming infinite numerical precision of computations). Since the number of FDTD iterations required for the convergence of a simulation is unknown in advance and the DGF generation and convolution computations require processor time, DGF has to be truncated with appropriate window function:

$$WG_{i,j,k}^n = W^n \cdot G_{i,j,k}^n, \quad (2)$$

$$W^n = \begin{cases} 0, & n < n_i \\ 1, & n_i \leq n < n_0 \\ w^n, & n_0 \leq n < n_t \\ 0, & n_t \leq n \end{cases}. \quad (3)$$

As the DGF values are nonzero only for $n \geq n_i$, the window length is defined as $n_s = n_t - n_i$. The value of n_i coefficient corresponds to the moment of the DGF wavefront arrival in the FDTD grid. The window function consists of a constant range ($W_n = 1$) and a tapering function range ($W_n = w^n$). Such a formula of windowing modifies only a tail of the DGF waveform. The constant range to total window length ratio (*CTR*) for the window functions was defined as a quotient of the number of samples with $W_n = 1$ and the total window size n_s :

$$CTR = \frac{n_0 - n_i}{n_s}. \quad (4)$$

The paper [10] reports evaluation of the accuracy of the DGF truncation for the most frequently applied window functions, i.e., Hann, Hamming, Gaussian and exponential windows, showing the best performance for the Hann window. In this contribution, the results of investigations are reported for the following window functions:

- Barlett:

$$w^n = 1 - \frac{n - n_0}{n_t - n_0}, \quad (5)$$

- Blackman:

$$w^n = 0.42 + 0.5 \cos\left(\pi \frac{n - n_0}{n_t - n_0}\right) + 0.08 \cos\left(2\pi \frac{n - n_0}{n_t - n_0}\right), \quad (6)$$

- Bohman:

$$w^n = \left(1 - \frac{n - n_0}{n_t - n_0}\right) \cos\left(\pi \frac{n - n_0}{n_t - n_0}\right) + \frac{1}{\pi} \sin\left(\pi \frac{n - n_0}{n_t - n_0}\right), \quad (7)$$

- flat top:

$$\begin{aligned} w^n = & 0.21557895 + 0.41663158 \cdot \cos\left(\pi \frac{n - n_0}{n_t - n_0}\right) + 0.277263158 \cdot \cos\left(2\pi \frac{n - n_0}{n_t - n_0}\right) \\ & + 0.083578947 \cdot \cos\left(3\pi \frac{n - n_0}{n_t - n_0}\right) + 0.006947368 \cdot \cos\left(4\pi \frac{n - n_0}{n_t - n_0}\right), \end{aligned} \quad (8)$$

- Kaiser:

$$\beta = \begin{cases} 0, & A < 21 \\ 0.5842(A-21)^{0.4} + 0.07886(A-21), & 21 \leq A \leq 50, \\ 0.1102(A-8.7), & A > 50 \end{cases}, \quad A \text{ — sidelobe attenuation (dB)} \quad (9)$$

$$w^n = \frac{I_0\left(\beta \sqrt{1 - \left(\frac{n-n_0}{n_t-n_0} - 1\right)^2}\right)}{I_0(\beta)}. \quad (10)$$

For the sake of comparison, results are also presented for the Hann window:

$$w^n = 0.5 \left[1 + \cos \left(\pi \frac{n - n_0}{n_t - n_0} \right) \right]. \quad (11)$$

It has already been demonstrated that CTR in the range 0.25–0.5 provides the best performance of the DGF windowing [10]. Therefore, $CTR = 0.5$ was taken for the reported below results. Although characteristics are not presented for the abrupt truncation of DGF ($CTR = 1$), it was verified that their accuracy is lower than for $CTR = 0.5$.

3. NUMERICAL RESULTS

Numerical tests were executed to evaluate the accuracy of the DGF windowing. The Courant numbers were taken as $s_x = s_y = s_z = 0.99/\sqrt{3}$ and the discretization step sizes $\Delta x = \Delta y = \Delta z = 1$ mm for the presented here results. The computations were executed in double floating-point precision. The relative error between vector results generated using the evaluated (\mathbf{E}_{ijk}^n) and reference (\mathbf{E}_{ref}^n) methods was calculated as:

$$Error = 20 \log_{10} \left[\frac{\max |\mathbf{E}_{ijk}^n - \mathbf{E}_{ref}^n|}{\max |\mathbf{E}_{ref}^n|} \right] \quad (\text{dB}). \quad (12)$$

Such a formula for error calculation is insensitive to distortions of a single field component which values are negligible in comparison to the maximal length of the field vector. The error (12) was evaluated in comparison with the convolution of a current source waveform with long DGF. Such a reference solution is equivalent to the FDTD simulation in the infinite Yee's grid, thus reference waveforms are not distorted by reflections from imperfect PMLs.

Figures 1–4 present the maximum relative error in the E -field vector computed using the convolution of the current source J_z with the WG_{ee} DGF as a function of the size of windows (5)–(11). Figs. 1–2 and Figs. 3–4 show results obtained respectively for the ramped sinusoid ($\lambda = 20\Delta x$) and differentiated Gaussian pulse (maximal frequency in the spectrum corresponding to the wavelength $\lambda = 20\Delta x$) excitations. The position of observation cell was varied in the (a) axial ($i, 0, 0$) and (b) diagonal (i, i, i) directions from the source. In the presented results, the best performance was obtained for the Hann window. As seen, the computations of the field require window length longer than minimal value providing an acceptable error level, e.g., the error below -50 dB requires 100

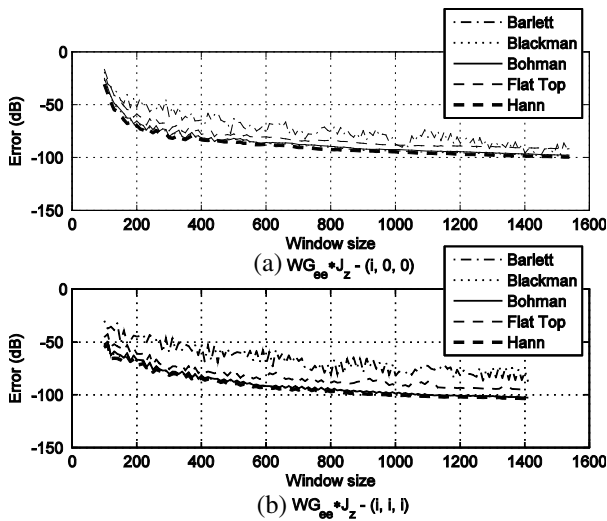


Figure 1: Maximum relative error in the E -field waveform computed using the convolution with the truncated WG_{ee} DGF as a function of the window size (ramped sinusoid excitation, $\lambda = 20\Delta x$, $i = 64$).

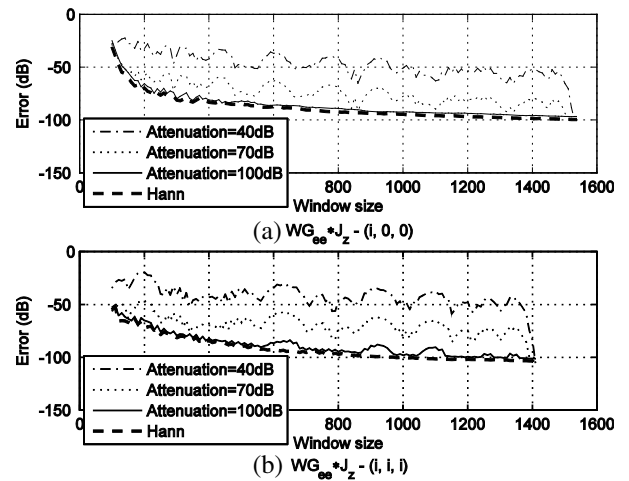


Figure 2: Maximum relative error in the E -field waveform computed using the convolution with the truncated WG_{ee} DGF as a function of the Kaiser window size and varying sidelobe attenuation (ramped sinusoid excitation, $\lambda = 20\Delta x$, $i = 64$).

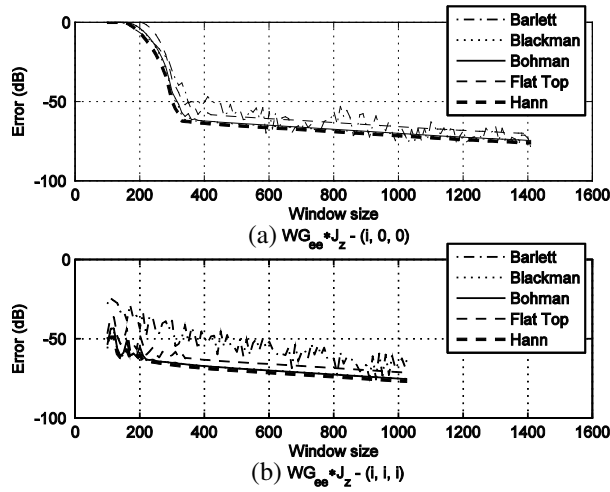


Figure 3: Maximum relative error in the E -field waveform computed using the convolution with the truncated WG_{ee} DGF as a function of the window size (differentiated Gaussian pulse excitation with the maximal frequency in the spectrum corresponding to the wavelength $\lambda = 20\Delta x$, $i = 192$).

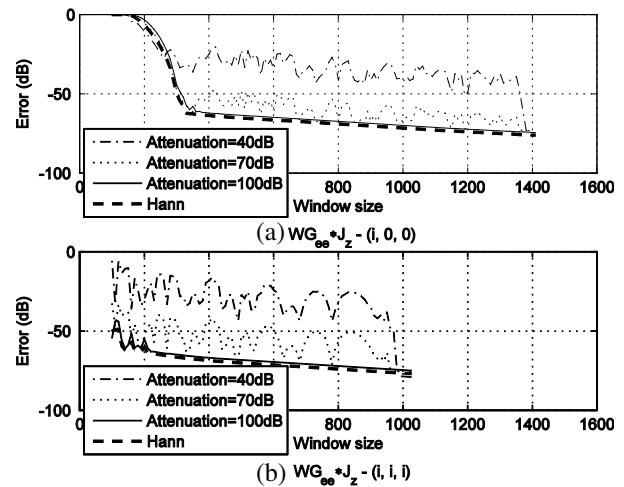


Figure 4: Maximum relative error in the E -field waveform computed using the convolution with the truncated WG_{ee} DGF as a function of the Kaiser window size and varying sidelobe attenuation (differentiated Gaussian pulse excitation with the maximal frequency in the spectrum corresponding to the wavelength $\lambda = 20\Delta x$, $i = 192$).

samples for the Hann window in Fig. 1. The error is slightly larger for observation points in the axial direction than for the diagonal direction in the grid. For pulse excitation (refer to Figs. 3–4), the error in the axial direction is increased for insufficient length of DGF. In this case, the remedy is to use longer window for the DGF truncation. The results show that sufficient accuracy of the DGF formulation of the FDTD method may require long DGF waveforms, whose generation takes significant computing runtimes.

4. CONCLUSIONS

The abrupt truncation of DGF waveforms results in an increase in the error of the electromagnetic field computations in comparison to the solution obtained for the DGF waveform truncated using a suitable window function. Truncation errors were compared for a wide range of parametric and nonparametric window functions demonstrating the best performance for the Hann window.

ACKNOWLEDGMENT

This work was realized within the HOMING PLUS Program of the Foundation for Polish Science, co-financed from the European Union Regional Development Fund.

REFERENCES

1. Vazquez, J. and C. G. Parini, "Discrete Green's function formulation of FDTD method for electromagnetic modelling," *Electron. Lett.*, Vol. 35, No. 7, 554–555, 1999.
2. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, 3rd Edition, Artech House, Boston, 2005.
3. De Hon, B. P. and J. M. Arnold, "Stable FDTD on disjoint domains — A discrete Green's function diakoptics approach," *Proc. The 2nd European Conf. on Antennas and Propag. (EuCAP)*, 1–6, 2007.
4. Stefanski, T. P., "Discrete Green's function approach to disjoint domain simulations in 3D FDTD method," *Electron. Lett.*, Vol. 49, No. 9, 597–599, 2013.
5. Holtzman, R. and R. Kastner, "The time-domain discrete Green's function method (GFM) characterizing the FDTD grid boundary," *IEEE Trans. Antennas Propag.*, Vol. 49, No. 7, 1079–1093, 2001.
6. Holtzman, R., R. Kastner, E. Heyman, and R. W. Ziolkowski, "Stability analysis of the Green's function method (GFM) used as an ABC for arbitrarily shaped boundaries," *IEEE Trans. Antennas Propag.*, Vol. 50, No. 7, 1017–1029, 2002.

7. Ma, W., M. R. Rayner, and C. G. Parini, "Discrete Green's function formulation of the FDTD method and its application in antenna modeling," *IEEE Trans. Antennas Propag.*, Vol. 53, No. 1, 339–346, 2005.
8. Mirhadi, S., M. Soleimani, and A. Abdolali, "An FFT-based approach in acceleration of discrete Green's function method for antenna analysis," *Progress In Electromagnetics Research M*, Vol. 29, 17–28, 2013.
9. Jeng, S.-K., "An analytical expression for 3-D dyadic FDTD-compatible Green's function in infinite free space via Z-transform and partial difference operators," *IEEE Trans. Antennas Propag.*, Vol. 59, No. 4, 1347–1355, 2011.
10. Stefanski, T. P., "Accuracy of the discrete Green's function formulation of the FDTD method," *IEEE Trans. Antennas Propag.*, Vol. 61, No. 2, 829–835, 2013.
11. Stefanski, T. P., "Fast implementation of the FDTD-compatible Green's function on multicore processor," *IEEE Antennas Wireless Propag. Lett.*, Vol. 11, 81–84, 2012.

Global Base Approach to H -polarized Wave Scattering by Conducting Circular Cylinders

K. Yashiro

Chiba University, Japan

Abstract— The scattering of a H -polarized plane wave by an arbitrary configuration of parallel perfect electric conducting (PEC) circular cylinders with any radii is analysed by method of moments (MoM). The scattered magnetic field is expressed as integral of the product of equivalent magnetic current and Green function, that is, the Hankel function of zero-th order. The unknown magnetic current density on each cylinder is expanded with global basis functions and the Galerkin's method is used to obtain a system of linear equations. Although the analysis objects are confined to circular cylinders, analytical expressions are obtained for the matrix elements by using Graf's addition theorem repeatedly. The comparison with previous work and numerical examples are presented.

1. INTRODUCTION

The structures of the arrangement of perfect electric conducting (PEC) circular cylinders often appear in microwave, millimeter wave and optical applications such as grating [1], photonic crystals [2], etc.. Hence it is desirable to analyse accurately the interaction of electromagnetic waves with large structure consisting of PEC circular cylinders.

The scattering of a plane wave by an arbitrary configuration of parallel circular cylinders was solved as boundary value problem to obtain a formal solution in terms of cylindrical wave functions [3]. The method of moments (MoM) [4] was also applied to the scattering by two parallel PEC circular cylinders [5]. The unknown current was expressed as a linear combination of pulse functions and the point matching technique was used.

In this paper, the MoM is implemented to simulate scattering from PEC circular cylinders. To reduce the number of the unknowns per cylinder with keeping the accuracy of solution, the unknown current density on each cylinder is expanded by periodic global base functions.

2. FORMULATION

The geometry of the problem is illustrated in Fig. 1. Suppose that the incident plane wave has only a z -component of the magnetic field. A time dependence $e^{j\omega t}$ is assumed and suppressed throughout. Then the incident plane wave is expressed as

$$H_z^{inc}(\vec{\rho}) = H_0 e^{-jk_0(x \cos \phi_0 + y \sin \phi_0)} \quad (1)$$

where k_0 is the wave number. Assumed that the equivalent magnetic current on the ν -th cylinder is $M_z^\nu(\vec{\rho})$. Then the scattered field is described by

$$H_z^{sc}(\vec{\rho}) = -\frac{k_0 \eta_0}{4} \sum_{\nu=1}^N \int_{C_\nu} M_z^\nu(\vec{\rho}') H_0^{(2)}(k_0 |\vec{\rho} - \vec{\rho}'|) d\rho' \quad (2)$$

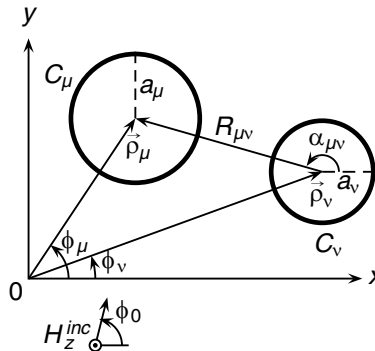


Figure 1: Geometry of problem.

where $H_0^{(2)}(x)$ is the Hankel function of the second kind of 0-th order and $\eta_0 (= \sqrt{\varepsilon_0/\mu_0})$ is the intrinsic admittance of free space. The total tangential component of electric fields vanishes on the surface of PEC cylinders. Hence,

$$E_{\theta_\nu}^{sc}(\vec{\rho}) + E_{\theta_\nu}^{inc}(\vec{\rho}) = \frac{-1}{j\omega\varepsilon_0} \left(\frac{\partial H_z^{sc}}{\partial r_\nu} + \frac{\partial H_z^{inc}}{\partial r_\nu} \right) = 0, \quad \vec{\rho} \in C_\nu \quad (\nu = 1, 2, \dots, N) \quad (3)$$

Here θ_ν is the azimuthal angle of the local reference system of which the origin is on the center axis of the ν -th circular cylinder. Equation (3) gives a system of boundary integral equations.

3. REDUCTION TO MATRIX EQUATION

Here we introduce the local reference system of which the origin is on the center axis of the ν -th circular cylinder, that is, $x = x_\nu + r_\nu \cos \theta_\nu$ and $y = y_\nu + r_\nu \sin \theta_\nu$. Since the magnetic current density should be periodic along the periphery of the cross section of each circular cylinder, each current is expanded as

$$M_z^\nu(\vec{\rho}') = M_z^\nu(\vec{\rho}_\nu + \vec{a}_\nu') = H_0 \sum_{n=-N_\nu}^{N_\nu} \chi_n^\nu e^{-jn\theta_\nu'} \quad (4)$$

Substituting (4) into (2), and using Graf's addition theorem [6]

$$H_n^{(2)}(k_0|\vec{\rho} - \vec{\rho}'|) \frac{\cos n\psi}{\sin m\varphi} = \sum_{m=-\infty}^{\infty} H_{n+m}^{(2)}(k_0R_\nu) J_m(k_0a_\nu) \frac{\cos m\varphi}{\sin m\varphi} \quad (5)$$

we integrate the resulting equation to obtain

$$H_z^{sc}(\vec{\rho}) = -\frac{\pi\eta_0}{2} H_0 \sum_{\nu=1}^N \sum_{n=-N_\nu}^{N_\nu} \chi_n^\nu k_0 a_\nu J_n(k_0 a_\nu) H_n^{(2)}(k_0 R_\nu) e^{-jn\theta_\nu} \quad (6)$$

where $R_\nu = |\vec{\rho} - \vec{\rho}_\nu|$, θ_ν is an angle between $\vec{\rho} - \vec{\rho}_\nu$ and x -axis. Three points $\vec{\rho}$, $\vec{\rho}'$, $\vec{\rho}_\nu$ form a triangle with sides, $|\vec{\rho} - \vec{\rho}'|$, $|\vec{\rho} - \vec{\rho}_\nu|$, $|\vec{\rho}' - \vec{\rho}_\nu|$, and ψ and φ are vertex angles opposite to $|\vec{\rho}' - \vec{\rho}_\nu|$ and $|\vec{\rho} - \vec{\rho}'|$, respectively.

The Galerkin's method is used to obtain a system of linear equations. By substituting (6) into (3), and using Graf's addition theorem again, the first term of (3) becomes

$$E_{\theta_\mu}^{sc}(\vec{\rho}_\mu + \vec{a}_\mu) = -j\frac{\pi}{2} H_0 \left\{ \sum_{n=-N_\mu}^{N_\mu} \chi_n^\mu k_0 a_\mu J_n(k_0 a_\mu) H_n^{(2)'}(k_0 a_\mu) e^{-jn\theta_\mu} \right. \\ \left. + \sum_{\substack{\nu=1 \\ \nu \neq \mu}}^N \sum_{n=-N_\nu}^{N_\nu} \sum_{k=-\infty}^{\infty} \chi_n^\nu k_0 a_\nu e^{-jn\alpha_{\mu\nu}} J_n(k_0 a_\nu) J_k'(k_0 a_\mu) H_{n+k}^{(2)}(k_0 R_{\mu\nu}) \Phi_k(\varphi) \right\} \quad (7)$$

$$\Phi_k(\varphi) = \begin{cases} e^{-jk\varphi} & \psi = \theta_\nu - \alpha_{\mu\nu} \geq 0 \\ e^{jk\varphi} & \psi = \alpha_{\mu\nu} - \theta_\nu \geq 0 \end{cases}$$

By multiplying (7) by $e^{jm\theta_\mu}$ and integrating the resulting equation along C_μ , it becomes

$$\int_{-\pi}^{\pi} E_{\theta_\mu}^{sc}(\vec{\rho}_\mu + \vec{a}_\mu) e^{jm\theta_\mu} a_\mu d\theta_\mu = -j\frac{\pi^2}{k_0} H_0 \left\{ \chi_m^\mu k_0 a_\mu J_n(k_0 a_\mu) k_0 a_\mu H_n^{(2)'}(k_0 a_\mu) \right. \\ \left. + \sum_{\substack{\nu=1 \\ \nu \neq \mu}}^N \sum_{n=-N_\nu}^{N_\nu} \chi_n^\nu k_0 a_\nu J_n(k_0 a_\nu) k_0 a_\mu J_k'(k_0 a_\mu) H_{n-m}^{(2)}(k_0 R_{\mu\nu}) e^{-j(n-m)\alpha_{\mu\nu}} \right\} \quad (8)$$

The incident wave is expressed in the local reference system (r_μ, θ_μ) as follows

$$H_z^{inc}(\vec{\rho}) = H_z^{inc}(\vec{\rho}_\mu + \vec{r}_\mu) = H_0 e^{-jk_0 \rho_\mu \cos(\phi_\mu - \phi_0)} \sum_{n=-\infty}^{\infty} (-j)^n J_n(k_0 r_\mu) e^{-jn(\theta_\mu - \phi_0)} \quad (9)$$

By differentiating (9) with respect to r_μ , the incident electric field on the surface of μ -th cylinder is obtained by

$$E_\theta^{inc}(\vec{\rho}_\mu + \vec{a}_\mu) = j\zeta_0 H_0 e^{-jk_0 \rho_\mu \cos(\phi_\mu - \phi_0)} \sum_{n=-\infty}^{\infty} (-j)^n J'_n(k_0 a_\mu) e^{-jn(\theta - \phi_0)} \quad (10)$$

where $\zeta_0 = 1/\eta_0$. By multiplying (10) by $e^{jm\theta_\mu}$ and integrating the resulting equation along C_μ , it becomes

$$\int_{-\pi}^{\pi} E_\theta^{inc} e^{jm\theta_\mu} a_\mu d\theta_\mu = j \frac{2\pi}{k_0} (-j)^m \zeta_0 H_0 k_0 a_\mu J'_m(k_0 a_\mu) e^{j\{m\phi_0 - k_0 \rho_\mu \cos(\phi_\mu - \phi_0)\}} \quad (11)$$

All the resulting equations obtained above are cast into a system of linear equations.

$$\begin{bmatrix} A_{11} & A_{12} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NN} \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_N \end{bmatrix} \quad (12)$$

where $A_{\nu\nu}$ and $A_{\mu\nu}$ are $(2N_\nu + 1) \times (2N_\nu + 1)$ diagonal matrix and $(2N_\mu + 1) \times (2N_\nu + 1)$ matrix, respectively, and X_ν and B_ν are $2N_\nu + 1$ column vectors, and their elements are given as follows:

$$a_{mn}^{\nu\nu} = -j \frac{\pi^2}{k_0} H_0 (k_0 a_\nu)^2 J_m(k_0 a_\nu) H_m^{(2)'}(k_0 a_\nu) \delta_{mn} \quad (13)$$

$$a_{mn}^{\mu\nu} = -j \frac{\pi^2}{k_0} H_0 k_0 a_\mu k_0 a_\nu J_n(k_0 a_\nu) J'_m(k_0 a_\mu) H_{n-m}^{(2)}(k_0 R_{\mu\nu}) e^{-j(n-m)\alpha_{\mu\nu}} \quad (14)$$

$$b_m^\mu = -j \frac{2\pi}{k_0} (-j)^m \zeta_0 H_0 k_0 a_\mu J'_m(k_0 a_\mu) e^{j\{m\phi_0 - k_0 \rho_\mu \cos(\phi_\mu - \phi_0)\}} \quad (15)$$

4. SCATTERED FAR FIELD

The scattered field from all cylinders is obtained by (6) after solving (12) numerically. Upon using the far-field approximation, $|\vec{\rho} - \vec{\rho}_\nu| \sim \rho - \vec{\rho}_\nu \cdot \vec{\rho}/\rho$ and $\theta_\nu \sim \phi$, the far-scattered field is expressed as

$$H_z^{sc}(\vec{\rho}) \sim H_0 \sqrt{\frac{2}{\pi k_0 \rho}} e^{-j(k_0 \rho - \frac{\pi}{4})} f(\phi) \quad (16)$$

where $f(\phi)$ is defined by

$$f(\phi) = -\frac{\pi\eta_0}{2} \sum_{\nu=1}^N \sum_{n=-N_\nu}^{N_\nu} j^n \chi_n^\nu k_0 a_\nu J_n(k_0 a_\nu) e^{jk_0 \vec{\rho}_\nu \cos(\phi - \phi_\nu)} e^{-jn\phi} \quad (17)$$

Here we call $f(\phi)$ the scattered far field amplitude.

The scattering width $\sigma(\phi)$ of the multiple cylinders is given by $4|f(\phi)|^2/k_0$. Thus, the total scattering width σ_{tot} is given by

$$\sigma_{tot} = \int_0^{2\pi} \sigma(\phi) d\phi = \frac{4}{k_0} \int_0^{2\pi} |f(\phi)|^2 d\phi \quad (18)$$

Using the principle of conservation of energy, the total scattering width σ_{tot} is related to the forward scattering amplitude $f(\phi_0)$, that is,

$$\sigma_{tot} = -\frac{8\pi}{k_0} \Re[f(\phi_0)] \quad (19)$$

The above relation is called the optical theorem.

5. NUMERICAL RESULTS

In order to verify the validity of the present method, some examples are considered. Firstly, we consider the scattering from a single PEC cylinder. In this case, a system of linear equations is solved analytically as the coefficient matrix is diagonalized. The solution for $H_z^{sc}(\vec{\rho})$ is identical with that obtained by the method of separation of variables.

Secondly, the scattering cross section of five PEC circular cylinders due to a H -polarized plane wave incident with $\phi_0 = 0$ was analysed in Ref. [7]. Each cylinder radius is 0.1λ and their centers are separated by 0.5λ . The numerical result obtained based on the above analysis is shown in Fig. 2. The result is compared with that in [7], where the interaction current was extracted by subtracting the current on an isolated cylinder from the total current and it was solved by the ordinary MoM. The solid curve is obtained by present method and the dotted one is duplicated from [7]. It is clear from the figure that the proposed solution is in complete agreement with the solution in [7].

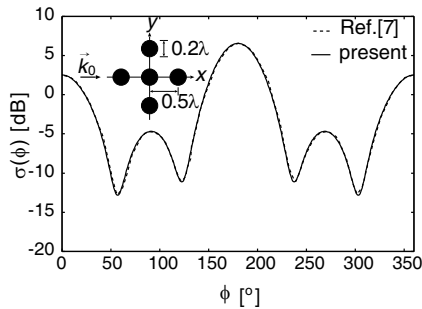


Figure 2: The scattering cross section of the five PEC circular cylinders.

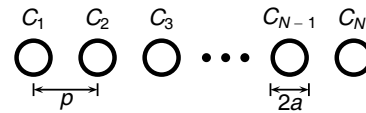


Figure 3: The finite linear array of the PEC circular cylinders.

Finally, let's consider the scattering by a finite linear array of PEC circular cylinders shown in Fig. 3. Suppose that each cylinder has a radius of a and the spacing between centers of adjacent cylinders is p . We confine ourselves to the case of grazing incidence. The total scattering width σ_{tot} of a linear array of ten PEC cylinders is plotted in Fig. 4 as a function of normalized frequency p/λ with the normalized radius as a parameter. We verified that the curves drawn by using the real part of forward scattered far field amplitude $f(\phi_0)$ coincided with those of σ_{tot} and the optical theorem (19) holds. It is seen from Fig. 4 that there are peaks or dips at $p \approx n\lambda/2$.

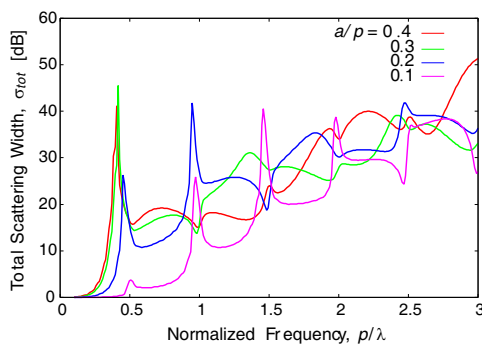


Figure 4: Total scattering width of a linear array of ten cylinders.

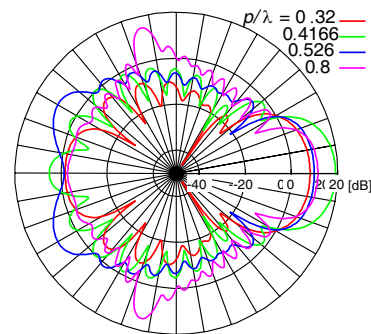


Figure 5: Far scattered field pattern for the case of $a/p = 0.3$.

The far scattered field pattern $|f(\phi)|$ is shown in Fig. 5. Each cylinder radius is $0.3p$. The forward scattering amplitude is largest for $p/\lambda = 0.32$ and 0.4166 . Especially, this tendency is remarkable at the first peak. The scattering amplitude has a large peak at $\phi = \pm 150^\circ$ in case of $p/\lambda = 0.526$, and it has a maximum at $\phi = \pm 119^\circ$ in case of $p/\lambda = 0.8$.

We computed the distribution of the total magnetic field around the array. Figs. 6(a) and 6(b) show distributions of $|H_z(\vec{\rho})|$ in case of $p/\lambda = 0.4166$ and $p/\lambda = 0.526$, respectively. The magnetic field concentrates in the gaps between cylinders and the wave propagates along the array for the case of $p/\lambda = 0.4166$, while the wave can not propagate along the array and is reradiated in the direction of $\phi \approx \pm 150$ in accordance with far scattered field pattern in Fig. 5.

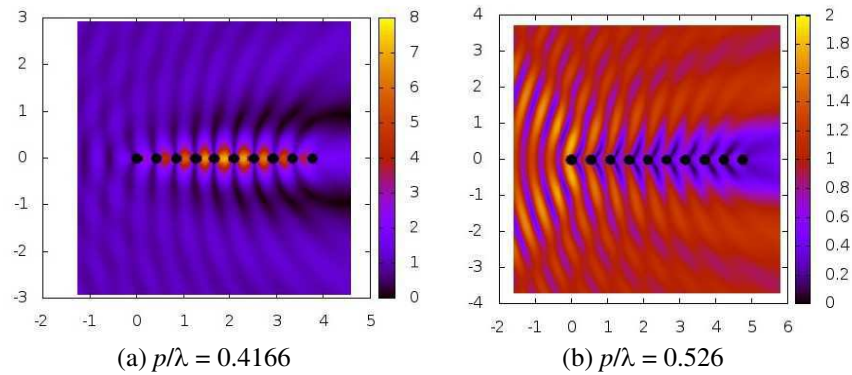


Figure 6: Field distribution for the case of $a/p = 0.3$.

The far scattered field pattern and the distribution of total magnetic field are shown respectively in Figs. 7 and 8 for the case where the total scattering width has a dip at $p/\lambda = 0.983$. It is seen from Figs. 7 and 8 that the wave is reflected strongly and the magnetic field in the gaps is weak.

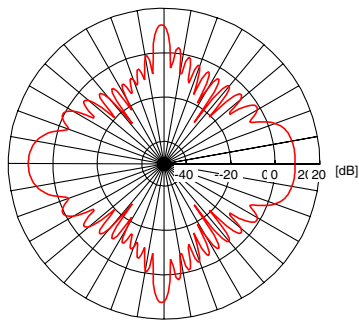


Figure 7: Far scattered field pattern for the case of $a/p = 0.3$ and $p/\lambda = 0.983$.

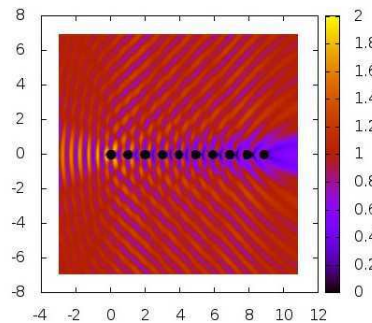


Figure 8: Field distribution for the case of $a/p = 0.3$ and $p/\lambda = 0.983$.

6. CONCLUSION

The scattering of H -polarized plane wave by an arbitrary configuration of parallel PEC circular cylinders is analysed by the method of moments. The equivalent magnetic current density on each cylinder is expanded with global basis functions and the Galerkin's method is used to obtain a system of linear equations. The elements of the coefficient matrix are obtained analytically.

REFERENCES

1. Twersky, N., "On scattering of waves by the infinite grating of circular cylinders," *IRE Trans. Antennas Propag.*, Vol. 10, No. 6, 737–765, Nov. 1962.
2. Nicorovici, N. A., R. C. McPhedran, and L. C. Botten, "Photonic band gaps for arrays of perfectly conducting cylinders," *Phys. Rev. E*, Vol. 52, No. 1, 1135–1145, Jul. 1995.
3. Twersky, V., "Multiple scattering of radiation by an arbitrary configuration of parallel cylinders," *J. Acoust. Soc. Ame.*, Vol. 24, No. 1, 42–46, Jan. 1952.
4. Harrington, R. F., *Field Computation by Moment Methods*, Macmillan, New York, 1968.
5. Elsherbeni, A. Z. and M. Hamid, "Scattering by parallel conducting cylinders," *IEEE Trans. Antennas Propagat.*, Vol. 35, No. 3, 355–358, 1987.
6. Abramowitz, M. and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 363, Dover, New York, 1972.
7. Elsherbeni, A. Z., "A comparative study of two-dimensional multiple scattering techniques," *Radio Science*, Vol. 29, Nos. 7–8, 1023–1033, Jul.–Aug. 1994.

Optical Interaction of Two Closely Spaced Nanoholes in Au Film

M. Janipour¹, T. Pakizeh², and F. Hodjat-Kashani¹

¹Department of Electrical Engineering, Iran University of Science Technology, Tehran 1684613114, Iran

²Faculty of Electrical and Computer Engineering, K. N. Toosi University of Technology, Tehran 16314, Iran

Abstract— Two closely spaced nanoholes analyzed as two tiny interacting magnetic dipoles depending on the incident beam polarization [1–3]. It is demonstrated that the optical coupling mechanism between two nanoholes inherently has a magnetic property, unlike the electric dipoles which are mainly coupled through electric-fields [4]. It is obvious that, in comparison with a single hole, the transmission experiences two peaks for Fig. 1(b) where the long and short wavelength peaks are proportional to anti-phase and in-phase magnetic dipoles, respectively. The long wavelength peak considerably blue-shifts by increasing the distance, whereas the short wavelength peak partially red-shifts. It is obvious from Fig. 1(a) and Fig. 1(b) that the electric field is confined in the air region of nanoholes and thus unlike electric dipoles, coupling through electric fields for nanoholes is very weak, while coupling through magnetic fields is strong. The peak of Fig. 1(d) is due to the in-phase interaction of dipoles and experiences a little blue shift relative to single hole, which is caused by electric coupling of dipoles.

1. INTRODUCTION

Since exploration of enhanced and extraordinary optical transmission through subwavelength nanoholes, much attention has been attracted to phenomena and applications related to subwavelength apertures drilled in a metal film [1]. The most important optical feature of metallic nanostructures is the bypassing of the diffraction limit in conventional optics. Using this unique trait and the electrical properties of metallic nanostructures, plasmonics could take a major step toward achieving nanoscale photonic and electronic devices. The extraordinary optical transmission through nanoholes is due to the arrangement of surface plasmon polaritons (SPPs) and refractive coupling effects with localized resonances [2]. The SPP coupling occurs when the separation distance between holes is comparable to the SPP wavelength, or in other words each nanohole locates in the far-field optical region of its neighboring hole. In the present study, the coupling mechanisms between two closely spaced (i.e., in the near-field region of each other) subwavelength holes are investigated.

2. THE THEORY AND MODEL

We deal with the optical coupling effects of two nanoholes which are placed closely. Two rectangular shaped nanoholes with equal sizes are considered. It is appropriate to investigate the electric and magnetic field powers in the near-field region to identify the electric or magnetic coupling effect of two nanoholes. The normalized electric and magnetic power amplitudes can be obtained using [5]:

$$A_e = \frac{-1}{\mathbf{P}} \int_v \mathbf{E} \cdot \mathbf{J} dv, \quad (1)$$

$$A_m = \frac{1}{\mathbf{P}} \int_v \mathbf{H} \cdot \mathbf{M} dv, \quad (2)$$

where, \mathbf{J} and \mathbf{M} are the volume distribution of the electric and magnetic current densities induced in the metal region around the nanohole, respectively. In Eq. (1) and Eq. (2), \mathbf{P}_0 is a normalization factor to the power flow, which is defined as [5]:

$$\mathbf{P}_0 = 2 \int_{S_0} (\mathbf{E} \times \mathbf{H}) \cdot \hat{\mathbf{n}} ds, \quad (3)$$

Figures 1(a) and 1(b) depict the field profile of $|H_y/H_0|^2$ and $|E_x/E_0|^2$ components for a rectangular nanohole in a thin gold film and the normalized power amplitudes in the metal medium, respectively. The integration region of S_0 in Eq. (3) is a rectangular area from $x = -200$ to 200 nm, and $y = -200$ to 200 nm. The simulation is based on a three dimensional finite-difference time-domain (FDTD)

method [6], with $\Delta x = \Delta y = \Delta z = 3$ nm, and considering that each nanohole dimension is $a = 100$ nm, $b = 200$ nm, and thickness $d = 100$ nm. It is obvious from Fig. 1(a), although the incident light is polarized along x -axis, the rectangular nanohole's magnetic field radiates along y -axis, and the electric field is confined in the dielectric region (inside the nanohole). This is in contrast to the nanoparticles, which radiate along the \mathbf{P} vector (polarization of the incident light). Fig. 2(a) depicts the normalized magnetic power amplitude, obtained from FDTD simulation and quasi-static approximation. The quasi-static approximation has a good agreement with the calculations. The normalized magnetic power amplitude, as a criterion for comparison between electric and magnetic powers, is much greater than the normalized electric power (i.e., the normalized electric power is in the order of 10^{-3}). Fig. 2(b) shows the normalized magnetic power amplitudes of the two rectangular nanoholes which are positioned in the near-field region of each other and are exposed by an x -polarized incident field.

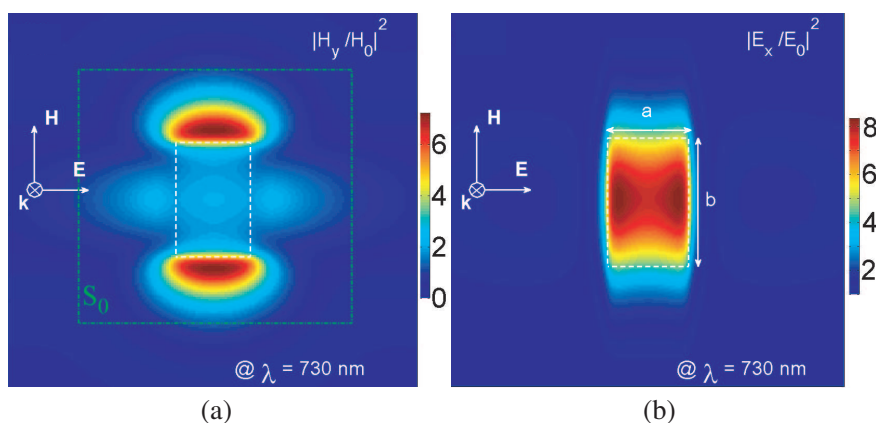


Figure 1: Intensities of (a) the magnetic field component ($|H_y/H_0|^2$), and (b) the electric field component ($|E_x/E_0|^2$) of the rectangular nanohole milled in an Au film with $a = 100$ nm, $b = 200$ nm and $d = 100$ nm at $\lambda = 730$ nm (LSPR wavelength).

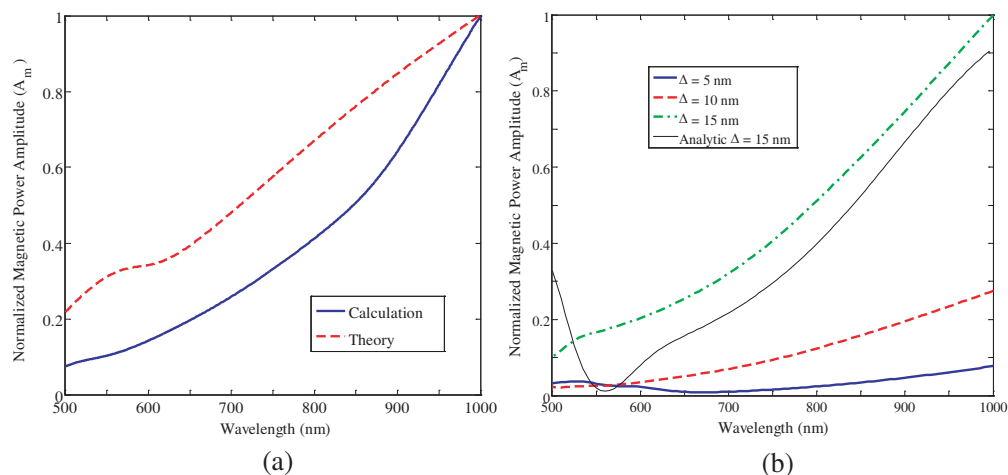


Figure 2: The normalized magnetic power amplitude for (a) single rectangular nanohole, and (b) two closely spaced nanoholes in gold film.

The separation distance between nanoholes experiences different values of $\Delta = 5, 10,$ and 15 nm, respectively. In order to confirm the results, the analytic calculation for the amplitude of the magnetic power, based on the quasi-static dipole approximation, is presented. As mentioned in Fig. 4(a) the electric power is negligible in the gold medium thus, the normalized magnetic power is investigated in the Δ region. According to Fig. 2(b) the limited increasing of the separation distance results in the increasing of the magnetic power amplitude and it seems that two dipoles with a magnetic nature are interacting.

In the case of two nanoholes, if both of holes lie along the major and/or minor axis the struc-

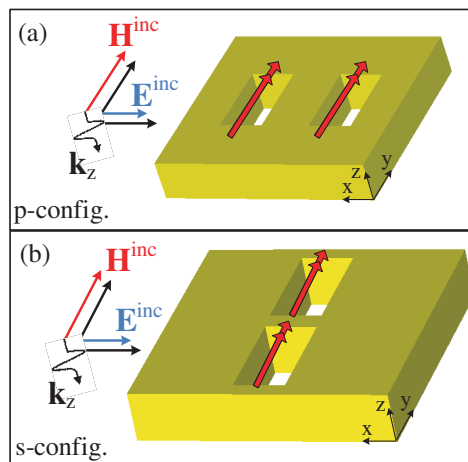


Figure 3: The schematic representation for arrangements of two nanoholes. (a) p-config., (b) s-config.

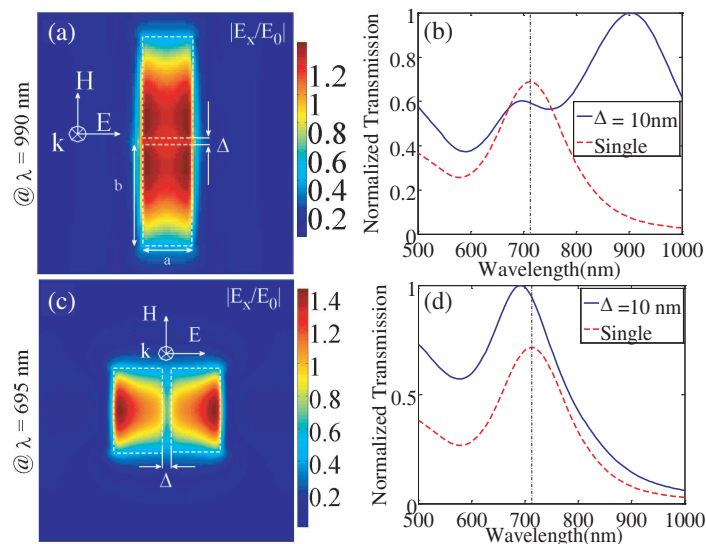


Figure 4: (a), (c) The field profile of electric field component of a rectangular Au nanoparticle in its LSP wavelength, (b), (d) normalized transmission for each configuration.

ture is called serial (s-config.) and/or parallel (p-config.) configuration, respectively. Fig. 3(a) and Fig. 3(b) illustrate the schematic representation for two nanoholes with p- and s-config. arrangements. In addition, the normalized electric field for s- and p-config. nanoholes is depicted in Fig. 4(a) and Fig. 4(c), correspondingly.

Figures 4(a) and 4(c) show the normalized transmission of the two s- and p-config. nanoholes with the separation distance Δ , respectively. According to Fig. 9(a) the second (long wavelength) resonant mode blue shifts and the energy of the coupled magnetic dipoles increases at this mode, by increasing the distance from $\Delta = 5$ to 15 nm. In contrast, the first resonant mode experiences a trivial red shift, by expanding Δ . The controlled increasing the distance leads to contribution of more free electrons in coupling mechanism and a constructive interaction occurs. The short wavelength minimum peak in transmission spectrum, at $\lambda = 580$ nm, is due to Au interband absorption [1], which decreases the normalized transmission amplitude. Moreover, increasing Δ enhances the antisymmetric plasmon hybrid resonances at the second mode wavelength.

This issue compensates the expected intensification of the Au interband absorption, through increasing the separation. Fig. 9(b) shows the normalized transmission for a pair of nanoholes in the p-config. structure. It can be seen that this configuration is not proper for increasing the interaction of the nanoholes, with similar dimensions and distances.

3. CONCLUSION

In summary, the nature of the coupling mechanism between the two nanoholes identified using electric and magnetic normalized power amplitudes and confirmed with quasi-static approximation method. According to the incident beam polarization, if the polarization of the incident magnetic field is directed along the nanohole's axis, two peaks are obvious which are related to the in-phase and anti-phase magnetic dipole radiations. Instead of changing the polarization, two structures for arranging the subwavelength apertures are defined as s- and p-config. For the s-config. arrangement of the nanoholes, and with the separation distance $\Delta = 10$ nm two distinct resonant peaks are excited at the normalized transmission spectrum at $\lambda = 697$, and 904 nm which are due to the in-phase and anti-phase interactions of the magnetic dipoles related to each nanohole. However, for the p-config. arrangement one resonant peak occurs at $\lambda = 695$ nm for the conditions same as the s-config.

REFERENCES

1. Pakizeh, T. and M. Käll, "Unidirectional ultracompact optical nanoantennas," *Nano Lett.*, Vol. 9, No. 6, 2343–2349, 2009.
2. Maier, S. A., P. G. Kik, H. A. Atwater, S. Meltzer, E. Harel, B. E. Koel, and A. A. G. Requicha, "Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides," *Nature Mater.*, Vol. 2, 229–232, 2003.
3. Bethe, H. A., "Theory of diffraction by small holes," *Phys. Rev.*, Vol. 66, Nos. 7–8, 163–182, 1944.
4. Bohren, C. F. and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley-Interscience, New York, 1983.
5. Pozar, D. M., *Microwave Engineering*, 3rd Edition, Wiley-Interscience, New York, 1997.
6. Taflov, C. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, Artech-House, 2005.

Novel All-optical Logic Gates Based on Microring Metal-insulator-metal Plasmonic Waveguides

Yaw-Dong Wu, Yung-Ta Hsueh, and Tien-Tsorng Shih

Department of Electronic Engineering

National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan, R.O.C.

Abstract— In this paper, we proposed novel all-optical logic gates based on microring metal-insulator-metal (MIM) plasmonic waveguides. The proposed all-optical logic gates were numerically studied by the finite-difference time-domain method (FDTD). In recent years, surface plasmon has attracted much attention due to the potential of applications. Surface plasmons are existed at the interface of metals and dielectrics. It has highly localized to a metal surface that can have applications as sub-wavelength waveguides to guide light below the diffraction limit in conventional optics. The MIM structures consist of a dielectric waveguide and two metallic claddings, which strongly confine the incident light in the insulator region. MIM waveguides are prospective for the design of nanoscale all-optical devices shown strong localization, as well as relatively simple fabrication. Some devices based on the MIM waveguides have been studied numerically and experimentally, such as the filters based on ring resonators, tooth-shaped plasmonic waveguide filters, nanodisk resonator, Y-shaped combiners, and wavelength selective waveguides. We designed a MIM structure which consists of the symmetric ring resonators and straight waveguides. By utilizing the coupling property between straight waveguides and ring resonator waveguides and modulating the radii of the ring resonators, we can design an all-optical NOT logic gate. By changing the state of the control port, we can make outgoing field propagating in the output waveguide or not. According to the numerical results, the proposed device could really function as an all-optical NOT logic gate. It would be a potential key component in the application of the all-optical signal processing system.

1. INTRODUCTION

Surface plasmons are existed between metal and dielectric of surface electromagnetic waves. One of the surface plasmon polariton (SPP) important characteristics is that electromagnetic wave can couple to propagating free electron oscillations at metal-dielectric interfaces [1]. SPPs have promising application on the devices in highly integrated optical circuits because they overcome the conventional diffraction limit and can manipulate light on sub-wavelength scales [2]. SPP has been proposed several sub-wavelength optical devices, such as all-optical switches [3], beam manipulator [4], modulators [5], sensors [6, 7], and metallic nanowires [8]. The MIM structures consist of a dielectric waveguide and two metallic claddings, which strongly confine the incident light in the insulator region [9]. In this paper, we designed a MIM structure which consists of the microring resonators and straight waveguides. We use two straight waveguides and two microring resonators to construct NOT logic gate based on MIM waveguide structure.

2. ANALYSES AND NUMERICAL RESULTS

In general, the interface between semi-infinite materials having positive and negative dielectric constants can effectively guide transverse magnetic (TM) surface waves. Because the width of the MIM plasmonic waveguide is much smaller than the wavelength, only the fundamental transverse magnetic (TM) waveguide mode can propagate. The dispersion equation for TM mode in the waveguide is given by [10]:

$$\varepsilon_d k_m + \varepsilon_m k_d \tanh\left(\frac{k_d}{2}w\right) = 0 \quad (1)$$

where k_d and k_m are define as: $k_d = (\beta^2 - \varepsilon_d k_0^2)^{\frac{1}{2}}$ and $k_m = (\beta^2 - \varepsilon_m k_0^2)^{\frac{1}{2}}$. ε_d and ε_m are, respectively, dielectric constants of the insulator and the metal. $k_0 = 2\pi/\lambda$ is the free-space wave vector. The propagation constant β is represented as effective index $n_{eff} = \beta/k_0$ of the waveguide for SPP.

In the paper, the dielectric is assumed to be air with $\varepsilon_d = 1$, and the metal to be silver. The dielectric constant ε_m of silver can be calculated by Drude model [11]:

$$\varepsilon_m(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega(\omega + i\gamma)} \quad (2)$$

where the dielectric constant at infinite angular frequency $\epsilon_\infty = 3.7$, the bulk plasma frequency $\omega_p = 1.38 \times 10^{16}$ Hz, which represents the natural frequency of the oscillations of free conduction electrons, the damping frequency of the oscillations $\gamma = 2.73 \times 10^{13}$ Hz, and ω is the angular frequency of the incident electromagnetic radiation. The SPPs are excited with inputting a TM-polarized plane wave. The transmission of the structure is defined as $T = P_{tr}/P_{in}$ [12]. P_{in} presents the total incident power, and P_{tr} is transmission power.

We use two straight waveguides and two ring resonators to construct NOT logic gate based on the MIM waveguide structure. The structure of the proposed NOT logic gate is shown in Fig. 1. The incident wavelength λ is 850 nm, the radii of the the ring resonator are $R_1 = 414$ nm and $r = 364$ nm, the width of the straight waveguide w is 50 nm, and the coupling distances d between the ring resonators and straight waveguides is 15 nm. In our proposed logic gate, we require one signal port and this signal port must always be in ON state. We can change the state of the input port to determine the state of the output port. When the state of the input port A is OFF, the state of the output port is ON. The magnetic field distribution is shown in Fig. 2(a). The transmission efficiency is shown in Fig. 2(b). When the state of the input port A is ON, the state of the output port is OFF. The magnetic field distribution is shown in Fig. 3(a) and the transmission efficiency is shown in Fig. 3(b). According to the simulation results, the truth table and transmission efficiency that we obtained is shown in Table 1.

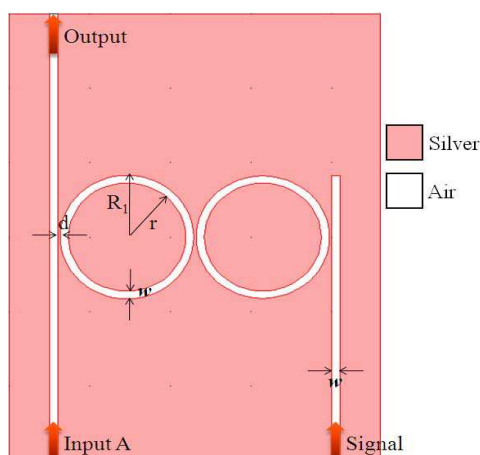


Figure 1: The proposed structure for plasmonic NOT gate.

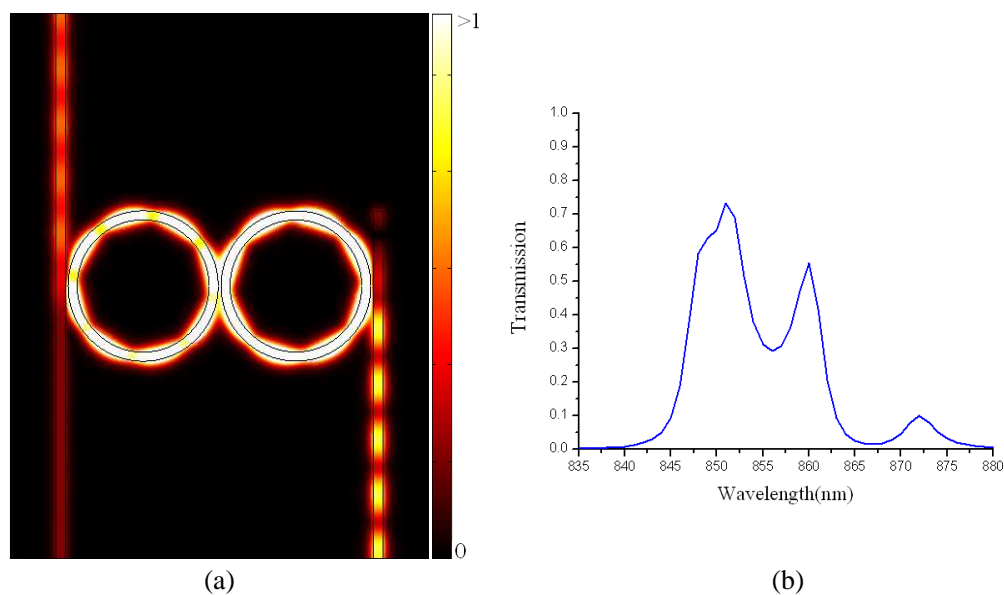


Figure 2: When the input port A = 0, (a) the magnetic field distribution and (b) the normalized transmission spectrum.

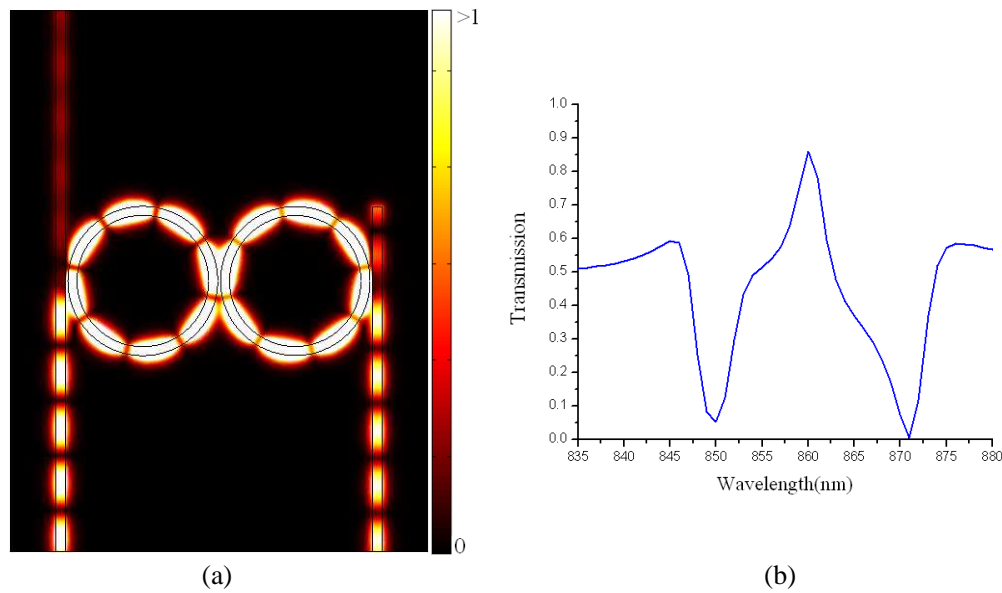


Figure 3: When the input port $A = 1$, (a) the magnetic field distribution and (b) the normalized transmission spectrum.

Table 1: The truth table of the NOT gate.

INPUT A	OUTPUT	T
0	1	65.35%
1	0	5.41%

3. CONCLUSION

In conclusion, we have proposed and numerically investigated a novel NOT logic gate in MIM plasmonic structure. By utilizing the coupling property between straight waveguides and ring resonator waveguides and modulating the radii of the ring resonators, we can design an all-optical NOT logic gate. By changing the state of the control port, we can make outgoing field propagating in the output waveguide or not. According to the numerical results, the proposed device could really function as an all-optical NOT logic gate. It would be a potential key component in the application of the all-optical signal processing system.

ACKNOWLEDGMENT

This work was partly supported by National Science Council R.O.C. and under Grant No. 101-2221-E-151-077.

REFERENCES

1. Raether, H., *Surface Plasmon on Smooth and Rough Surfaces and Gratings*, Springer-Verlag, 1998.
2. Lezec, H. J., A. Degiron, E. Devaux, R. A. Linke, L. Martin-Moreno, F. J. Garcia-Vidal, and T. W. Ebbesen, "Beaming light from a subwavelength aperture," *Science*, Vol. 297, No. 5582, 820–822, 2002.
3. Wang, B. and G. P. Wang, "Surface plasmon polariton propagation in nanoscale metal gap waveguides," *Opt. Lett.*, Vol. 29, No. 17, 1992–1994, 2004.
4. Min, C., P. Wang, X. Jiao, Y. Deng, and H. Ming, "Beam focusing by metallic nano-slit array containing nonlinear material," *Appl. Phys. B*, Vol. 90, No. 1, 97–99, 2008.
5. Nikolajsen, T., K. Leosson, and S. I. Bozhevolnyi, "Surface plasmon polariton based modulators and switches operating at telecom wavelengths," *Appl. Phys. Lett.*, Vol. 85, No. 24, 5833, 2004.

6. Enoch, S., R. Quidant, and G. Badenes, “Optical sensing based on plasmon coupling in nanoparticle arrays,” *Opt. Express*, Vol. 12, No. 15, 3422–3427, 2004.
7. Van Oosten, D., M. Spasenović, and L. Kuipers, “Nanohole chains for directional and localized surface plasmon excitation,” *Nano Lett.*, Vol. 10, No. 1, 286–290, 2010.
8. Dickson, R. M. and L. A. Lyon, “Unidirectional plasmon propagation in metallic nanowires,” *J. Phys. Chem. B*, Vol. 104, 6095–6098, 2000.
9. Veronis, G., Z. Yu, S. E. Kocabas, D. A. B. Miller, M. L. Brongersma, and S. Fan, “Metal-dielectric-metal plasmonic waveguide devices for manipulating light at the nanoscale,” *Chin. Opt. Lett.*, Vol. 7, 302–308, 2009.
10. Dionne, J. A., L. A. Sweatlock, H. A. Atwater, and A. Polman, “Plasmon slot waveguides: Towards chip-scale propagation with subwavelength-scale localization,” *Phys. Rev. B*, Vol. 73, No. 3, 035407, 2006.
11. Han, Z., E. Forsberg, and S. He, “Surface plasmon Bragg gratings formed in metal-insulator-metal waveguides,” *IEEE Photon. Technol. Lett.*, Vol. 19, 91–93, 2007.
12. Lin, X. and X. Huang, “Tooth-shaped plasmonic waveguide filters with nanometric sizes,” *Opt. Lett.*, Vol. 33, No. 23, 2874–2876, 2008.

Characteristic Planes of Microstrip and Unilateral Finline Tee-junctions

A. Casanueva¹, A. Leon¹, A. Mediavilla¹, and J. Helszajn²

¹Universidad de Cantabria, Santander, Spain

²Heriot-Watt University, Edinburgh, UK

Abstract— A unique set of reference planes which enters into the description of a three port E or H -plane tee-junction are its characteristics planes. These planes correspond to the positions of a short circuit at one typical port that will decouple a second port from one input port. The purpose of this paper is to establish these planes in the cases of microstrip and unilateral finline circuits. This is done in each case for a number of different geometries. The spacing between the first two characteristic planes across the junction is 360 deg. In the case of the H -plane tee junction, it is 180 deg in the case of the E -plane geometry.

1. INTRODUCTION

The parameters of E and H -plane tee-junctions are usually characterized at either the intersection between the symmetry planes of the main and side waveguides or at the physical opening of the junction. One other possibility is to adopt the characteristics planes for this purpose. Such planes provide a universal set of terminals in this class of network in both, microstrip line and waveguide. The concept of characteristics planes was first introduced in [1]. It has recently been applied in the descriptions of E and H -plane waveguide tee-junctions in WR-75 [2]. There is one such plane in the description of a wye-junction with three fold symmetry and there are two such planes in the definition of a right angled tee-junction. The purpose of this paper is to characterize microstrip in the H -plane and finline tee-junctions in the E -plane using this notation. This is done for both, parametric values of impedance levels and relative dielectric constants of the substrates. The scattering matrices of the circuits are separately calculated. This is done at an arbitrary frequency interval between 3 and 6 GHz. These sort of junctions enter in the construction of filters circuits and also in the design of rat races and other microwave circuits. The two geometries under consideration are illustrated in Figures 1 and 2. The geometry in Figure 1 is an example of an H -plane arrangement, that in Figure 2 is one of an E -plane structure. The eigenvalue problem of the H -plane circuit has been touched in [3]. Some typical works on tee-junctions are described in [4–15].

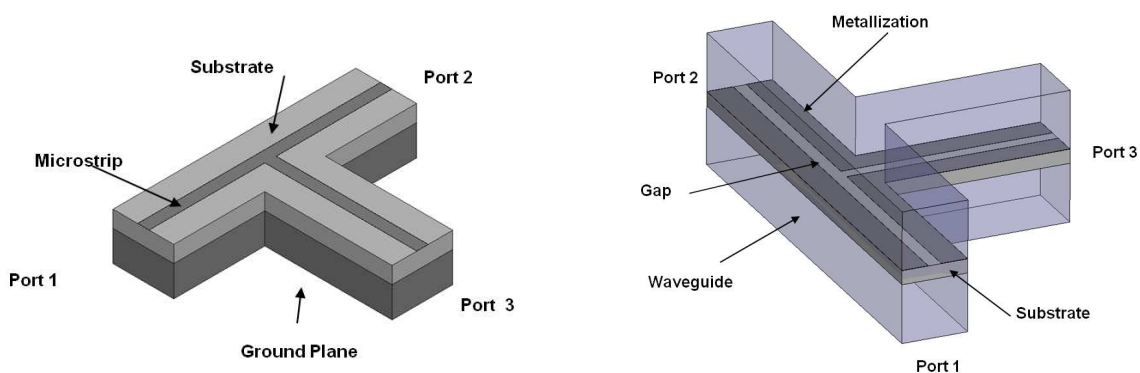


Figure 1: Schematic diagram of microstrip tee-junction.

Figure 2: Schematic diagram of finline tee-junction.

2. CHARACTERISTIC PLANES OF H -PLANE TEE JUNCTION

A common 3-port microstrip junction is an H -plane right angled tee geometry. Its scattering matrix is

$$\bar{S} = \begin{bmatrix} \alpha & \delta & \gamma \\ \delta & \alpha & \gamma \\ \gamma & \gamma & \beta \end{bmatrix} \quad (1)$$

The network under consideration is reciprocal so that the matrix is symmetric about the main diagonal. A question arising with the description of this type of junction is the locations of its reference planes. One unique possibility is to adopt characteristic planes for this purpose. These planes have the positions of a short circuit at one typical port that will decouple a second port from an input one. There are two such planes in the description of the tee junction.

$$\alpha - \delta = -1 \quad (2a)$$

$$\beta - \frac{\gamma^2}{\delta} = -1 \quad (2b)$$

The first equation defines the characteristic planes in ports 1 and 2 of the junction; the second fixes that in port 3. Figures 3 and 4 illustrate typical in microstrip junction, the locations of the characteristic planes and the experimental arrangement.

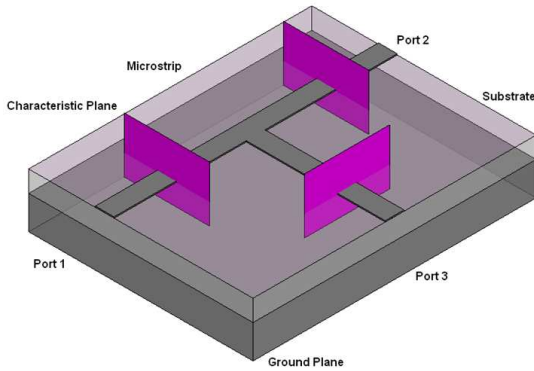


Figure 3: Characteristics planes of microstrip tee junction.

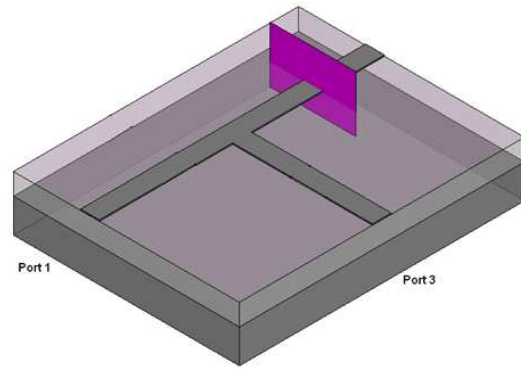


Figure 4: Experimental arrangement for measurement of a typical characteristic plane.

The two characteristic planes are equi-distance from the opening of the junction provided

$$\alpha - \delta = \beta - \frac{\gamma^2}{\delta} \quad (3)$$

This condition is satisfied provided

$$\frac{\gamma}{\delta} = \sqrt{2} \quad (4)$$

The nature of the scattering matrix of the E -plane tee junction is different and is dealt with separately.

3. CHARACTERIZATION OF H -PLANE MICROSTRIP TEE-CIRCUITS

It is necessary in order to tabulate universal data in an unambiguous way in the descriptions of right angled tee-junctions to express the results at its characteristics planes. This may be done by defining electrical angles.

$$\theta_i = \frac{2\pi d_i}{\lambda_g} \quad i = 1, 2, 3, \dots \quad (5)$$

The coordinate system adopted by J. T. Allanson et al. is that at the intersection between the symmetry planes of the main waveguide and the side waveguide. Other authors, such as Mansour have chosen the reference planes to coincide with the opening of the waveguides of the junction. The convention adopted here is to place the reference terminals at the characteristics planes of the junction. It is indicated in Figure 6 in the case of microstrip.

In order to carry out the fast efficient eigenmodes dispersion analysis for multilayer and multiconductor planar tlines, the hybrid fields are expressed by superposing TE-to and TM-to- with Hertzian scalar potentials. The solution to the boundary value problem describing electromagnetic wave propagation in the z direction in the structure under investigation can be found using SVD method [18].

The guide wavelength in the transmission line may be accurately deduced by recognizing that the characteristic planes repeat every half wavelength at a single frequency. It provides therefore

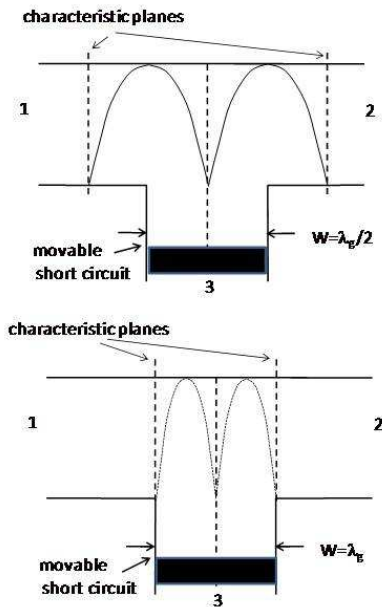


Figure 5: Characteristics planes at two different frequencies.

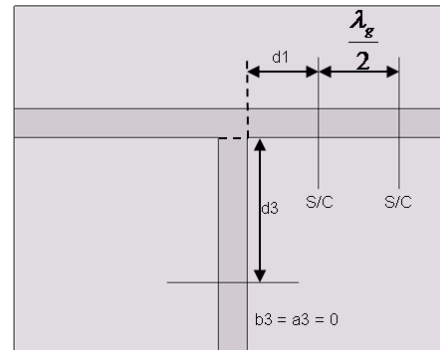


Figure 6: Definition of characteristic planes in right angled microstrip tee-junction.

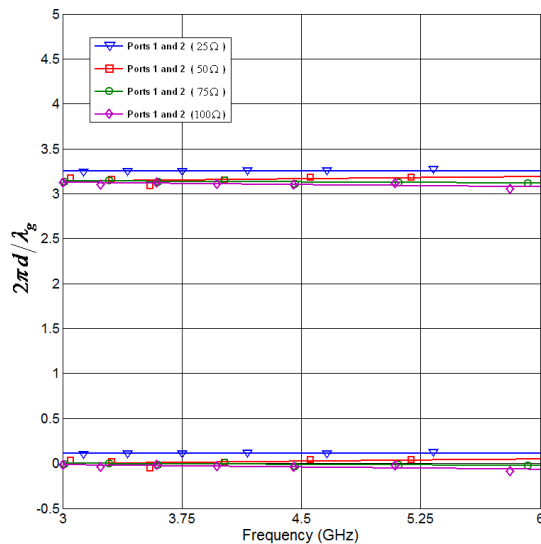


Figure 7: Calculated characteristic planes at port 1 or 2 of the microstrip tee-junction over the frequency range of 3 to 6 GHz.

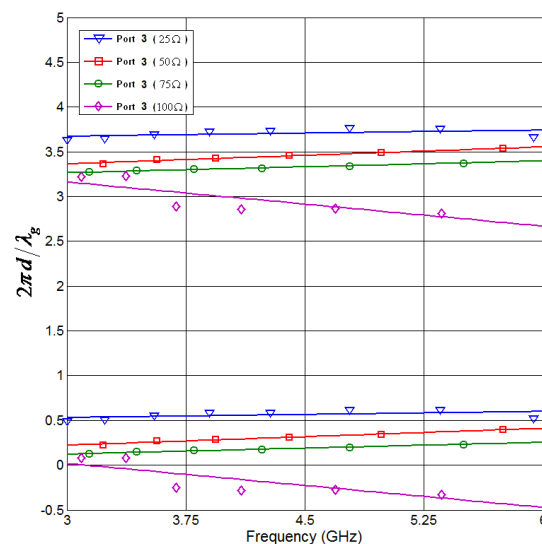


Figure 8: Calculated characteristic planes at port 3 of the microstrip tee-junction over the frequency range of 3 to 6 GHz.

one means of measuring or calculating the wavelength of the transmission line, including fringing effects. This allows the experimental reference plane to be chosen at will and translate it to that of the junction.

The characteristics planes of the structure in Figure 1 are summarized in Figures 7 and 8.

This is done for parametric values of impedance between 25 and 100 Ω over a frequency band between 3 and 6 GHz. The dielectric constant of the substrate is 3.12. The corresponding standing wave patterns are illustrated in Figures 9(a) and (b). The midband scattering parameters at the midband and edge band frequencies at the first characteristic planes are

$$\left. \begin{aligned} \alpha &= 0.942 \angle -176.13^\circ \\ \beta &= 0.920 \angle -179.14^\circ \\ \gamma &= 0.286 \angle 87.78^\circ \\ \delta &= 0.147 \angle 67.14^\circ \end{aligned} \right\} \text{1st}$$

and for the other characteristic plane

$$\left. \begin{aligned} \alpha &= 0.302 \angle -21.25^\circ \\ \beta &= 0.390 \angle 129.32^\circ \\ \gamma &= 0.638 \angle 53.01^\circ \\ \delta &= 0.692 \angle 157.72^\circ \end{aligned} \right\} \text{2nd}$$

The scattering parameters obtained in this way satisfy the unitary condition.

$$\begin{aligned} |\beta|^2 + 2|\gamma|^2 &= 1.010 && \text{1st} \\ &= 0.966 && \text{2nd} \\ |\alpha|^2 + |\gamma|^2 + |\delta|^2 &= 0.991 && \text{1st} \\ &= 0.977 && \text{2nd} \end{aligned}$$

$|\gamma|$ these should be chosen to 1, 0.

The amplitudes of these quantities are of course independent of the choice of the reference planes adopted. Figure 9 illustrates the electromagnetic field patterns with short circuits pistons at the characteristic planes.

Exchanging the generator and short circuit conditions between ports 1 and 3 leaves in a reciprocal circuit the field patterns unchanged.

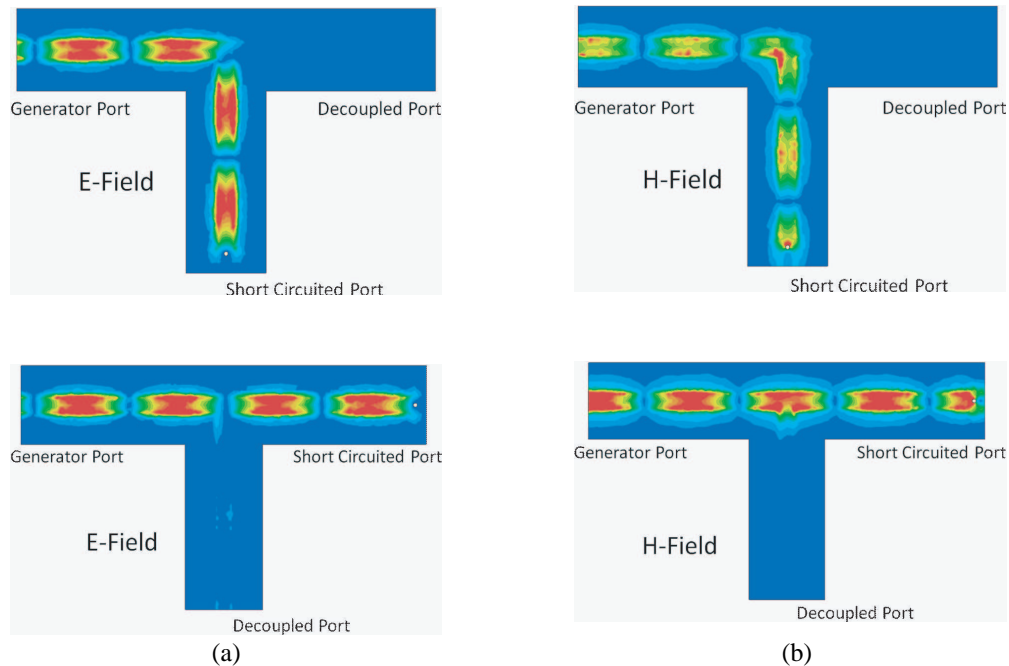


Figure 9: (a) Standing wave solution of electric field in the H -plane microstrip. (b) Standing wave solution of magnetic field in the H -plane microstrip.

4. E -PLANE UNILATERAL FINLINE TEE-JUNCTION

The unilateral finline waveguide is an example of an E -plane geometry it is described by the a and b dimensions of the rectangular waveguide, by the gap (W) and by the thickness (H) and by the dielectric constant of the substrate (ϵ_r). The relative dielectric constant employed in this work is 2.2 and the waveguide is WR-75. The schematic diagram under consideration is indicated in Figure 10. The definitions of the reference planes are here arbitrarily taken at the openings of the waveguide of the junction rather than at those of the finline circuit. This is shown in Figure 11. The analysis of the unilateral finline can be solved using various numerical methods. The initial conditions of the structure are usually based on some existing approximate closed form expressions [16, 17] in conjunction with a spectral domain approach (SDA). It is here, however, fixed by having recourse to a full-wave simulator. This type of planar transmission line differs from the microstrip case in that it is a dispersive structure.

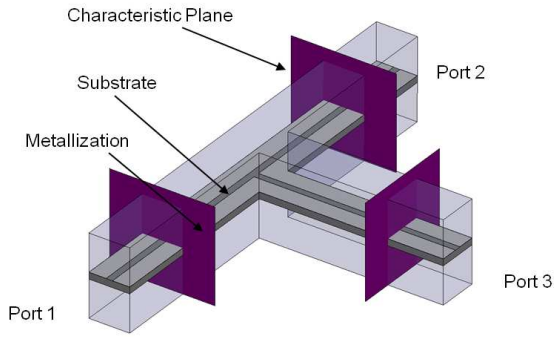


Figure 10: Characteristics planes of the E -plane finline tee-junction.

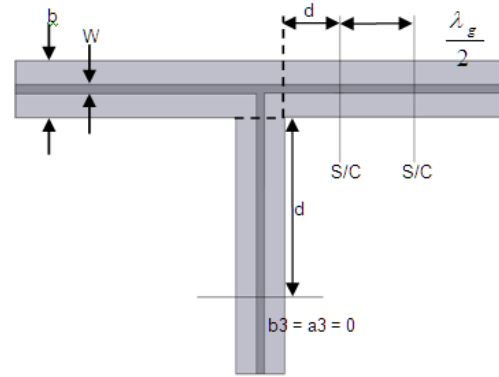


Figure 11: Definition of characteristic planes of right angled finline tee-Junction.

The scattering matrix of differs from that of the H -plane geometry in that the phases of S_{31} and S_{32} are 180 degrees out of phase with those S_{13} and S_{31} [2].

$$\bar{S} = \begin{bmatrix} \alpha & \delta & \gamma \\ \delta & \alpha & -\gamma \\ \gamma & -\gamma & \beta \end{bmatrix}$$

The reflection coefficients at the points of this main waveguide and at the side one is in this instance given by

$$\rho_{1,2} = \alpha + \delta = -1 \quad (6a)$$

$$\rho_3 = \beta + \frac{\gamma^2}{\delta} = -1 \quad (6b)$$

Figures 12 and 13 indicate the connection between the characteristics planes between 9 to 15 GHz of the unilateral finline circuit in Figure 2 at ports 1 or 2 and at port 3. In the construction of the frequency response of this type of circuits the guide wavelength adopted is, of course, that at the midband frequency.

The midband scattering parameters at the characteristics planes are

$$\alpha = 0.339 \angle 64.74^\circ$$

$$\beta = 0.398 \angle 113.76^\circ$$

$$\gamma = 0.652 \angle 114.89^\circ$$

$$\delta = 0.676 \angle -87.13^\circ$$

This gives

$$|\beta|^2 + 2|\gamma|^2 = 1.008612$$

and

$$|\alpha|^2 + |\gamma|^2 + |\delta|^2 = 1.04$$

The standing wave patterns of the different possible experimental arrangements are shown in Figures 14(a) and (b).

The lack of negative signs in the description of the scattering matrix of the H -plane tee junction compared to that of the E -plane geometry produces on electric field in the inner box of the junction which is zero whereas it produces the dual condition in the latter arrangement.

One important consequence of this situation is that the spacing between the dominant characteristic planes is 360 deg. In the case of H -plane junction and 180 deg. in the other case.

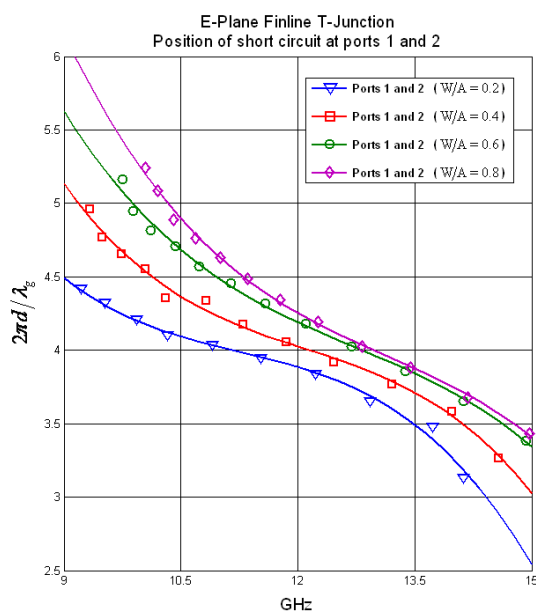


Figure 12: Calculated characteristic planes of the E -plane unilateral finline tee-junction at ports 1 and 2 between the frequency range of 9 to 15 GHz.

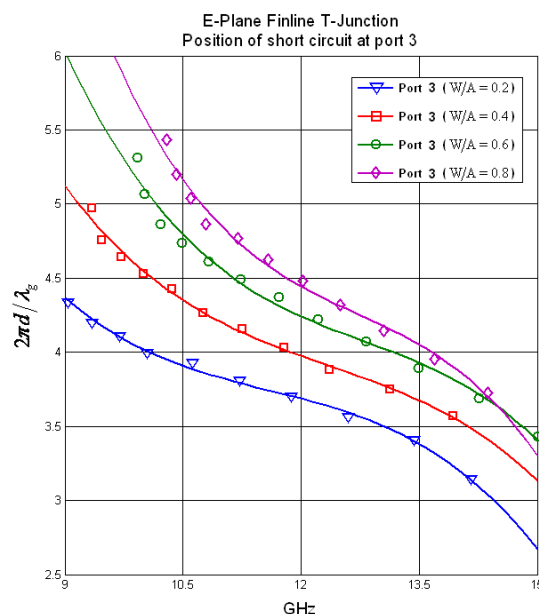


Figure 13: Calculated characteristic planes of the E -plane unilateral finline tee-junction at port 3 between the frequency range of 9 to 15 GHz.

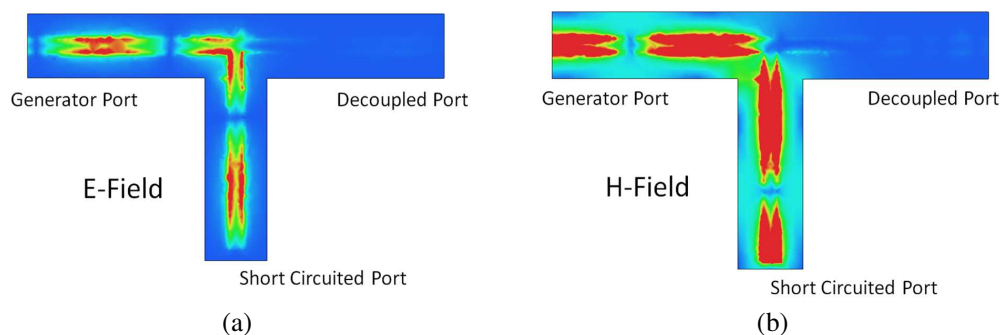


Figure 14: (a) Standing wave solution of electric field in E -plane unilateral finline. (b) Standing wave solution of magnetic field in E -plane unilateral finline.

5. CONCLUSIONS

The purpose of this paper is to establish the characteristic planes of microstrip and finline circuits. A knowledge of these planes allow unique representations of these type of circuits. The former is an example of an H -plane or shunt network and the latter is one of an E -plane or series tee-junction.

ACKNOWLEDGMENT

The authors would like to thank the Spanish Ministry of Science and Innovation (MICINN) by the financial support provided through project TEC2011-29126-C03-01 and CONSOLIDER-INGENIO 2010 CSD2008-00068.

REFERENCES

1. Allanson, J. T., R. Cooper, and T. G. Cowling, "The theory and experimental behaviour of right-angled junctions in rectangular-section wave guides," *IEE Proc.*, 1946.
2. Helszajn, J., M. Caplin, J. Frenna, and B. Tsounis, "Characteristics planes, scattering matrices and eigenvalues of E and H -plane waveguide tee junctions," *IEEE MTT International Conference*, 2012.
3. Montgomery, C. G., R. H. Dicke, and E. M. Purcell, *Principles of microwave Circuits*, McGraw-Hill Co., New-York, 1948.

4. Franco, A. G. and A. A. Oliner, "Symmetric strip transmission line tee junction," *IRE Trans. Microwave Theory and Tech.*, Vol. 10, No. 2, 118–124, 1962.
5. Mansour, R. R. and J. Dude, "Analysis of microstrip T-junction and its application to the design of transfer switches," *IEEE Conference*, 1992.
6. Uher, J., J. Bornemann, and U. Rosenberg, *Waveguide Components for Antenna Feed Systems: Theory and CAD*, Artech House, Norwood, MA, 1993.
7. Yao, H.-W., A. E. Abdelmonem, J.-F. Liang, X.-P. Liang, K. A. Zaki, and A. Martin, "Wide-band waveguide T-junctions for diplexer applications," *IEEE Trans. Microwave Theory and Tech.*, Vol. 41, No. 12, 2166–2173, 1993.
8. Branner, G. R., B. P. Kumar, and D. G. Thomas, Jr., "Design of microstrip T junction power divider circuits for enhanced performance," *IEEE Conference*, 1995.
9. Jia, H., K. Yoshitomi, and K. Yasumoto, "Rigorous analysis of E/H -plane junctions in rectangular waveguides using Fourier transform technique," *Progress In Electromagnetics Research*, Vol. 21, 273–292, 1999.
10. Rong, Y., H. Yao, K. A. Zaki, and T. G. Dolan, "Millimeter-wave Ka-band H -plane diplexers and multiplexers," *IEEE Trans. Microwave Theory and Tech.*, Vol. 47, No. 12, 2325–2330, Dec. 1999.
11. Jia, H., K. Yasumoto, and K. Yoshitomi, "Fast and efficient analysis of inset dielectric guide using Fourier transform technique with a modified perfectly matched boundary," *Progress In Electromagnetics Research*, Vol. 34, 143–163, 2001.
12. Hashemi-Nasab, M. and A. Cheldavi, "Coupling model for two orthogonal microstrip lines in two layer PCB board (quasi-TEM approach)," *Progress In Electromagnetics Research*, Vol. 60, 153–163, 2006.
13. Saenz, E., A. Cantora, I. Ederra, R. Gonzalo, and P. de Maagt, "A metamaterial T-junction power divider," *IEEE Microwave and Wireless Components Letters*, Vol. 17, No. 3, 2007.
14. Yang, R.-Y., C.-M. Hsiung, C.-Y. Hung, and C.-C. Lin, "Design of a high band isolation diplexer for GPS and WLAN system using modified stepped-impedance resonators," *Progress In Electromagnetics Research*, Vol. 107, 101–114, 2010.
15. He, F. F., K. Wu, W. Hong, L. Han, and X. Chen, "A planar magic-T structure using substrate integrated circuits concept and its mixer application," *IEEE Trans. Microwave Theory and Tech.*, Vol. 59, No. 1, 72–78, Jan. 2011.
16. Nakajima, M., "A simplified formula for the characteristic impedance of microstrip lines," *Int. J. Electron.*, Vol. 44, 61–64, 1978.
17. Bhat, B. and S. K. Koul, *Analysis, Design and Applications of Finlines*, Artech House, Massachusetts, 1987.
18. Casanueva, A. and J. L. García, "An innovative fast powerful method for tackling electromagnetic eigenvalue problems for multistrip transmission lines," *IEEE Trans. Microwave Theory and Tech.*, Vol. 50, 36–40, Jan. 2002.

A K-band Direct Injection-locked Frequency Divider Using Harmonic LC Tanks

Chien-Pai Wu and Yen-Chung Chiang
National Chung Hsing University, Taiwan

Abstract— A K-band injection-locked frequency divider implemented in the TSMC 0.18 μm CMOS process technology is proposed in this paper. By using the harmonic LC tank technique, the third order harmonic of the output signal from the frequency divider is enhanced and a balanced performance between the locking range and the output swing can be achieved. Compared with the conventional direct injection-locked frequency divider, the locking speed of the proposed frequency divider can also be improved. The chip size of the proposed frequency divider is $0.711 \times 0.776 \text{ mm}^2$. The measured locking range is from 24.8 to 26.6 GHz and the core circuit draws a 5.4 mW power from a 1.8-V supply.

1. INTRODUCTION

Recently, the rapid development of wireless communication technology has made our daily life easier and funnier. Multi-media, videos and information data can be transferred via portable devices with a high data rate. Some systems operating above 20 GHz, including the local multipoint distribution service (LMDS), the short range vehicle crash prevention radars, wireless local area networks, and other ISM band applications, have been proposed. Therefore, the ratio-frequency (RF) applications are getting more and more important. Furthermore, due to the progress of semiconductor technology, RF front-end system can be integrated in a single IC, and then employed in cellular phones, personal computers and other hand-held devices. In high-speed wireless communication systems, the phase-locked loop (PLL) is a key component utilized as a frequency synthesizer and/or a quadrature signal generator [1]. PLL can generate a stable and controllable output signal whose frequency is locked by the input reference signal. The design considerations of a PLL include the fast-locking duration and the low-jitter performance. A PLL is generally composed of the voltage control oscillator (VCO), the phase-frequency detector (PFD), the charge pump (CP), and the frequency divider (FD). Among these sub-blocks, one essential circuit is the FD which takes a periodic input signal and generates a periodic output signal at a frequency being a fraction of that of the input.

Several existing techniques are used for frequency division: e.g., common-mode logic (CML) based [2], dynamic logic based [3], injection locked FDs (ILFD) [4], and Miller dividers [5]. As indicated in [6], every technique has its own pros and cons. For example, the mechanism of a CML FD is easy to design and to understand, however, it suffers from higher power consumption and lower operation frequency. On the other hand, the ILFDs dissipate lower power with better noise performance, but their locking range is relatively small. Figures 1 and 2 show two conventional ILFDs: current tail injection and direct injection, respectively. The conventional current tail injection ILFD shown in Figure 1 is basically a LC tank VCO with the input signal injected by current tail

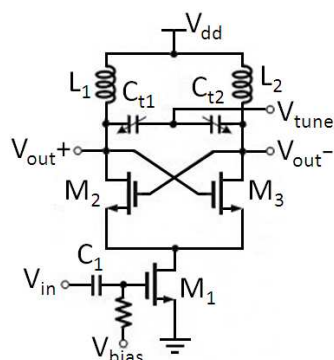


Figure 1: Schematic of a conventional ILFD by current tail injection.

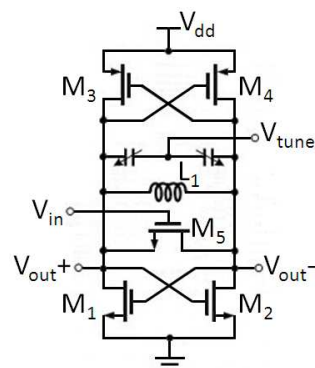


Figure 2: Schematic of a conventional ILFD by direct injection.

transistor M_1 . The limited locking range of such ILFD is due to the inefficient injecting path from the tail transistor M_1 . To overcome this problem, a direct injection ILFD as shown in Figure 2 was proposed in [6]. The input signal is directly injected into the oscillating loop thus the locking range can be improved.

The CMOS process technology has become a strong candidate for wireless communication system due to its low cost and easy integration. In this paper, we implement a ILFD in the 0.18 μm CMOS process in which we combine the direct injection ILFD with the harmonic LC tank technique for improving the locking speed and the locking range. The detailed design consideration is given in Section 2, and the experimental results are shown in Section 3. Finally, a simple conclusion is made in Section 4.

2. CIRCUIT DESIGN

Although the differential architecture proposed in [6] can improve the locking range, it still causes other problems. The output power of differential injection-locked frequency divider is low, especially at high frequency operation. For this reason, it is hard to drive the next stage and the required locking time is relatively long. To improve the locking speed and output power, we adopt the harmonic LC tanks to the differential ILFD to enhance the third order harmonic. Figure 3 shows the proposed ILFD in which we also use the complementary cross-coupled pair architecture formed by transistors M_1 , M_2 , M_3 and M_4 as the basic oscillating core for increasing the required transconductance for start-up condition.

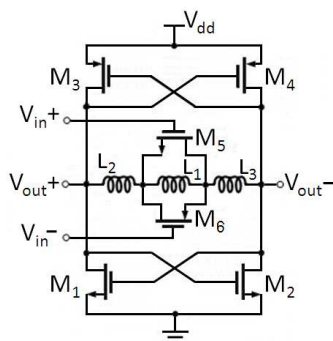


Figure 3: Schematic of the proposed direct injection frequency divider with harmonic tank.

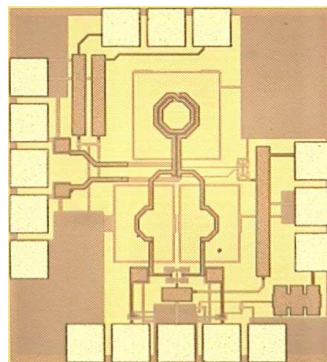


Figure 4: Micrograph of the proposed ILFD.

Instead of using only fundamental inductor L_1 , the third harmonic LC tank consist of parasitic capacitors and high quality factor (Q) inductors L_2 and L_3 is added as shown in Figure 3. The additional inductors combined with the fundamental inductor L_1 will form a network with two peaks in the frequency response: one at the output frequency and the other at the third harmonic frequency. The fundamental inductor L_1 is implemented as the symmetric type inductor with coil width being 9 μm , inner radius being 43.5 μm , and the turn number being 2. The additional harmonic inductors L_2 and L_3 are implemented as standard inductors with coil width being 9 μm , inner radius being 30 μm , and the turn number being 0.5. Although the high quality inductors will narrow the locking range of ILFD, the harmonic tanks can generate third harmonic frequency component and then improve the output power. From the simulation results, the locking speed is also twice faster than the traditional differential ILFD.

Although varactors can extend the free running oscillation frequency range and thus increase the locking range of frequency divider, we do not adopt varactors in our design to show the performance of the harmonic tank. Since we would like to take the input as differential signal, the injection devices are NMOS M_5 and PMOS M_6 transistors, respectively. The sizes of input transistors are chosen to be balance between the overall network frequency response and the locking range and both transistors are 1.5 $\mu\text{m}/0.18 \mu\text{m}$ 15 fingers in size. We also adopt the forward body bias technique in the transistors M_5 and M_6 . We use two open-drain stages which do not shown in Figure 3 as output buffers and external bias-T devices are used in the measurement.

3. MEASURED RESULTS

The proposed ILFD is implemented in the $0.18\ \mu\text{m}$ CMOS process technology. The micrograph of the circuit is as shown in Figure 4 and the chip size is $0.711 \times 0.776\ \text{mm}^2$. The core circuit of the proposed ILFD draws a $5.4\ \text{mW}$ dc power from a 1.8-V power supply. The measured free running frequency is $12.77\ \text{GHz}$ with $-7.3\ \text{dBm}$ output power. Figures 5(a) and 5(b) show the measured output spectrums for the injection locked operation upper bound and lower bound, respectively. It can be seen that the locking range is from 24.8 to $26.6\ \text{GHz}$. The measured output power is about $-8.99\ \text{dBm}$ to $-10.38\ \text{dBm}$.

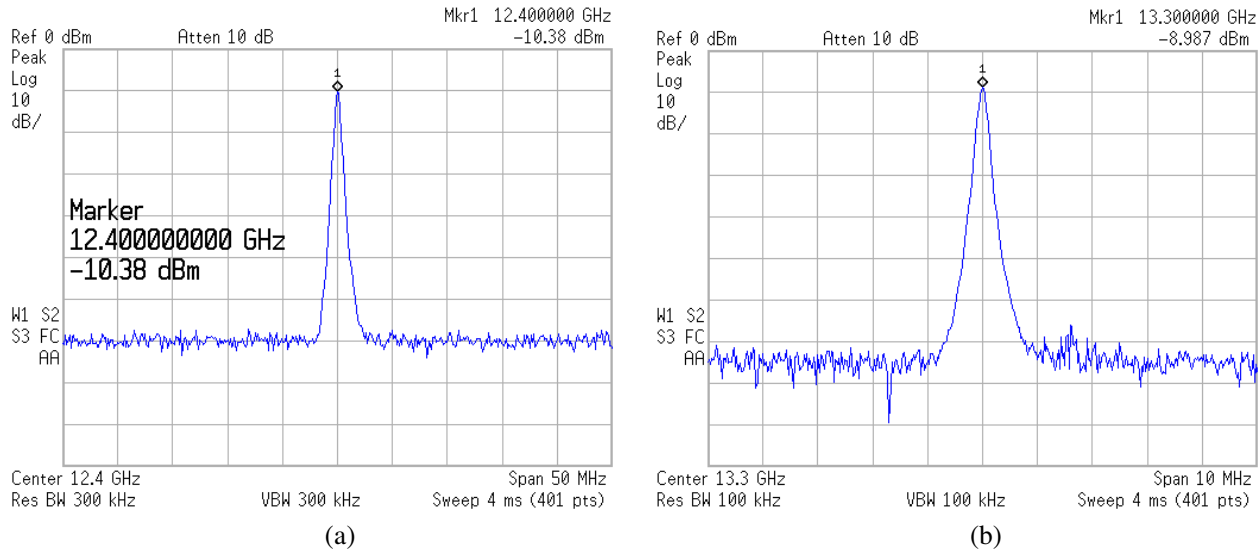


Figure 5: Measured output spectrums for (a) the lower bound of locking range and (b) the upper bound of locking range.

Figure 6 shows the measured phase noise of the proposed ILFD for the free-running output signal, the injected input signal, and the divided output signal. Although the phase noise of the free-running signal is not good, it can be seen that the divided output signal has phase noise about $6\ \text{dB}$ below the phase noise of the injected signal for most of the offset frequencies as expected. Figure 6 shows the measured input sensitivity curve of the proposed ILFD. The locking range is from 24.8 to $26.6\ \text{GHz}$ at the $0\ \text{dBm}$ input power level, while the locking range is from 24.4 to $26.9\ \text{GHz}$ if the input power is increased to $5\ \text{dBm}$.

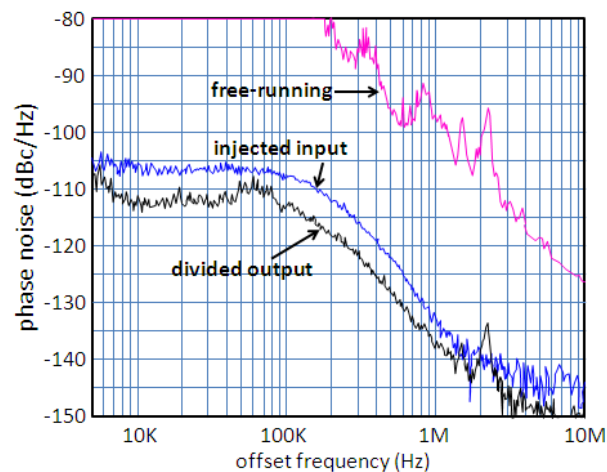


Figure 6: Measured phase noise of the proposed ILFD for free-running, injected input, and divided output signals.

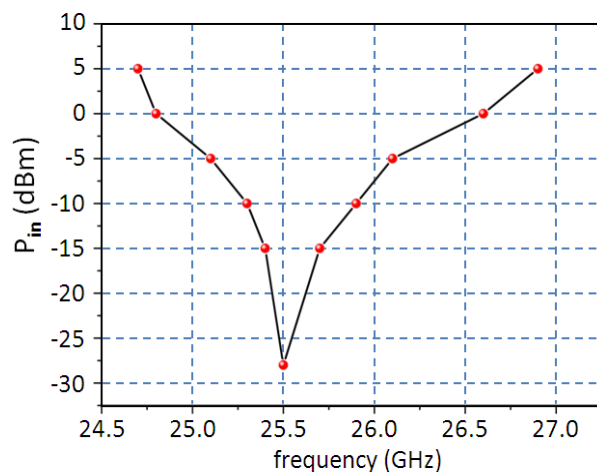


Figure 7: Measured input sensitivity of the proposed ILFD.

4. CONCLUSIONS

In this paper, we have proposed an injection-locked frequency divider designed for the K-band applications in the 0.18 μm CMOS process technology. By using the harmonic LC tank technique, the third order harmonic of the output from the frequency divider is increased and the locking speed can also be improved. The measured locking range is from 24.8 to 26.6 GHz under 0 dBm input power and the core circuit draws 5.4 mW power from a 1.8-V supply.

ACKNOWLEDGMENT

This work was supported in part by the Ministry of Education, Taiwan, R.O.C. under the ATU plan and by the National Science Council of the Republic of China under Grant NSC101-2220-E-005-010. The authors would like to thank the National Chip Implementation Center (CIC) in Hsinchu, Taiwan for the chip fabrication, measurement and other technical supports.

REFERENCES

1. Chuang, Y.-H., S.-H. Lee, R.-H. Yen, S.-L. Jang, J.-F. Lee, and M.-H. Juang, "A wide locking range and low voltage CMOS direct injection-locked frequency divider," *IEEE Microw. Wireless Compon. Lett.*, Vol. 16, No. 5, 299–301, 2006.
2. Craninckx, J. and M. Steyaert, "A 1.75 GHz/3 V dual-modulus divide-by-128/129 prescaler in 0.7- μm CMOS," *Proc. ESSCIRC'95*, 254–257, 1995.
3. Huang, Q. and R. Rogenmoser, "Speed optimization of edge-triggered CMOS circuits for gigahertz single-phase clocks," *IEEE J. Solid-State Circuits*, Vol. 31, No. 3, 456–463, 1996.
4. Quere, R., E. Ngoya, M. Camiade, A. Suarez, M. Hessane, and J. Obregon, "Large signal design of broadband monolithic microwave frequency dividers and phase-locked oscillators," *IEEE Trans. Microw. Theory Tech.*, Vol. 41, No. 11, 1928–1938, 1993.
5. Lee, J. and B. Razavi, "A 40 GHz frequency divider in 0.18- μm CMOS technology," *VLSI Circuits Tech. Dig.*, 259–262, 2003.
6. Tiebout, M., "A CMOS direct injection-locked oscillator topology as high-frequency low-power frequency divider," *IEEE J. Solid-State Circuits*, Vol. 39, No. 7, 1170–1174, 2004.

Electromagnetic Signal Processing for Feature Extraction and Classification of Lossy Dielectric Targets

Gonul Turhan-Sayan and Basak I. Barut

Department of Electrical and Electronics Engineering
Middle East Technical University (METU), Ankara, Turkey

Abstract— This paper presents the application of a real-time electromagnetic target classification technique to recognize dispersive dielectric objects with varying loss characteristics. The suggested classifier design technique is operative at the resonance region and is based on the use of Singularity Expansion Method (SEM) to represent a given electromagnetic scatterer by its natural response. A multi-aspect database of wide band scattered signals are processed to extract target features with the ultimate aim of target characterization in an aspect invariant manner. The Wigner distribution (WD), a quadratic time-frequency transformation, is used to extract pole-related feature vectors from non-stationary scattered signals. The Principal Components Analysis (PCA) is further used to fuse multi-aspect feature vectors to obtain a single unified feature vector for each library object. A suitable late-time design interval needs to be chosen to obtain fused features with minimized aspect sensitivity that is a vital requirement for improved classifier accuracy.

1. INTRODUCTION

Electromagnetic object recognition is a challenging problem due to the aspect and polarization dependent nature of scattered electromagnetic signals. The Singularity Expansion Method (SEM), originally established in 1970's by C.E. Baum, is a well-known modeling tool used to represent the late-time portion of the scattered electromagnetic response waveforms. These aspect and polarization dependent late-time signals are basically composed of the superposition of damped sinusoidal signals oscillating at the complex natural resonance (CNR) frequencies of the object. The CNR frequencies are the poles of the object's system function in complex frequency domain [1]. Collection of these poles can describe the related finite-size object uniquely in an aspect and polarization independent manner as they are determined by the size, shape and material properties of a given object. As the extraction of poles from measured late-time scattered data is highly susceptible to noise, use of alternative indirect feature extraction techniques is found more feasible. This paper is based on the recently suggested WD-PCA method [2, 3] where the extracted object features describe the spectral content of natural response over a properly chosen late-time interval. This technique makes use of the Wigner-Ville Distribution (WD), a quadratic time-frequency transformation, to extract pole-related object features at each available aspect/polarization combination. Then, the Principal Component Analysis (PCA) is used for data reduction and feature fusion. In WD-PCA method, electromagnetic scattered data measured or simulated at only a few different aspects would be enough to design the classifier for a library of preselected objects. Each library object is represented in the classifier's feature database by a single fused feature vector with minimized aspect variance. In the real-time object recognition phase, on the other hand, a scattered signal received at an arbitrary and unknown aspect would be sufficient to classify the test object. In literature, the WD-PCA technique has been successfully applied so far to recognize perfect conductor, perfect dielectric or dielectric coated conducting objects [2–4]. In this paper, this technique will be demonstrated for the first time in literature for objects made of dispersive and lossy dielectric materials.

2. THEORY

Basic theory and fundamental design steps of the WD-PCA based object classification technique are briefly described in this section. In the first step of this design procedure, a database of electromagnetic scattered signals must be generated to contain a total of $K \times M$ different time-domain scattered signals at K reference combinations of aspect angle/polarization for each one of M library targets. Scattered signals can be either measured or synthesized by numerical simulations over a common preselected frequency band in resonance region. Next, the auto Wigner-Ville

distribution (WD) is computed for each design signal $x(t)$ as

$$W_x(t, f) = \int_{-\infty}^{\infty} e^{-j2\pi f\tau} x(t + \tau/2) x^*(t - \tau/2) d\tau \quad (1)$$

where the output represents an energy density function over the joint time-frequency plane, in approximate sense. To characterize the natural response behavior of the target, it is needed to partition the total time span T_0 of the signals into Q equally-wide intervals and use the WD output over a selected late time interval [3] to construct a late-time feature vector (LTFV) defined as $\bar{e} = [\bar{E}_{q^*} \bar{E}_{q^*+1}]$ where \bar{E}_q is the q th partition vector of length $(N/2)$ defined on the q th time interval with its entries computed as

$$E_q(f_p) = \int_{(q-1)\frac{T_0}{Q}}^{q\frac{T_0}{Q}} W_x(t, f_p) dt \quad \text{for } q = 1, 2, \dots, Q, \quad p = 1, 2, \dots, \frac{N}{2} \quad \text{and} \quad f_p = (p-1)\frac{1}{T_0} \quad (2)$$

where N is the number of samples of the discrete time domain scattered signal. The methods for choosing parameters Q and q^* are discussed in detail in [3]. Then, the PCA technique is used to extract main patterns common to LTFV features (whose aspect variances are reduced moderately) computed at K different aspect/polarization combinations for a given library object. To implement the PCA based multi-aspect feature fusion technique, a real valued feature matrix F of size $K \times N$ is formed for each object having the late-time feature vectors \bar{e}_k , $k = 1, \dots, K$ as its rows. Then, the covariance matrix S_F of this feature matrix F is computed which is a $K \times K$ nonsingular, symmetric matrix. Next, an orthonormal matrix $U = [u_1 \ u_2 \ \dots \ u_K]$ is formed where u_i 's are the normalized eigenvectors (of size $K \times 1$) corresponding to the eigenvalues λ_i of the covariance matrix S_F . In this process, the U matrix is formed after ordering the computed eigenvalues such that $\lambda_1 > \lambda_2 > \dots > \lambda_K$ and it can be used to diagonalize the covariance matrix and to transform the correlated feature vectors $\bar{e}_1, \bar{e}_2, \dots, \bar{e}_K$ into a set of uncorrelated vectors z_1, z_2, \dots, z_K by

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_K \end{bmatrix} = U^T \begin{bmatrix} \bar{e}_1 - \text{mean}(\bar{e}_1)I_N \\ \bar{e}_2 - \text{mean}(\bar{e}_2)I_N \\ \vdots \\ \bar{e}_K - \text{mean}(\bar{e}_K)I_N \end{bmatrix} \quad (3)$$

where I_N is an all-ones row vector of length N , the superscript T denotes the transpose operator and the resulting matrix Z (of size $K \times N$) is composed of transformed vectors, z_i 's which are the principal components of the feature matrix F . Each principal component z_i is a zero-mean vector with variance λ_i . Then, the fused feature vector (FFV) can be constructed as a weighted sum of principal components to characterize the associated target in the classifier's feature database.

The same procedure is repeated for each object in the classifier library to completely design the classifier's feature database. Then, during the real-time classification phase, the LTFV of the measured test signal is computed as described above and its correlation coefficient with respect to each FFV in the classifier library is computed to determine the library target with the highest matching score.

3. APPLICATIONS AND RESULTS

In this section, we will demonstrate the design of an object classifier to recognize $M = 3$ different spherical targets T_1, T_2, T_3 which have the same radius ($a = 10$ cm) but they are made of dispersive and lossy dielectric materials with different complex relative permittivity parameters $\varepsilon_r(\omega) = \varepsilon_r'(\omega) + j\varepsilon_r''(\omega)$. In this study, we have chosen three typical types of human breast tissue with permittivity functions described by the Cole-Cole model

$$\varepsilon_r(\omega) = \varepsilon_\infty + \frac{\Delta\varepsilon}{1 + (j\omega\tau)^{(1-\alpha)}} + \frac{\sigma_s}{j\omega\varepsilon_0} \quad (4)$$

with the values of model parameters fitted for each tissue type [5] as shown in Table 1.

Table 1: Cole-Cole parameters for three typical types of breast tissue.

Tissue Type	ε_∞	σ_s (S/m)	$\Delta\varepsilon$	τ (ps)	α
T_1	9.941	0.462	26.60	10.90	0.003
T_2	7.821	0.713	41.48	10.66	0.047
T_3	6.151	0.809	48.26	10.26	0.049

Then, scattered signals for these dispersive library objects are computed in frequency domain (in response to an x -polarized uniform plane wave propagating in $+z$ -direction as described in Figure 1) using the Mie series over the frequency range from almost DC to 19.1 GHz at $\phi = 90$ degrees azimuth angle and for 12 different bistatic aspect angles for $\theta = 15, 30, 45, 60, 75, 90, 105, 120, 135, 150, 165$ and 179 degrees. Out of these 12 different aspects, only $K = 4$ reference aspects at $\theta = 45, 90, 135$ and 179 degrees are chosen to be used in classifier design. Accordingly, a total of $K \times M = 12$ scattered time-domain signals are used in classifier design after being transformed into time domain by using IFFT. The classifier is designed over the optimum time interval [9.21 ns–10.89 ns] using the design parameters $N = 1024$, $T_0 = 26.81$ ns, $Q = 32$ and $q^* = 12$. An example of the scattered time-domain signals is shown in Figure 2 for the object T_3 at a bistatic aspect angle for $\theta = 150$ degrees. Real and imaginary parts of permittivity functions for all three objects are plotted in Figures 3(a) and 3(b), respectively.

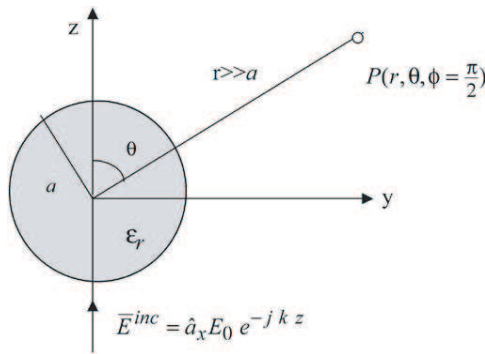
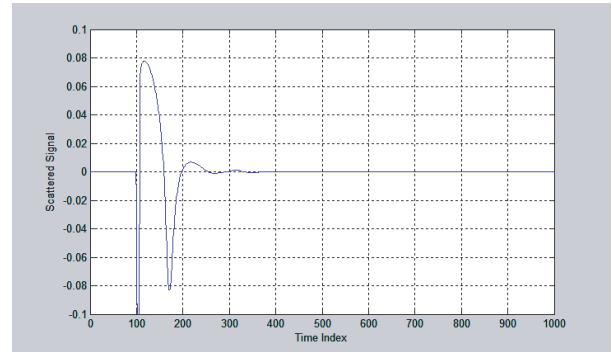
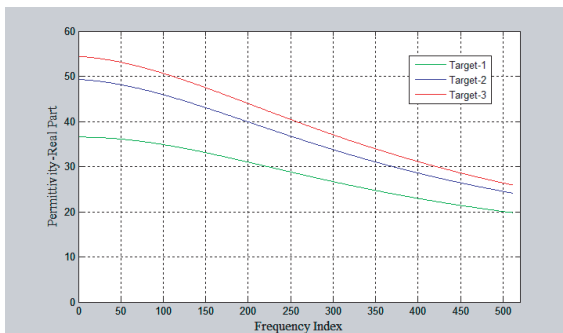
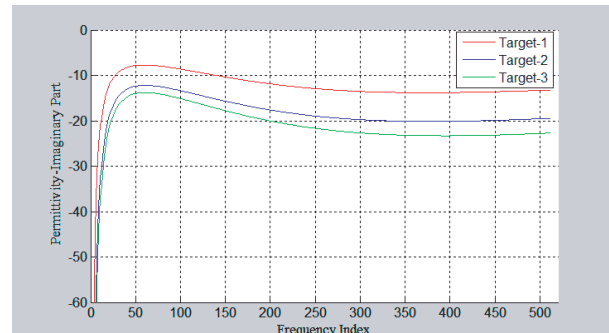


Figure 1: Simulation setup for scattered signal computations under plane wave excitation.

Figure 2: Time-domain scattered signal for target 3 at $\theta = 150$ degrees (for 30 degrees bistatic angle).

(a)



(b)

Figure 3: (a) Real parts $\varepsilon'_r(\omega)$ and (b) imaginary parts $\varepsilon''_r(\omega)$ of relative permittivity for targets T_1 , T_2 and T_3 .

The fused feature vectors (FFV's) obtained by the suggested WD-PCA technique for each library object are given in Figure 4. Finally, the test results for 24 different test signals (each of 3 targets observed at 8 different non-design aspects) are summarized in the contour plot of Figure 5 where the correlation coefficients between each of 24 testing LTFVs and each of 3 FFVs of the classifier feature database are computed with a result of 100 percent accuracy rate. The

diagonal “matched” submatrices have very large correlation coefficient values while the off-diagonal “mismatched” submatrices show much lower correlation figures, as expected.

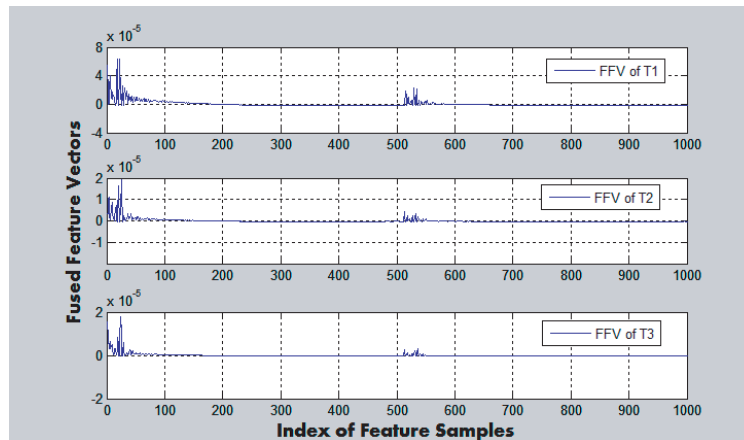


Figure 4: Fused feature vectors (FFVs) of the designed classifier for targets T_1 , T_2 and T_3 .

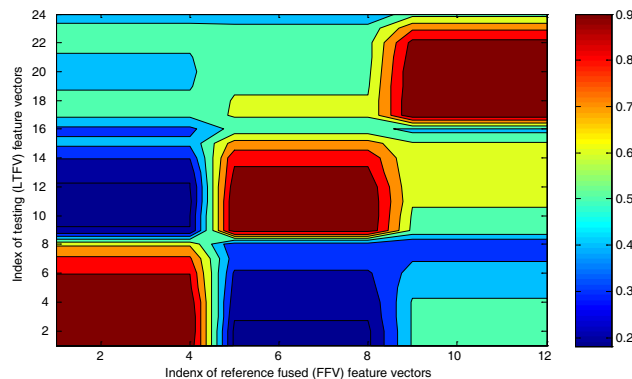


Figure 5: Contour plot of correlation coefficients computed for all possible pairs of testing LTFVs and FFVs.

4. CONCLUSIONS

In this paper, design of a WD-PCA based object classifier for three different lossy dispersive dielectric spheres is demonstrated together with performance testing with 100 percent accuracy. Noise performance of the classifier will be simulated in a future work. Also, usefulness of the WD-PCA method will be investigated in detection and classification of a benign or malignant tumor within such realistic breast tissues as the ultimate goal of our future studies.

REFERENCES

1. Baum, C. E., E. J. Rothwell, K. M. Chen, et al., “The singularity expansion method and its application to target identification,” *Proc. IEEE*, Vol. 79, No. 10, 1481–1492, Sept. 1991.
2. Turhan-Sayan, G., “Natural resonance-based feature extraction with reduced aspect sensitivity for electromagnetic target classification,” *Pattern Recognition*, Vol. 36, No. 7, 1449–1466, Jul. 2003.
3. Turhan-Sayan, G., “Real time electromagnetic target classification using a novel feature extraction technique with PCA-based fusion,” *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 2, February 2005.
4. Turhan-Sayan, G. and E. Ergin, “Non-destructive recognition of dielectric coated conducting objects by using WD type time-frequency transformation and PCA based fusion,” *International Journal of RF and Microwave Computer-Aided Engineering*, Vol. 23, No. 4, 403–409, Jul. 2013.
5. Zhu, G. K. and M. Popovic, “Comparison of radar and thermoacoustic technique in microwave breast imaging,” *Progress In Electromagnetics Research B*, Vol. 35, 1–14, 2011.

Bivariate Statistical Analysis for Electromagnetic Reverberation Chamber

Suna Choi and Seungkeun Park

Electronics and Telecommunications Research Institute (ETRI), Koera

Abstract— The probability density functions (PDFs) of electromagnetic fields inside real reverberation chambers has been categorized into eight cases and the mathematical expressions for all the cases were recently reported. In this paper, we compile the complete expressions for the magnitude and phase PDFs of bivariate normal distribution. Furthermore, we present the assessment of these PDFs using a RC located at Ajou University, South Korea. To investigate how close to the ideal case is the measured data, T-tests and F test are used to test the hypotheses for zero means, equal variances and zero correlation between the real and imaginary part of the field. The results show that the circumstance of the RC is far from the ideal case. Then, KS and AD GoF tests are applied to the magnitude and phase PDFs for investigating the agreement between the measurement data and PDFs of the eight cases. According to the GoF tests, the magnitude PDFs of the type 3 and type 7 are very highly rejected. Also, the rejection rate of magnitude PDFs of the type 4 is lower than those of most other types, and the type 8 shows the lowest rejection rate. On the other hand, the phase PDFs of even types are much less rejected than those of odd types, and the phase PDFs of the type 8 shows the lowest rejection rate. The results of the paper are helpful in better understanding the behavior of fields in RCs.

1. INTRODUCTION

Reverberation chambers (RCs) have been attracting attention as electromagnetic (EM) testing facilities, performance evaluations of multiple-input-multiple-output handset antennas, etc. The correct use of RCs is based on the full understanding of the field behaviors inside the RCs.

It is commonly accepted that the real and imaginary components of EM fields in RCs follow normal distributions. However, recent researches gave the PDFs of the field magnitude and phase in the viewpoint of “good-but-imperfect” approach on the basis of bivariate normal distributions (BNDs) in order to describe imperfect reverberation conditions [1, 2]. They categorized BNDs into eight cases, and suggested mathematical expressions describing the magnitude and phase of electric fields by deriving the marginal PDFs of the BNDs. Regarding the given approach to mathematically characterize EM signals, a similar but more generalized analysis was also proposed from a communication society [3]. Most recently, the analytical formulas which were not introduced in [1–3] is presented [4]. Thus, we compile the complete mathematical expressions describing the magnitude and phase of electric fields.

Moreover, we present the experimental results of electric fields in a real RC. Also, the idealness of means, variances and correlation between the real and imaginary part of the field inside the RC are investigated. Moreover, using the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) goodness-of-fit (GoF) tests, we evaluate the magnitude and phase distributions of electric fields for the eight cases of BNDs [5, 6].

2. MAGNITUDE AND PHASE PDFS OF BNDs

For characterizing the electric fields of a RC, the BND of real and imaginary parts ξ and η is represented as, $N(\mu_\xi, \mu_\eta; \sigma_\xi, \sigma_\eta; \rho)$ where μ_ξ and μ_η are the means for ξ and η respectively, σ_ξ and σ_η are the standard deviations and ρ is the correlation coefficient between ξ and η . The BNDs are categorized into eight types according to the values of $\mu_\xi, \mu_\eta, \sigma_\xi, \sigma_\eta$ and ρ .

The magnitude r and phase ϕ of the field are represented as follows:

$$r = \sqrt{\xi^2 + \eta^2}, \quad r \geq 0 \quad (1)$$

$$\psi = \frac{\pi}{2} \operatorname{sgn}(t) \cdot [1 - \operatorname{sgn}(s)] + \operatorname{sgn}(s) \cdot \tan^{-1} \left(\frac{t}{|s|} \right), \quad -\pi \leq \psi \leq \pi \quad (2)$$

where $\operatorname{sgn}(x) = x/|x|$ for $x \neq 0$, and $\operatorname{sgn}(x) = 0$ for $x = 0$.

According to the PDFs presented in [1, Table 1], six PDFs require numerical integrations when obtaining the MPDF of the field magnitude and phase because analytical expressions for these

PDFs are not found in closed forms. In [3], the magnitude and phase PDFs of BNDs for several cases were derived in analytical forms which involve the modified Bessel function of the first kind or the complementary error function. The phase PDFs for type VI and magnitude PDFs of Type VII are presented in [4]. In reviewing the PDFs reported in [1–4], the expressions for the magnitude and phase PDFs of BNDs are completed as shown in Table 1.

Table 1: summary of the PDFs for phase and magnitude.

Case	Phase PDF $f_\phi(\phi)$	Magnitude PDF $f_r(r)$
1. $N(0, 0; \sigma, \sigma; 0)$	[1, Eq. (8)]	[1, Eq. (8)]
2. $N(\mu_\xi, \mu_\eta; \sigma, \sigma; 0)$	[1, Eq. (10)]	[1, Eq. (11)]
3. $N(0, 0; \sigma_\xi, \sigma_\eta; 0)$	[1, Eq. (13)]	[1, Eq. (16)]
4. $N(\mu_\xi, \mu_\eta; \sigma_\xi, \sigma_\eta; 0)$	[4, Eq. (1)]	[1, Eq. (19)]
5. $N(0, 0; \sigma, \sigma; \rho)$	[1, Eq. (21)]	[2, Eq. (6)]
6. $N(\mu_\xi, \mu_\eta; \sigma, \sigma; \rho)$	[1, Eq. (24)]	[2, Eq. (3)]
7. $N(0, 0; \sigma_\xi, \sigma_\eta; \rho)$	[1, Eq. (26)]	[4, Eq. (2)]
8. $N(\mu_\xi, \mu_\eta; \sigma_\xi, \sigma_\eta; \rho)$	[2, Eq. (1)]	[2, Eq. (2)]

3. MEASUREMENT AND ANALYSIS

To verify the availability of the novel BN approach, complex electric field components were measured in the RC located at Ajou University, South Korea, while the RC was performing in stirred mode. The dimension of the RC is 2.4 m × 2.3 m × 1.6 m and the vertical Z-fold shape stirrer has size of 0.6 m × 0.5 m × 1.0 m. The chamber walls are made of stainless steel and welded. The first resonant frequency is around 97 MHz, and the lowest useable frequency might occur at slightly above 291 MHz, according to [7]. In this experiment, three different frequencies of 1.8, 2.4, and 3.8 GHz were selected, and about 5,000 samples per each frequency were measured using a vector network analyzer and dipole antennas.

To evaluate how close to the ideal case is the measured data, we test the ideal hypothesis for the parameters of means, variances and correlation between the real and imaginary part of the field measured in the RC. In ideal field condition of a RC, the means are zero, variances are equal and the correlation is zero. T tests and F test are used to test the ideal hypotheses and about 50 tests per each frequency are performed with 5% level of confidence. Fig. 1 shows the rejection rates resulting from the measurements performed in the RC. The rejection rates of zero mean hypotheses are 19, 28 and 22% for 1.8, 2.4 and 3.8 GHz, respectively, which is more highly rejected than other hypotheses. The rejection rates of equal variance hypotheses are 9, 12 and 15% and those of zero covariance hypotheses are 15, 14 and 13%. The rejection rates seem non-negligibly high. Therefore, it is considered that the circumstance of the RC is far from the ideal case.

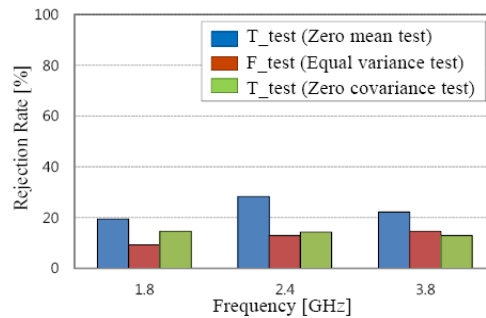


Figure 1: Ideal hypothesis Tests for means, variances and covariance.

For investigating the agreement between the measurement data and the eight cases, KS and AD GoF tests are applied to the PDFs for the magnitude and phase. KS GoF test makes an assessment of whether there is sufficient evidence to reject the null hypothesis that the measurement data and the theoretical PDF are the same. Fig. 2 presents rejection rates resulting from the KS GoF tests

with 5% of confidence level for the eight cases. The rejection rates of the magnitude PDFs of the type 4 are 4~11% and those of type 8 are 1~4%, which are less than those of other types. The magnitude PDFs of the type 3 and type 7 are very highly rejected, which have rejection rates of 35~73%, equally. The phase PDFs of even types are much less rejected than those of odd types. That is accord with the previous test result that the hypothesis of zero mean is highly rejected, as shown in Fig. 1. The rejection rate of phase PDFs of the type 8 is close to 0% while that of type 1 is 79~100%.

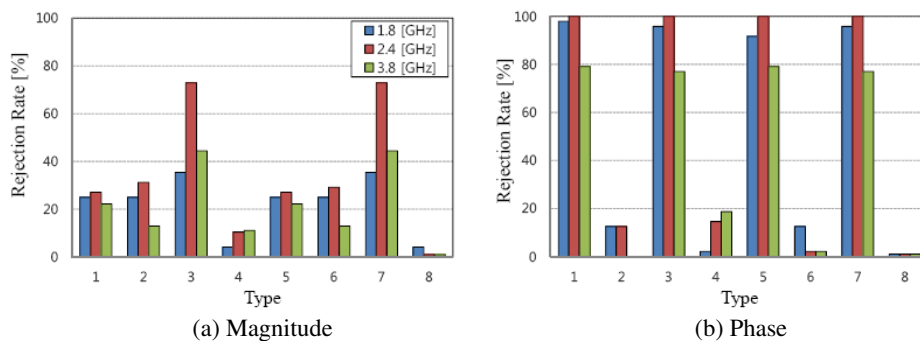


Figure 2: KS GoF test results for magnitudes and phases of eight cases.

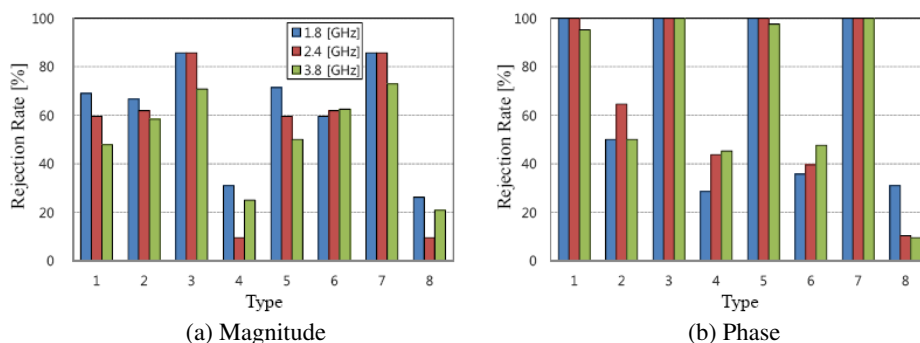


Figure 3: AD GoF test results for magnitudes and phases of eight cases.

The AD GoF test generally gives more stringent results than KS test. Accordingly, the results of AD GoF test shows higher rejection rate than KS GoF test. However, the tendency is similar to KS test as shown in Fig. 3. The rejection rates of magnitude PDFs of the type 8 are 9~26%, while those of type 1 are 48~69%. The rejection rates of phase PDFs of the type 8 are 9~31%, while those of type 1 are 95~100%.

The results show that the magnitude and phase of the field inside the RC are far from the ideal distributions. Therefore, some experimental errors should be taken into consideration when applying ideal assumptions. Other PDFs of complex cases are closer accordance with the experimental data.

4. CONCLUSION

We compile the complete expressions for the magnitude and phase PDFs of BNDs. Also, eight cases of BNDs are assessed using the experiment and the GoF tests. In order to guarantee the accuracy of the test results from RCs, the statistical distributions of EM fields inside the real RCs should be considered. The results of the paper are helpful in better understanding the behavior of fields in RCs.

ACKNOWLEDGMENT

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

REFERENCES

1. Serra, R. and F. G. Canavero, "Bivariate statistical approach for good-but-imperfect electromagnetic reverberation," *IEEE Trans. Electromagn. Compat.*, Vol. 53, No. 3, 554–561, Aug. 2011.
2. Serra, R., F. Leferink, and F. G. Canavero, "Good-but-imperfect electromagnetic reverberation in a VIRC," *Proc. IEEE Int. Symp. Electromagn. Compat.*, 954–959, Long Beach, USA, Aug. 2011.
3. Dharmawansa, P., N. Rajatheva, and C. Tellambura, "Envelope and phase distribution of two correlated Gaussian variables," *IEEE Trans. Commun.*, Vol. 57, No. 4, 915–921, Apr. 2009.
4. Choi, S. and S. Park, "A bivariate normalization approach for characterizing reverberation chambers," *IEEE Trans. Electromagn. Compat.*, accepted.
5. Eadie, W. T., D. Dryard, F. E. James, M. Roos, and B. Sadoulet, *Statistical Methods in Experimental Physics*, North-Holland, Amsterdam, 1971.
6. Anderson, T. W., *Anderson-Darling Tests of Goodness-of-fit*, International Encyclopedia of Statistical Science, Springer, 2011.
7. IEC 61000-4-21, *Electromagnetic Compatibility, Part 4: Testing and Measurement Technology, Section 21: Reverberation Chamber Test Methods*.

Design of the Electric-fields Probe in the Personal Exposure Meter

Sang Il Kwak¹, B. C. Kim¹, H. D. Choi¹, and Young Joong Yoon²

¹Radio Technology Research Department
Electronics and Telecommunications Research Institute (ETRI)
218 Gajeongno, Yuseong-gu, Daejeon, Korea

²Department of Electrical and Electronics Engineering, Yonsei University, Seoul, Korea

Abstract— This paper shows a development of the isotropic electric-fields probe in the personal exposure meter. The personal exposure meter is a device for measurement of E -fields near the human body. To detect the exposed E -fields, a 3-axis E -fields probe is used in the personal exposure meter. The probe is consists of three orthogonal infinitesimal dipole antennas with schottky diodes and lossy transmission lines using chip resistors. The proposed probe bandwidth is 10 MHz to 6 GHz and the isotropic deviation of the proposed probe radiation pattern is less than 0.5 dB in the simulation results. Finally, the fabricated probe will be shown and the properties of the probe such as a linearity and radiation pattern of the probe are presented.

1. INTRODUCTIONS

The electronic devices for communications, broadcasting and medical treatments are widely used in our life. For the human health and the mitigated EMC problems, the reduction techniques of electromagnetic field have been researched widely [1, 2]. To evaluation of exposed an E -field from the electronic devices, the E -fields probe is needed.

The personal exposure meter is a monitoring device for measuring on the total exposed E -field. The personal exposure meter consists of the E -fields probe with the lossy transmission line and a readout device. The E -fields probe comprises arranged perpendicularly three dipoles, a schottky diode and the resistive transmission lines. For precision measurement of the E -fields, the isotropic radiation pattern of the probe is required. To obtain an isotropic radiation pattern at a wideband frequency, the electrically small dipole antenna is used. The schottky diode transformed the detected a continues-wave incident field at the dipole antennas into a DC voltage. The transformed voltage is conveyed over the resistive transmission line to readout device. Since the transmission line is not received the electromagnetic fields directly or indirectly and act as a low pass filter, the highly lossy components such as chip resistors, NiCr and carbon materials are used. For simple fabrication, the chip resistors are used in this paper. The readout device is consists of an amplifier, signal process components and display components.

In this paper, the isotropic E -fields probe for the personal exposure meter is designed. In general, the isotropic deviation of a commercial 3-axis probe is less than 0.5 dB [3]. The proposed probe bandwidth is 10 MHz to 6 GHz and the isotropic deviation of the proposed probe radiation pattern is less than 0.5 dB in the simulation results. Finally, the fabricated probe will be shown and the properties of the probe such as a linearity and radiation pattern of the probe are performed.

2. PROBE DESIGN

Figure 1 shows the layout of the proposed E -fields probe. To measure the wideband services such as ISM devices (RFID, medical devices etc.), broadcasting and communications devices, the target frequency bandwidth is 10 MHz to 6 GHz. Using the Ref. [4], the size of the dipole antenna is calculated at target frequency. The size of the probe antenna is 3 mm \times 2 mm. A HSC4-5330 schottky diode is used for transformation and the spatial angle of the three dipoles is 54.74^o for orthogonal property. To obtain highly resistive transmission line, chip resistors are used.

3. RESULTS

To obtain the isotropic radiation characteristic of the probe, the parameter studies are performed at previous study [5]. The length of the transmission line is 65 mm and total height of the probe is 73.7 mm. Fig. 2 shows the simulated radiation pattern of the probe. As shown in Fig. 2 and Ref. [5], the isotropic deviation of the proposed probe is less than 0.5 dB.

The measurement of the probe is using the GTEM cell (TESEQ GTEM 250) in Fig. 3. The GTEM cell can create uniform E -fields and the input power can be adjusted. Fig. 4 shows the

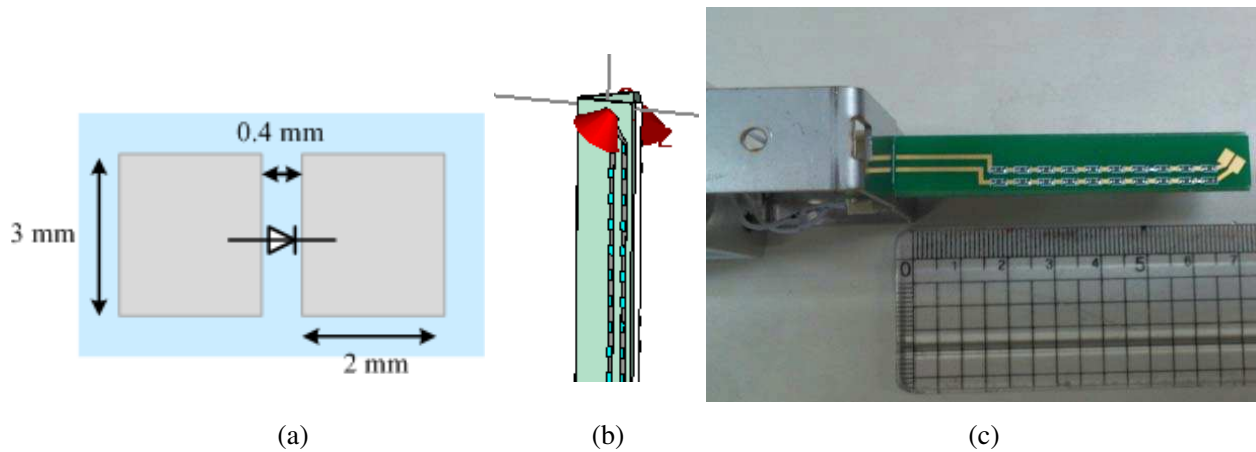


Figure 1: Layout of the proposed probe: (a) Dipole with a diode. (b) Probe (3-axis). (c) Picture of the probe with a readout device.

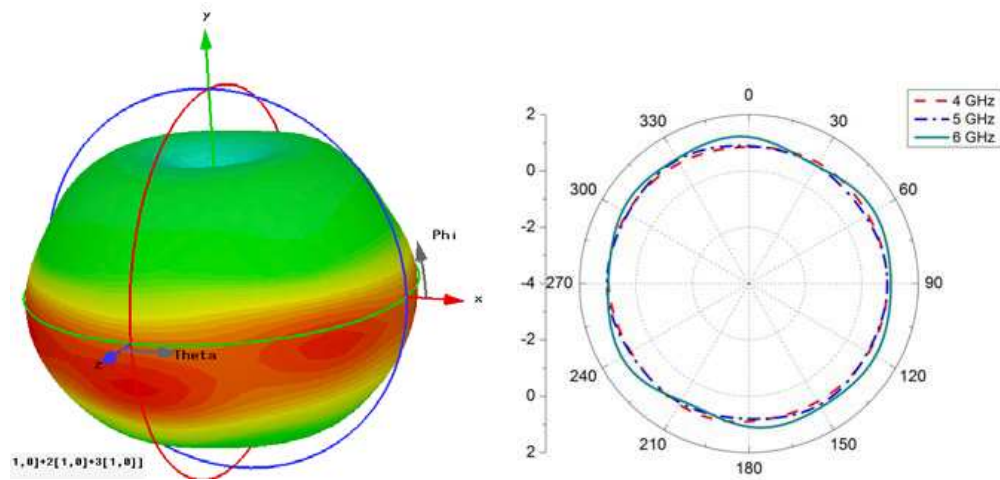


Figure 2: Simulated radiation pattern of the proposed probe.

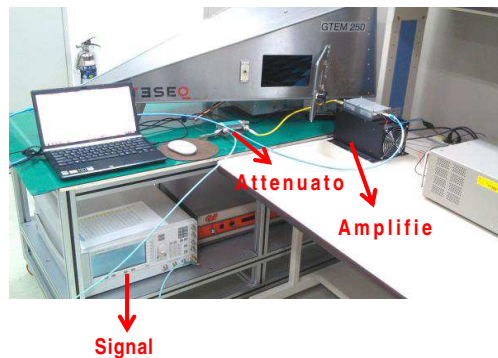


Figure 3: Measurement set-up.

measured output voltage of the probe according to the input power. As the input power is increased, the output voltage increases linearly. Also, the occupational exposure limits to protect human body at 2 GHz is 134 V/m. The proposed probe can measure the E -fields up to 267 V/m at 2 GHz. The measured radiation pattern of the probe is also measured using the GTEM cell. As the input power is fixed, the designed probe is rotated in the GTEM cell and measured output E -fields. The results will be shown at the conference.

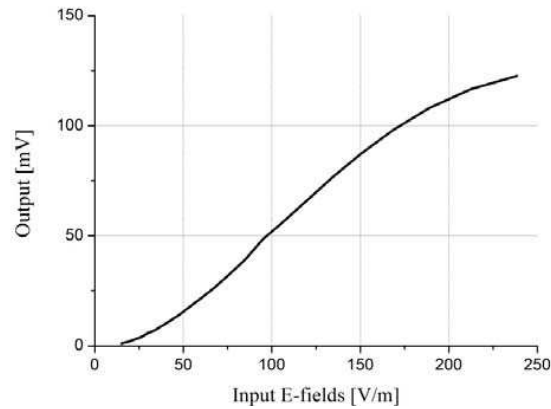


Figure 4: Measured output voltage according to the input E -fields at 2 GHz.

4. CONCLUSION

In this paper, the isotropic electric fields probe in the personal exposure meter is designed. To detect the exposed E -fields, a 3-axis E -fields probe is used in the personal exposure meter. The probe consists of three orthogonal infinitesimal dipole antennas with schottky diodes and lossy transmission lines using chip resistors. The bandwidth of the probe is 10 MHz to 6 GHz. The designed probe output voltage has a linear characteristic for the input power. Finally, the fabricated probe can be used in the personal exposure meter.

ACKNOWLEDGMENT

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2013.

REFERENCES

1. Kwon, J. H., S. I. Kwak, D. U. Sim, and J. G. Yook, "Partial EBG structure with DeCap for ultra-wideband suppression of simultaneous switching noise in a high-speed system," *ETRI Journal*, Vol. 32, No. 2, Apr. 2010.
2. Kwak, S. I., D. U. Sim, and J. H. Kwon, "Design of optimized multilayer PIFA with the EBG structure for SAR reduction in mobile applications," *IEEE Trans. Electromagnetic Compatibility*, Vol. 53, No. 2, May 2011.
3. <http://www.ets-lindgren.com/specs/HI-6005>.
4. Bassen, H. I. and G. S. Smith, "Electric field probes — A review," *IEEE Trans. Antennas and Propg.*, Vol. 31, 710–718, Sep. 1983.
5. Kwak, S. I., B. C. Kim, and H. D. Choi, "Analysis of isotropic deviation of the E -fields probe in the personal exposure measurement meter," *ICTC 2012*, 235–237, Oct. 2012.

Capacitively-loaded Inductively-coupled Fed Loop Antenna with an Omnidirectional Radiation Pattern for UHF RFID Tags

Qi Liu, Yufeng Yu, and Qiangsheng Huang

Center for Optical and Electromagnetic Research, Zhejiang University, Hangzhou 310027, China

Abstract— A capacitively-loaded inductively-coupled fed loop RFID tag antenna with an omnidirectional radiation pattern is proposed. With an inductively-coupled feed structure, it is convenient and accurate to achieve conjugate impedance matching condition. Concentration of surface current on the radiating loop is beneficial to improving the gain of antenna. The radiation pattern which is omnidirectional in antenna plane makes the proposed tag practicable mounted on top or bottom of target objects. The length of the antenna is only 60 mm, and the maximum simulated read range of the prototype is 14 m with 4 W of EIRP.

1. INTRODUCTION

The RFID (Radio Frequency Identification) technology is a non-contact automatic identification technology, which enables automatic target recognition and relevant data acquisition through radio frequency signals. In an RFID system, the tag antenna is a vital part and its performance greatly affects the reading range and accuracy of the RFID system.

Dipole-type antennas, whose radiation patterns are perpendicular to the surfaces of the antennas, are adopted for most tags applied in catering business [1]. However, if the target object is low-profile, or irregular-profile (e.g., plates), tags could no longer be attached on the profile of the target object. In this case, an RFID antenna with a maximum radiation direction in the same plane with the antenna plane would be suitable. Tags with such antennas can be mounted to the bottom of target objects, the labeling is easy-operating and the working performance is satisfactory in this case.

[2] and [3] proposed RFID antennas with such omnidirectional radiation patterns. They are both bent dipoles in shape and $60 \times 60 \text{ mm}^2$ in size. The reading range of the antenna proposed in [3] is 10–14 centimeters with output power of 14 dBm and reader antenna gain of 2 dBi, while the reading range of antenna proposed in [4] is 24–32 centimeters with output power of 27 dBm and reader antenna gain of 2 dBi.

In this paper, a UHF RFID tag antenna with an omnidirectional radiating pattern is proposed. This antenna consists of a radiating loop and a concentric inductively-coupled feeding loop. The radiating loop loaded with four periodically spaced interdigital capacitors realizes uniform current distribution along the loop even when it is no longer electrically small [4, 5]. The feeding loop inserted with an interdigital capacitor which is inductively coupled to the radiating loop provides a convenient way for conjugate impedance matching [6, 7]. In this way, the proposed tag antenna can achieve an omnidirectional radiation pattern and good conjugate impedance matching with arbitrary size. With a size of $0.18\lambda_0 \times 0.18\lambda_0$ (λ_0 is the corresponding wavelength of the operating frequency), the antenna achieves a simulated reading range of above 10 meters with total transmitted power of 4 W. This maximum reading range is much further than that in 3 and 4, which is 1.4 m and 0.7 m, respectively, with the same total transmitted power. The obtained operating impedance bandwidth ($< -3 \text{ dB}$) can cover the FCC band.

2. ANTENNA DESIGN

The geometry of the proposed antenna is depicted in Fig. 1. It is seen that the antenna consists of an outer and inner loop, which is the radiator and the feed, respectively. The two loops are made of 0.01 mm aluminum and printed on a Polyethylene Terephthalate (PET) substrate with a thickness of 0.05 mm and relative permittivity of 3.2. It is noted that the loops and substrate are represented by orange and gray colors in Fig. 1.

Four periodically spaced interdigital capacitors are inserted into the radiating loop to provide a small phase delay between the adjacent sections, which enables a uniform current distribution along the loop even when the loop is no longer electrically small. An interdigital capacitor is inserted into the feeding loop for impedance matching. An RFID chip Impinj Monza 4 is mounted on the feeding loop in the symmetrical position of the inner interdigital capacitor. The details of the interdigital capacitors are shown in the inset. Detailed dimensions of the proposed antenna are: $L_s = 45 \text{ mm}$,

$L_r = 43$ mm, $L_f = 23$ mm, $W = 3$ mm, $L_g = 0.3$ mm, $w_i = 0.3$ mm, $g = 0.24$ mm, $L_{out} = 7.9$ mm, $L_{inner} = 8.9$ mm.

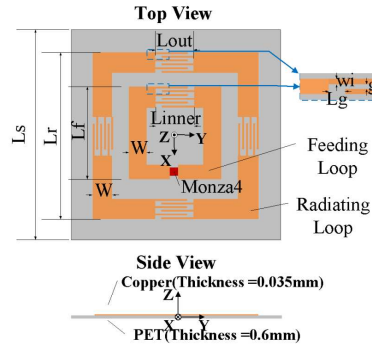


Figure 1: Geometry of the proposed RFID antenna.

3. SIMULATED RESULTS AND DISCUSSION

The antenna is designed at the frequency of 915 MHz, and at this frequency its size is $0.18\lambda_0 \times 0.18\lambda_0$. The simulated input impedance is shown in Fig. 2. It can be seen that the impedance of the proposed antenna is $11.7 + 146j$ Ohms at 915 MHz. Because the impedance of Monza 4 chip is $11 - 143j$ Ohms, good conjugate impedance matching condition can be achieved. The simulated PRC (Power Reflection Coefficient) [8] is plotted in Fig. 3. The lowest point of the simulated PRC is at the frequency of 915 MHz. The -3 dB bandwidth of the proposed antenna is 901–928 MHz, which covers the FCC band (902–928 MHz). The simulated radiation patterns in x - y and x - z planes are shown in Fig. 4. It is seen that the gain deviation is less than 1 dB in x - y plane, which indicates good omnidirectional radiation characteristic. The simulated reading range is plotted in Fig. 5. It

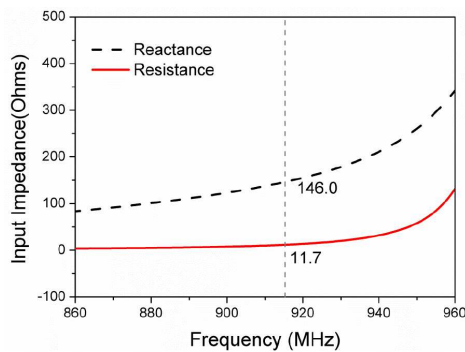


Figure 2: Simulated input impedance of the proposed antenna.

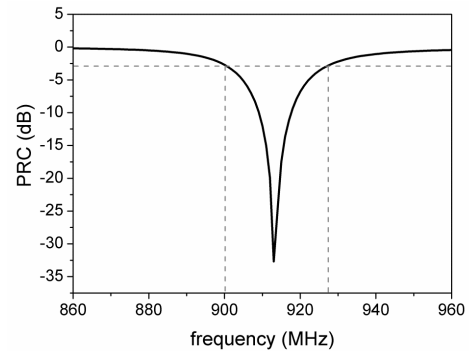


Figure 3: Simulated power reflection coefficient of the proposed antenna.

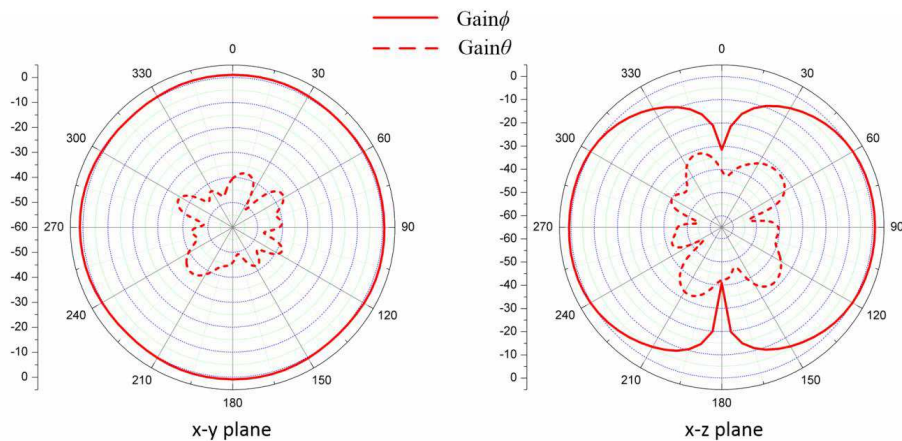


Figure 4: Simulated radiation patterns in x - y plane and x - z plane.

is seen that the reading range in FCC band is above 9 m and the maximum reading range is 14 m at 915 MHz.

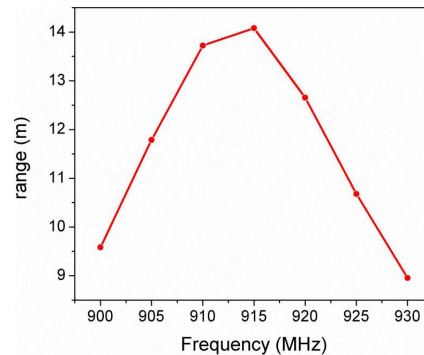


Figure 5: Simulated reading range from 900 MHz to 930 MHz.

4. CONCLUSIONS

In this paper, an RFID tag antenna has been proposed using capacitively-loaded inductively-coupled fed loop. The proposed antenna can obtain good conjugate impedance matching condition in arbitrary size. Compared to other tag antennas in similar size, it has a higher gain and further reading range. The radiation pattern which is omnidirectional in the antenna plane makes it a good choice for applications on target objects with irregular curved flanks in RFID applications.

ACKNOWLEDGMENT

This work is partially supported by the National High Technology Research and Development Program (863 Program) of China (No. 2012AA030402), the National Natural Science Foundation of China (Nos. 61178062 and 60990322), the Program of Zhejiang Leading Team of Science and Technology Innovation, Swedish VR grant (#621-2011-4620), and AOARD.

REFERENCES

1. Bjorninen, T., A. Z. Elsherbeni, and L. Ukkonen, "Low-profile conformal UHF RFID tag antenna for integration with water bottles," *IEEE Antennas and Wireless Propagation Letters*, Vol. 10, 1147–1150, 2011.
2. Kaiwen, K. T., X. Qing, C. K. Goh, and L. Zhu, "A UHF omni-directional RFID antenna," *Proc. IEEE Antennas and Propagation Society International Symposium (APSURSI)*, 1–4, Toronto, ON, July 2010.
3. Lukas, J., X. Qing, Z. N. Chen, and A. Alphones, "UHF omnidirectional bent dipole antenna for item-level RFID applications," *Proc. Int. IEEE AP-S Symp.*, 745–748, San Diego, CA, July 2008.
4. Qing, X. M., C. Goh, and Z. N. Chen, "A broadband near-field UHF RFID antenna," *IEEE Trans. Antennas Propag.*, Vol. 58, No. 12, 3829–3838, December 2010.
5. Wei, K., Z. J. Zhang, and Z. H. Feng, "Design of a wideband horizontally polarized omnidirectional printed loop antenna," *IEEE Antennas Wireless Propag. Lett.*, Vol. 11, 49–52, 2012.
6. Wu, J., J. X. Li, X. S. Cui, and L. H. Mao, "Circular loop antenna for UHF RFID tags with inductively coupled structure," *2011 International Conference on Control, Automation and Systems Engineering (CASE)*, 1–4, Singapore, July 30–31, 2011.
7. Son, H. W. and C. S. Pyo, "Design of RFID tag antennas using and inductively coupled feed," *Electron. Lett.*, Vol. 41, No. 18, 994–996, September 2005.
8. Nikitin, P. V., K. V. S. Rao, S. F. Lam, V. Pillai, R. Martinez, and H. Heinrich, "Power reflection coefficient analysis for complex impedances in RFID tag design," *IEEE Trans. Microw. Theory Tech.*, Vol. 53, No. 9, 2715–2721, September 2005.

Design of Coplanar Sensor for the Permittivity Measurement of Thin Dielectric Samples

Manisha Shete and M. Jaleel Akhtar

Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India

Abstract— This paper explores the nondestructive permittivity measurement of thin dielectric samples using the designed coplanar waveguide sensor. A conductor backed coplanar line is designed for 50 ohms, and the sample whose permittivity is to be measured is placed freely on top of this sensor. This structure of the coplanar sensor with sample placed on it is referred as MCCPW (multilayered conductor backed coplanar waveguide), and the scattering parameters of the MCCPW structure are measured using the VNA (Vector network analyzer). To reduce effect of higher order modes, the sample is placed at the center of the coplanar sensor. From measured scattering parameters, the effective permittivity of the MCCPW structure loaded with the test sample is calculated using the proposed reflection-transmission approach. The fill factors are calculated from the geometry of the MCCPW structure. In this paper, instead of using the costly microwave probe stations for the measurement of scattering parameters of coplanar lines, two SMA to coplanar end launchers are used. A simple de-embedding technique based on the transmission matrix method is used to extract the scattering parameters at the coplanar sample interface. Two coplanar sensors with length 25.4 mm and 50 mm are designed and fabricated. The designed sensor is tested by extracting the permittivity of a number of reference samples from both the simulated and the experimental data.

1. INTRODUCTION

There are a number of techniques presently available for measuring the complex permittivity and permeability of samples in the microwave frequency band. Each technique is limited to specific frequencies, materials, applications etc. by its own constraint. The most commonly used techniques are:

- Transmission/reflection line technique.
- Open ended coaxial probe technique.
- Free space technique.
- Resonant technique.

In this work, the most popular transmission/reflection technique is modified and used for the planar sensor. The planar circuit used here is based on the coplanar waveguide because of its advantages compared to other planar circuits such as microstrip and stripline [1]. The coplanar lines has advantage over microstrip as it is less dispersive than microstrip, and its impedance remains constant from 0.5 GHz to 40 GHz, which makes them quite useful for the broad band permittivity measurement [2–6]. In this paper, a grounded coplanar waveguide is designed for 50 Ω , and the air gap between side-grounds and central conductor is kept small in order to facilitate the even mode propagation. The proposed method requires no special sample preparation as the only requirement is to keep the air gap between sensor and sample to be minimum. Earlier proposed works usually suggest using test specimens of less than half wavelength in order to avoid the ambiguity [7–9]. However, it is observed that longer samples provide better accuracy in the extracted relative permittivity. To resolve the ambiguity for longer samples, algorithms presented in [10, 11] are used to calculate the actual solution. The proposed conductor backed coplanar sensor employs a set of easily available SMA to coplanar launchers instead of the expensive probe station commonly being used for these types of measurements. The characterization of these two coplanar launchers is done using the in-house developed algorithm based on the transmission matrix based approach.

In this work, two coplanar sensors of different lengths are designed in order to characterize longer dielectric samples over a wide frequency band. The designed coplanar sensor is validated by simulating the structure with the help of full wave numerical simulation software, the CST microwave studio [12]. The simulated scattering coefficients of a number of standard samples are employed in the proposed algorithm to estimate their dielectric properties in order to compare them with their actual values. After validating the design against the simulation data, the designed coplanar sensors are fabricated in order to measure the dielectric properties of a number of substrates and other thin dielectric samples with the help of a Vector Network Analyzer (VNA).

2. THEORY

2.1. Extraction of Sample Permittivity from Scattering Parameters

The first step of the proposed approach requires the design of the conductor backed coplanar waveguide sensor having constant impedance of 50Ω in the frequency range of operation using analytical formulas [9]. It is desired to have the even mode field propagation as it is less dispersive. The coplanar waveguide is simulated using the full wave 3-D electromagnetic field simulator, the CST Microwave Studio [12] to validate the initial design of the sensor. The test specimen is placed on the designed sensor for the measurement of scattering coefficients using a VNA, and the effective permittivity of the test specimen placed is determined from the measured scattering coefficients using modified version of the analytical formulas [9]. The actual complex permittivity of the sample is extracted from the effective permittivity using the formulas of the fill factors of each region, i.e., the coplanar line, the sample, and the air surrounding it [8].

Now, the problem in using directly the formulas given in [9] for computation of the effective permittivity is that it does not have unique solution because of the presence of logarithmic function of the complex number. This can be understood by re-writing the logarithmic function of propagation constant P in terms of the absolute value and the phase term as follows:

$$\log\left(\frac{1}{P}\right) = -\log(|P|) - i(\phi + 2n\pi) \equiv \gamma l \equiv (\alpha + j\beta)l \quad (1)$$

The above equation can be separated to obtain:

$$\alpha = \frac{\log|P|}{l} \quad (2)$$

$$\beta \equiv \frac{2\pi}{\lambda_g} = \frac{\phi + 2\pi n}{l} \quad (3)$$

where $|P|$ and ϕ are the magnitude and phase of P , respectively, l is the length of the sample, and n is an integer. It can be easily seen here that although α has a unique solution, but there is ambiguity in the solution of β because of the number of roots associated with (3). This problem was taken into account in earlier works [3, 6] by keeping the sample length less than the wavelength in order to make n always equal to zero. However, it is observed that the longer samples provide accurate results as compared to the shorter samples, especially for low loss cases.

In view of the above, it is useful to compute the correct value of n for the determination of permittivity of longer samples for accurate measurements. In this work, the method proposed earlier such as in [10, 11], has been adapted to make it applicable for the coplanar sensor so that the test sample of length greater than a wavelength can be measured using the coplanar line. Once the correct value of n is determined, the earlier proposed formulation [9, 11] can be used to calculate the effective permittivity in terms of the measured scattering data.

3. NUMERICAL SIMULATION

The coplanar sensor is first simulated using the electromagnetic simulator, the CST Studio in order to optimize the design for 50Ω feed impedance in the frequency range of interest. The sample is kept at the center of the coplanar sensor so that high order modes generated near the interface are attenuated. Figure 1 shows the coplanar waveguide with thin dielectric sample simply placed at its center. The reference planes are at the excitation ports of the coplanar waveguide, and the scattering parameters are defined at the reference planes. The scattering coefficients of a number of known dielectric samples placed on the coplanar sensor are simulated using the CST studio, and relative permittivities of these samples are extracted using the proposed approach. In all cases, the extracted values of the permittivity match quite well with their actual values. Figure 2 shows the reconstructed permittivity of 10 mm Alumina sample placed on the designed sensor. It can be observed from this figure that when the actual value of n is not resolved, then the permittivity becomes totally unstable at higher frequencies. On the contrary, the proposed approach which resolves the ambiguity for larger samples by finding the the correct value of n , provides the stable value of permittivity of the test specimen over the whole frequency range as observed from this figure. It is also observed from this figure that there is no ambiguity in the reconstruction of permittivity at lower frequency values.

This is because of the fact that at lower frequencies, the wavelength is large enough so that the electrical length of the sample does not exceed this value thereby providing stable reconstructions.

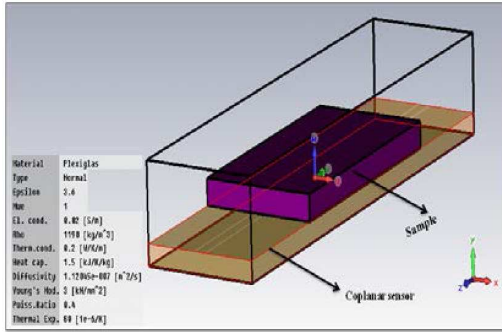


Figure 1: The test sample placed on the coplanar sensor.

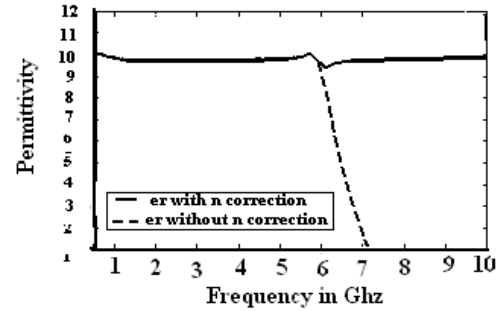


Figure 2: Real permittivity of alumina $l = 10$ mm placed on the coplanar line with and without correction.

4. EXPERIMENTAL RESULTS

For facilitating the dielectric measurement of samples, two coplanar sensors of different lengths are fabricated using the FR4 substrate. The metallization thickness is 18 microns, and the air gap width of 200 micron is maintained to comply with the fabrication facility of the institute. The test specimen is placed on the designed sensor, and the two ends of the sensor are connected to two ports of the Vector Network Analyzer through two SMA to coplanar end launchers. The standard two port calibration short open, load and through (SOLT) is carried out at the SMA end of the coplanar line over the frequency range of 0.5–6 GHz in order to facilitate the accurate measurement of the scattering coefficients in the specified frequency band.

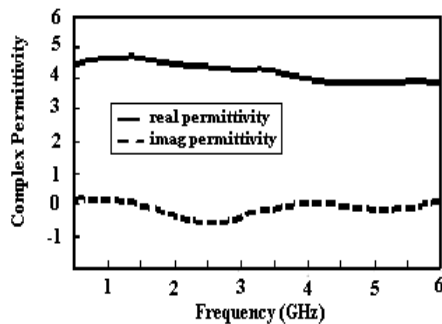


Figure 3: Measured permittivity of the FR4 sample.

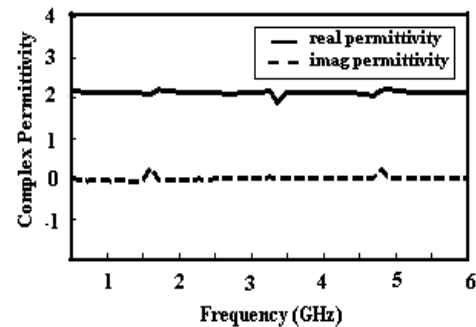


Figure 4: Measured permittivity of the Teflon sample.

In Figure 3, both the real and imaginary parts of permittivity of FR4 sample are extracted by placing it on the coplanar sensor of length 25.4 mm. Figure 4 represents the extracted complex permittivity of Teflon sample by placing it on the coplanar sensor of 50 mm length. For the longer Teflon sample shown in Figure 4, the spikes can be observed at few points, which are due to correction of n applied at these points. If the value of n is not corrected at these points, then the values of reconstructed permittivity would be totally ambiguous.

5. CONCLUSION

A practical form of the coplanar waveguide sensor for the permittivity measurement of thin dielectric samples has been designed, fabricated and tested. The proposed sensor employs simple de-embedding technique in combination with the cheaper end launchers in lieu of the expensive microwave probe stations for the measurement of scattering coefficients. The ambiguity associated with the longer samples has been resolved, and hence it facilitates the measurement of permittivity of test samples of practical lengths, which is especially useful for higher permittivity specimens. The extracted permittivities of a number of standard samples have been measured using the proposed approach, and their values match very closely with the data available in literature.

REFERENCES

1. Barry, W., "A broad band, automated, stripline technique for simultaneous measurement of complex permittivity and permeability," *IEEE Trans. Microwav. Theory Tech.*, Vol. 34, No. 1, 80–84, Jan. 1986.
2. Raj, A., W. S. Holmes, and S. R. Judah, "Wide bandwidth measurement of complex permittivity of liquids using coplanar line," *IEEE Transactions on Instrumentation and Measurement*, Vol. 50, No. 4, 802–809, Aug. 2001.
3. Hu, J., A. Sligar, C.-H. Chang, S.-L. Lu, and R. K. Settaluri, "A grounded coplanar waveguide technique for microwave measurement of complex permittivity and permeability," *IEEE Transactions on Magnetics*, Vol. 42, No. 7, 1929–1931, Jul. 2006.
4. Booth, J. C., N. D. Orloff, J. Mateu, M. Janezic, M. Rinehart, and J. A. Beall, "Quantative permittivity measurements of nanoliter liquid volumes in microfluidic channels to 40 GHz," *IEEE Transactions on Instrumentation and Measurement*, Vol. 59, No. 12, 3279–3288, Dec. 2010.
5. Hinjosa, J., "S-parameter broadband measurements on-coplanar and fast extraction of the substrate intrinsic properties," *IEEE Microwave and Wireless Components Letters*, Vol. 11, No. 2, 80–82, Feb. 2001.
6. Kang, B., J. Cho, and C. Cheon, "Nondestructive measurement of complex permittivity and permeability using multilayered coplanar waveguide structures," *IEEE Microwave and Wireless Components Letters*, Vol. 15, 381–382, May 2005.
7. Bediar, S. S. and I. Wolff, "Fast, accurate and simple approximate analytic formulae for calculating the parameters of supported coplanar waveguides for MMICs," *IEEE Trans. Microwav. Theory Tech.*, Vol. 40, No. 1, 41–48, Jan. 1992.
8. Simons, R. N., *Coplanar Waveguide Circuits, Components and Systems*, John Wiley & Sons, 2004.
9. Weir, W. B., "Automatic measurement of complex dielectric constant and permeability at microwave frequencies," *Proc. IEEE*, Vol. 62, No. 1, 33–36, Jan. 1974.
10. Akhtar, M. J., L. E. Feher, and M. Thumm, "A closed-form solution for reconstruction of permittivity of dielectric slabs placed at the center of a rectangular waveguide," *IEEE Geosciences and Remote Sensing Letters*, Vol. 4, No. 1, 122–126, Jan. 2007.
11. Zhou, H., G. Lu, Y. Li, S. Wang, and Y. Wang, "An improved method of determining permittivity and permeability by S parameters," *PIERS Proceedings*, 768–773, Beijing, China, Mar. 23–27, 2009.
12. Computer Simulation Technology (CST), Darmstadt, Germany, 1998–2003, [online] Available: <http://www.cst.com>.

Inverse Modeling of Pseudo-interdigital Bandpass Filters Using Artificial Neural Networks

Erdem Demircioglu¹, Murat H. Sazli¹, Orhan Sengul²,
S. Taha Imeci³, and Mesut Gokten⁴

¹Department of Electrics and Electronics Engineering, Ankara University, Tandoğan, Ankara, Turkey

²R & D Department, T. C. Başbakanlık, Ankara, Turkey

³Department of Electrics and Electronics Engineering, İstanbul Commerce University
Küçükyalı, İstanbul, Turkey

⁴Department of R & D and Satellite Manufacturing, Turksat AS, Golbasi, Ankara, Turkey

Abstract— A neural network trained to model original EM problems can be called as the forward model where the model inputs are physical or geometrical parameters and outputs are electrical parameters. Conversely neural network techniques are applicable to inverse modeling of microwave circuit design. In opposition to conventional statistical electromagnetic signal processing applications, inverse modeling techniques acquire electrical parameters as model input and geometrical properties as the output. Pseudo-interdigital (PID) bandpass microstrip filters offer compact and planar solutions to wide bandwidth filtering applications. They avoid the through vias required for short circuiting in conventional interdigital filters. Miniaturized microstrip bandpass filters are in demand for systems requiring small size and light weight. The coupling of the resonators in filter design must be adjusted using EM simulators. There are no analytical or numerical methods proposed for accurate determination of resonator spacing. In this study, the inverse modeling is applied to accurately determine the resonators' locations consistent with desired filter specifications.

1. INTRODUCTION

In recent decades, knowledge-aided design (KAD), approaches have been developed on the RF microwave modeling field for design and optimization. An artificial neural network (ANN) model for a device or circuit can be established by learning from microwave data which is acquired by measurement and simulation results, through a process called training. Once the ANN is trained, it can be applied for microwave design to provide instant answers to tasks it learned [1]. Successful implementation of nonlinear devices turns ANN applications into a research area for the modeling of passive and active microwave components [1, 4], adaptive beam forming [5, 6], cavities [7], stripline synthesis [8] and hybrid couplers [9].

A neural network trained to model the input-output relation between physical or geometrical parameters and electrical parameters respectively, is called as the forward model. The inverse model utilizes the adverse relation with electrical and physical parameters. The optimization and direct inverse models are proposed to solve the inverse problem [10]. The optimization method includes the repetitive EM simulations to find the optimum geometry. The direct inverse model establishes the electrical parameters as network input neurons and physical or geometrical parameters in the neural model desired output. Parallel-coupled microstrip filters with half-wavelength resonators are common elements in many microwave systems. However the size would be incompatible with the systems, where the reduced element size is a vital design parameter. Conventional microstrip inter-digital filters offer compact designs with short-circuited via holes. The fabrication of via holes requires non planar methods. The planar fabrication and compact design requirements offer pseudo-interdigital filter structures [11].

In this study, microstrip PID bandpass filter geometry is modeled using ANN to find the optimum gap lengths. In the common design procedure of pseudo-inter digital filters, the gap lengths between resonators are tuned in a trial and error method to calculate the coupling between resonators. In this paper, the inverse modeling of PID bandpass filters is applied and gap lengths are determined using design requirements such as coupling and substrate characteristics.

2. NEURAL NETWORK MODELING OF PID BANDPASS FILTERS

Multilayer perceptron (MLP) neural networks embrace the input layer, one or more hidden layers including computation nodes and an output layer with calculation ability. The input signal propagates through the network in a forward direction, on a layer-by-layer basis. MLP has been applied

successfully to solve nonlinear problems by training them in a supervised manner with a highly popular algorithm known as the error back-propagation algorithm [12].

The PID bandpass filters are proposed to achieve compact designs compared to conventional hair pin filters. The focused PID component layout is given in Figure 1. The gaps, S_1 and S_2 , between resonators are determined as sweeping parameters to alter the coupling coefficient. The $S_1 + S_2$ is fixed to 1 mm, and the S_1 and S_2 varies between 0.2 mm to 0.8 mm with 0.1 mm steps [11]. The dielectric constant of the filter substrate varies in a range of 2.2 to 10.8 and the substrate thickness alters from 20 mils to 125 mils. As the S_1 and S_2 changes, the electrical and magnetic resonance frequencies of the filters vary and the stop bandwidth is fine tuned. There is no analytical method is proposed to adjust the gaps between resonators. In this study, a KAD-based ANN model is proposed to determine the gap widths correspond to the desired coupling coefficients.

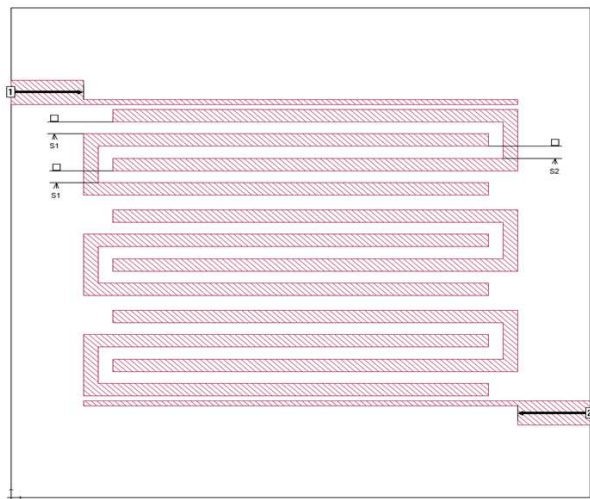


Figure 1: The pseudo-interdigital bandpass filter layout.

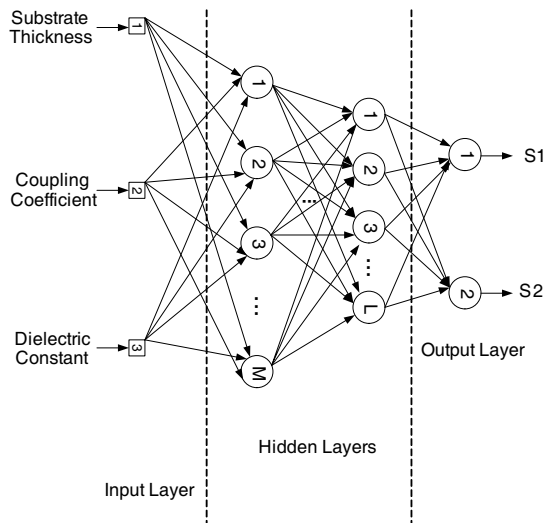


Figure 2: The proposed ANN architecture for the inverse modeling of PID filter.

In opposition to forward modeling methods, inverse modeling utilizes microwave design specifications such as coupling coefficient, insertion, return losses and bandwidth etc. through the input neurons and determine the component's geometrical dimensions. The proposed ANN architecture for the inverse modeling of PID filter is shown in Figure 2.

3. SIMULATION RESULTS

The inverse ANN modeling is applied accurately to determine the resonator spacing. The dielectric constant, substrate thickness and coupling coefficients are employed as model inputs and the resonator gap widths, S_1 and S_2 are established as outputs. The estimated spacing values are pre-

Table 1: The ANN model predictions for various substrates.

ϵ_r (Dielectric Constant)	Substrate Thickness (mils)	Coupling Coefficient	S_1 (mm)	S_1 (ANN Output)	S_2 (mm)	S_2 (ANN Output)
10.8	50	0.168	0.5	0.5067	0.5	0.4902
10.2	50	0.235	0.4	0.3739	0.6	0.6296
6.15	75	0.066	0.7	0.7580	0.3	0.2755
2.94	120	0.231	0.5	0.4978	0.5	0.5156
2.94	30	0.257	0.3	0.3066	0.7	0.7246
2.2	125	0.0957	0.7	0.7351	0.3	0.2847
10.2	75	0.134	0.6	0.6147	0.4	0.3859
2.2	31	0.264	0.3	0.2578	0.7	0.7347
6.15	100	0.170	0.5	0.4970	0.5	0.5101

sented in Table 1. It can be concluded that the dielectric constant has higher impact than substrate thickness for determining coupling coefficients and resonator spacing.

The proposed neural model includes a single hidden layer and five neurons in that layer. As the neuron number in each layer increases or the number of hidden layers is amplified, ANN model becomes over trained and the model accuracy decreases. The overall test set mean square error (MSE) for various neuron numbers in the hidden layer is given in Table 2.

Table 2: The overall mean square error values for various neuron numbers in the hidden layer.

Neuron numbers in hidden layer	Mean square error
3	0.0033
5	0.0359
10	0.2956
20	0.4665
50	0.3157
100	0.3640

4. CONCLUSIONS

In this work, an inverse ANN model is proposed to determine adequate PID bandpass filter's resonator spacing for achieving desired coupling coefficients. PID bandpass filters provide compact design methodology, however neither an analytical nor a numerical method is offered in the literature to determine resonator spacing. EM-simulators with trial and error schemes must be employed. ANN outputs provide low MSE and sufficient results which can be improved using properly selected hidden layer numbers and neurons in these layers. Complex architectures with many hidden layers and neurons have vital drawbacks of high computation time and over trained models. These trade-offs must be evaluated when the neural networks are realized to solve microwave modeling problems.

REFERENCES

1. Wang, F. and Q. J. Zhang, "Knowledge based neural models for microwave design," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 12, 2333–2343, 1997.
2. Zaabab, A. H., Q. J. Zhang, and M. S. Nakhla, "Neural network modeling approach to circuit optimization and statistical design," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 43, No. 6, 1349–1358, 1995.
3. Watson, P. M. and K. C. Gupta, "Design and optimization of CPW circuits using EM-ANN models for CPW components," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 45, No. 12, 2515–2523, 1997.
4. Creech, G. L., B. J. Paul, C. D. Lesniak, T. J. Jenkins, and M. C. Calcaterra, "Artificial neural networks for accurate microwave CAD application," *IEEE International Microwave Symposium Digest*, 733–736, San Francisco, CA, 1996.
5. Zaharis, Z. D., K. A. Gotsis, and J. N. Sahalos, "Adaptive beamforming with low side lobe level using neural networks trained by mutated boolean PSO," *Progress In Electromagnetics Research*, Vol. 127, 139–154, 2012.
6. Zaharis, Z. D., K. A. Gotsis, and J. N. Sahalos, "Comparative study of neural network training applied to adaptive beamforming of antenna arrays," *Progress In Electromagnetics Research*, Vol. 126, 269–283, 2012.
7. Stanković, Z., B. Milovanovic, and N. Doncov, "Hybrid empirical-neural model of loaded microwave cylindrical cavity," *Progress In Electromagnetics Research*, Vol. 83, 257–277, 2008.
8. Yildiz, C., K. Guney, M. Turkmen, and S. Kaya, "Neural models for coplanar strip line synthesis," *Progress In Electromagnetics Research*, Vol. 69, 127–144, 2007.
9. Demircioglu, E. and M. H. Sazli, "Behavioral modeling of a C-band ring hybrid coupler using artificial neural networks," *Radioengineering*, Vol. 19, No. 4, 645–652, 2010.
10. Kabir, H., Y. Wang, M. Yu, and Q. J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 56, No. 4, 867–879, 2008.

11. Hong, J. S. and M. J. Lancaster, "Development of new microstrip pseudo-interdigital bandpass filters," *IEEE Microwave and Guided Wave Letters*, Vol. 5, No. 8, 261–263, 1995.
12. Haykin, S., *Neural Networks — A Comprehensive Foundation*, Prentice Hall, 1999.

Metasurfing Substrate Integrated Waveguides to Mold the Radiation from Leaky Waves

Jose Luis Gomez-Tornero¹, George Goussetis², and Jay Guo³

¹Technical University of Cartagena, Spain

²Heriot-Watt University, UK

³CSIRO ICT Centre, Australia

Abstract— We show our latest developments in the application of the meta surfing concept to Substrate Integrated Waveguide (SIW) technology. It is described how this simple and versatile guiding medium can be conveniently modified and modulated to control the conversion from the conventional TE₁₀ propagating mode to surface or leaky wave with tailored propagation and radiation features. Theoretical aspects and practical designs are described with the general objective of wavefront control (both in amplitude and phase), showing results at Ku band (12–18 GHz) and W band (75–111 GHz), and illustrating a wide range of physical phenomena and interesting potential applications.

1. INTRODUCTION

The control of wavefronts of propagating/radiating electromagnetic fields is a general aspect of the electrodynamics theory & engineering, with many potential applications [1]. Particularly, the “metasurfing” concept has been coined referring to the engineering of artificial thin materials (also called metasurfaces) which can tailor the EM propagation [2]. In this work, we describe how these concepts can be applied to the well-known Substrate Integrated Waveguide (SIW) [3], so that its TE₁₀ propagating mode can be converted to a surface or leaky wave with tailored propagation and radiation features in the microwave/millimeterwave band [4]. These techniques can also be related to microwave holography in which surface-to-leaky wave transformation plays a fundamental role [5–8].

Figure 1(a) shows a top-view scheme of a SIW which cross-section dimensions are modulated along its length (z -axis). Particularly, two independent geometrical functions are defined: the width $W(z)$ and the separation between conducting vias in one of its sides $P(z)$. If $P(z)$ is large enough, the guided energy of the TE₁₀ propagating mode couples to the parallel-plate region surrounding the SIW side in the form of a vertically polarized TE surface wave (as shown in Fig. 1(b)). This makes the coupled SIW TE surface mode to have a complex wavenumber as a result of the attenuation due to substrate leakage. In addition, this surface wave can be transformed into a horizontally polarized leaky wave radiating into the upper free space, provided the top metal is truncated as shown in Fig. 1(b).

Figures 1(c) and 1(d) show, respectively, practical realizations of this SIW antenna in Ku band (12–18 GHz) [5, 9] and W band (75–111 GHz) [10] showing the feeding mechanism on the right side (SMA with microstrip-to-SIW at Ku band, and CPW to SIW at W-band). Full-wave simulations showing the propagation of the fundamental TE₁₀-like SIW mode (which attenuates as it propagates due to radiation losses) and the generation of a scanned fan-beam in the far-field, are depicted in Figs. 1(e) and 1(f) respectively. The following section describes different examples of radiated wavefront control/shaping. The efficient and robust synthesis procedure will be illustrated during the talk and it is not described here due to space restrictions; it is based on closed-form equations based on a simple but accurate Transverse Equivalent Network (TEN) of the modulated SIW cross section [5, 11].

2. WAVEFRONT CONTROL AND RADIATION PATTERN SYNTHESIS

The synthesis is based on the assumption of a modulated leaky wave (LW) described by its normalized complex propagation constant which varies along the length of the SIW [4, 5, 12]:

$$k(z)/k_0 = \beta(z)/k_0 - j\alpha(z)/k_0 \approx \sin \theta_{RAD}(z) - j\alpha(z)/k_0 \quad (1)$$

The first type of design is the most typical in LW antennas [12], since it involves the modulation or tapering of the LW to reduce the sidelobe level for a given scanning angle. As shown in Fig. 2(a), this is synthesized by keeping a coherent wave with uniform pointing angle θ_{RAD} along the whole

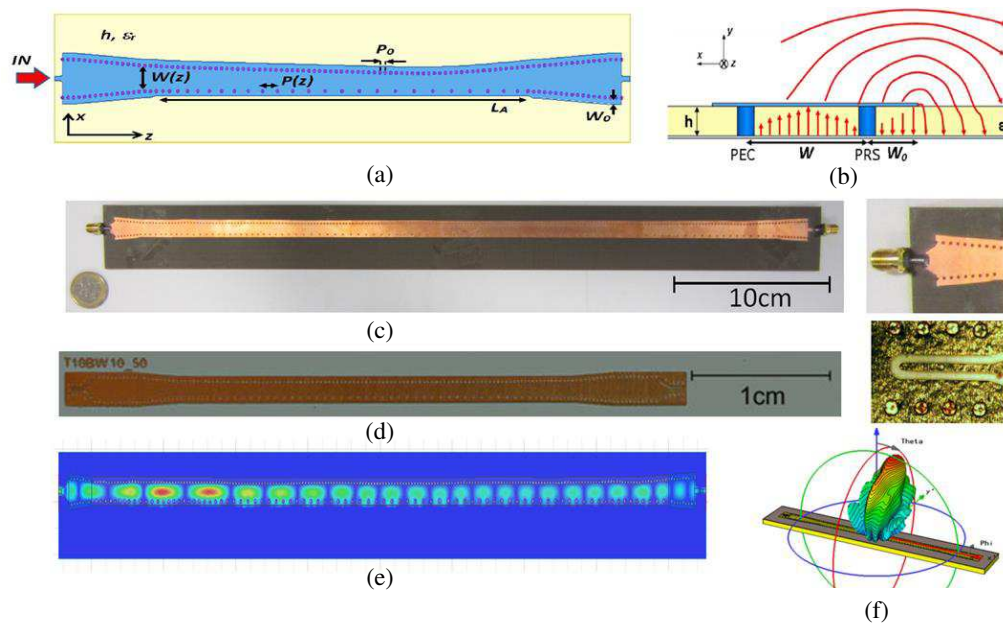


Figure 1: (a) Scheme of SIW LWA. (b) Cross section and electric fields. (c) Manufactured example for 15 GHz (d) and for 90 GHz ((coaxial to SIW transition and CPW to SIW transition shown on the right). (e) HFSS simulation of modulated propagating leaky mode (observing the modulation of the guided wavelength). (f) Simulated radiated fan beam.

antenna length (LW phase constant is kept unchanged), while its intensity is varied in a sine-like shape (LW leakage constant is adequately controlled). This type of designs finds application for directive telecom links and/or FMCW radar.

The second example involves the shaping of the far-field main beam, so that a selective angular response is obtained as shown in Fig. 2(b). For this purpose the scanning angle is not kept constant but it is modulated following an increasing function of the pointing angle as the LW advances through the antenna [4, 5]. This creates a diverging beam, which exhibits quasi-uniform low-ripple radiated power density per unit of angle in a certain angular region, and high rejection out from it, as shown in Fig. 2(b). The far-field filtering response can also be enhanced by inserting radiation null in adjacent angular regions [13] as illustrated in Fig. 2(c) [5]. This beam shaping can be used in telecoms/FMCW radar, or analog signal processing.

The LW radiated fields can be also modulated with a decreasing function of the local scanning angle, so that a converging phase front which focuses EM energy in the near-field regime [4, 5, 7, 8] is created as shown in Fig. 2(d). This has been demonstrated in SIW and microstrip technologies [14], thus envisaging its application for low-cost, high-performance microwave focusing systems for heating/sensing and/or hyperthermia medical applications [7, 8]. Also, this near-field concept has been recently applied by the authors to conceive a new type of quasi-optical multiplexer in SIW technology [15]. It must be highlighted that, in this latter case, the near-field focusing pattern is synthesized inside the substrate that host the SIW to provide a single-layer low-profile compact SIW device. For this purpose, the SIW couples to a surface wave with controlled transformation and coupling for optimum advanced design. Also it must be highlighted that this device benefits from the inherent frequency-scanning properties of LWs [16] to obtain spectral-spatial decomposition and thus the multiplexing behavior in a simple manner.

Another application of LW wavefront modulation techniques is the synthesis of electrically-long conformal radiators, to obtain high-directive scanning response despite the curved geometry [17]. This has also been demonstrated in SIW LWA technology, as shown in Fig. 2(e) [18]. In this application, the radiated wavefront must be readjusted to correct for the curved shape of the SIW, so that a resulting plane wave emanates from the conformed and modulated antenna.

A radial array configuration of SIWs has also been proposed [19] to extend the far-field shaping techniques from 1D to 2D, as shown in Fig. 2(f). In this way, a highly directive pencil beam radiating at broadside can be easily synthesized and simply fed from a single feeding, showing high interest for many applications. Also, this radial array configuration can be used for 3D shaping of the near-field focus, as proposed in [20] and illustrated in Fig. 2(g). Again, this interesting and

original device might be useful for microwave heating/sensing or microwave power transmission, and therefore suitable to associated industrial, scientific and medical applications.

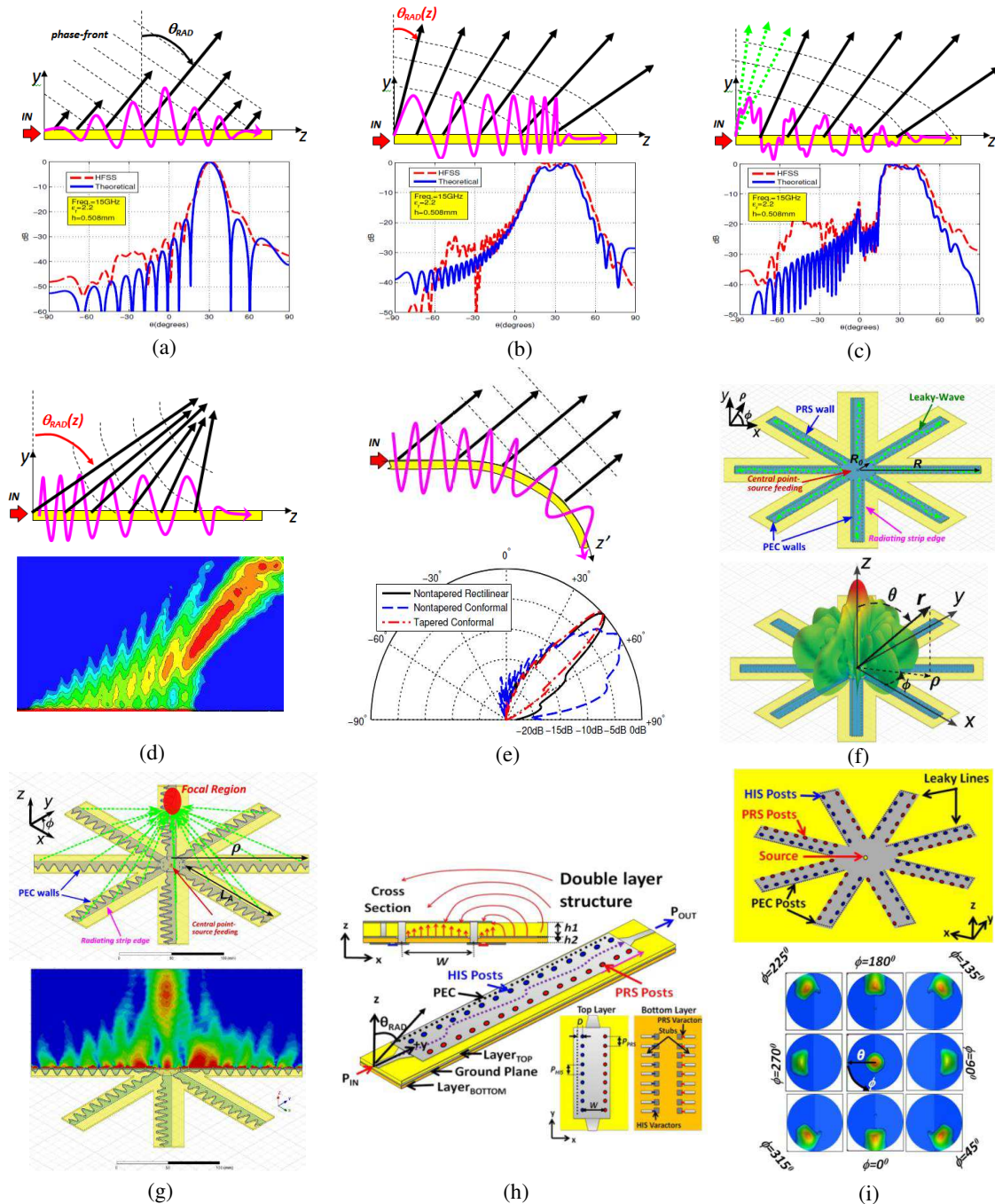


Figure 2: Different results and applications of static and dynamic (electronic) control of leaky-wave radiation in SIW technology: (a) Low-sidelobe scanning radiation pattern synthesis. (b) Broad beam shaping synthesis. (c) Synthesis of radiation nulls (dark angular regions). (d) 2D near-field focusing (microwave near-field lens). (e) Design of conformal scanning antennas. Radial array of leaky-wave antennas for (f) 3D beam far-field shaping (g) 3D near-field shaping. Electronically reconfigurable (h) 1D SIW radiator (i) 2D array of SIW radiator.

Leaky waves suffer from strong frequency dispersion of their performance [1]. In this sense all the proposed SIW designs are narrow band, which is an undesired feature for most telecoms applications. In this sense, the authors are working in the increase of the instantaneous squint-free bandwidth, as shown in Fig. 3 [21]. However, this frequency-dependent behavior can be an attractive characteristic for other applications such as FMCW radar [9], electrical prisms [16], bandpass

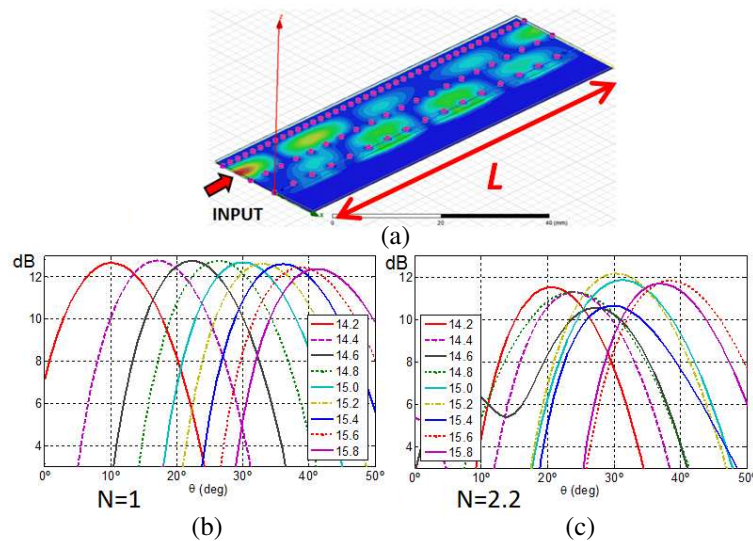


Figure 3: (a) 3D EM CAD model and propagating leaky fields for SIW antenna with reduced beam-squint response based on coupled SIW cavities [21] radiation patterns VS frequency. (b) with a single cavity and (c) with $N = 2$ cavities.

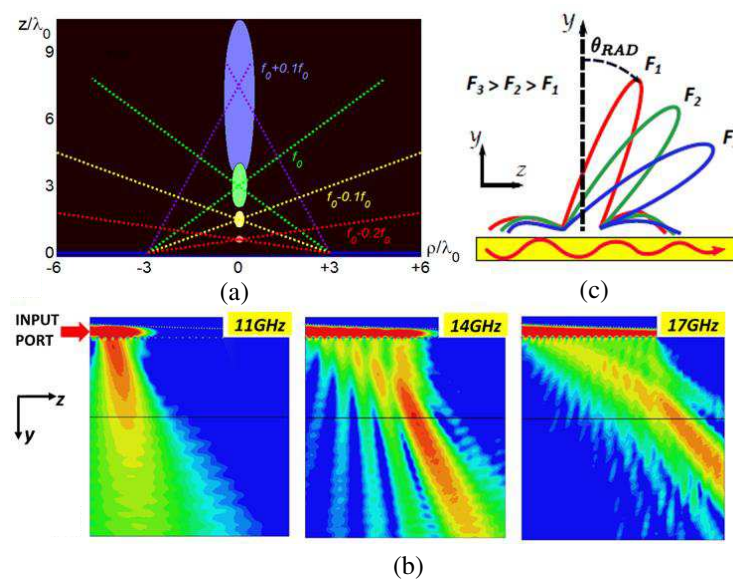


Figure 4: Different examples of application of LW frequency dispersion. (a) Electrical-prism showing frequency steering of the near-field focus [16]. (b) SIW quasi-optical multiplexer [15] and (c) RFID reader with frequency-beam-scanning for enhanced localization at indoors [22].

angular filters and multiplexers [15], RFID readers [22], and analog signal processing. This point is illustrated in Fig. 4 and it will be discussed in the oral presentation. Moreover, it is important to highlight that this strongly-dispersive response can be used to conceive electronic controllable SIW radiators as the ones shown in Fig. 2(h) and Fig. 2(i) [23], which benefit from simplicity and low cost if compared to more conventional solutions based on phased-array techniques [24].

3. CONCLUSION

The metasurfing concept can be applied to simple wavefront control in SIW technology. In this way, a variety of novel low-cost planar SIW devices have been presented, showing flexible synthesis and shaping of the near- and far-field patterns. It must be noticed that compared to conventional phased-arrays [24], it is avoided the use of complicated corporate feeding networks. Also, a clear advantage is found when comparing this modulated SIW solution with state-of-the-art reflectarrays [25] and transmitarrays [26]. This is the elimination of the external feeder thanks to the

integrated feeding mechanism of the proposed SIW devices. An extensive summary of our latest results in this field has been described; the reader is referred to the accompanying bibliography for more detailed information on each one of the presented designs and on the synthesis procedure.

REFERENCES

1. Pendry, J., D. Schurig, and D. Smith, “Controlling electromagnetic fields,” *Science*, Vol. 312, No. 5781, 1780–1782, 2006.
2. Maci, S., G. Minatti, M. Casaletti, and M. Bosiljevac, “Metasurfing: Addressing waves on impenetrable metasurfaces,” *IEEE Antennas Wireless Propag. Lett.*, Vol. 10, 1499–1502, 2011.
3. Xu, F. and K. Wu, “Guided-wave and leakage characteristics of substrate integrated waveguide,” *IEEE Trans. Microw. Theory Tech.*, Vol. 53, No. 1, 66–73, Jan. 2005
4. Gomez-Tornero, J. L., “Unusual tapering of leaky-wave radiators and their applications,” *Proc. of the 5th European Conf. on Antennas and Propagat. (EUCAP)*, 821–824, Rome, Italy, Apr. 2011.
5. Martínez-Ros, A. J., J. L. Gómez-Tornero, and G. Goussetis, “Holographic pattern synthesis with modulated substrate integrated waveguide line-source leaky-wave antennas”, *IEEE Trans. Antennas Propag.*, in press.
6. Nannetti, M., F. Caminita, and S. Maci, “Leaky-wave based interpretation of the radiation from holographic surfaces,” *2007 IEEE Antennas and Propagation Society Int. Symp.*, 5813–5816, Jun. 2007.
7. Gómez-Tornero, J. L., D. Blanco, E. Rajo-Iglesias, and N. Llombart, “Near-field focusing with holographic two-dimensional tapered leaky-wave slot antennas,” *Proc. of the 6th European Conf. on Antennas and Propag. (EUCAP)*, 234–238, Prague, Czech Republic, Mar. 2012.
8. Gómez-Tornero, J. L., D. Blanco, E. Rajo-Iglesias, and N. Llombart, “Holographic surface leaky-wave lenses with circularly-polarized focused near-fields. Part I: Concept, design and analysis theory,” *IEEE Trans. Antennas Propag.*, in press.
9. Martínez-Ros, A. J., J. L. Gomez-Tornero, and G. Goussetis, “Planar leaky-wave antenna with flexible control of the complex propagation constant,” *IEEE Trans. Antennas Propag.*, Vol. 60, No. 3, 1625–1630, Mar. 2012.
10. Zelenchuk, D., A. J. Martínez-Ros, T. Zvolensky, J. L. Gomez-Tornero, G. Goussetis, N. Buchanan, D. Linton, and V. Fusco, “W-band planar wide-angle scanning antenna architecture,” *Journal of Infrared, Millimeter, and Terahertz Waves*, Vol. 34, No. 2, 127–139, Jan. 2013.
11. Martínez-Ros, A. J., J. L. Gómez-Tornero, and F. Quesada-Pereira, “Efficient analysis and design of novel SIW leaky-wave antenna,” *IEEE Antennas and Wireless Propag. Lett.*, Vol. 12, 496–499, Mar. 2013.
12. Oliner, A. A. and D. R. Jackson, “Leaky-wave antennas,” *Antenna Engineering Handbook*, J. L. Volakis, Ed., 4th Edition, Ch. 11, Mc-Graw-Hill, New York, Jun. 2007.
13. Gómez-Tornero, J.L., A.J. Martínez-Ros, and R. Verdú-Monedero, “FFT synthesis of radiation patterns with wide nulls using tapered leaky-wave antennas,” *IEEE Antennas and Wireless Propag. Lett.*, Vol. 9, 518–521, 2010.
14. Martínez-Ros, A. J., J. L. Gómez-Tornero, F. J. Clemente-Fernández, and J. Monzó-Cabrera, “Microwave near-field focusing properties of width-tapered microstrip leaky-wave antenna,” *IEEE Trans. Antennas Propag.*, in press.
15. Martínez-Ros, A. J. and J. L. Gómez-Tornero, “Quasi-optical multiplexing using leaky-wave near-field focusing techniques in substrate integrated waveguide technology,” *Proc. of the IEEE Int. Microwave. Symp. (APS)*, Seattle, WA, USA, Jun. 2013 (in press).
16. Gómez-Tornero, J. L., F. Quesada-Pereira, A. Alvarez-Melcón, G. Goussetis, A. R. Weily, and Y. J. Guo, “Frequency steerable two dimensional focusing using rectilinear leaky-wave lenses,” *IEEE Trans. Antennas Propag.*, Vol. 59, No. 2, 407–415, Feb. 2011.
17. Gómez-Tornerom J. L., “Analysis and design of conformal tapered leaky wave antennas,” *IEEE Antennas and Wireless Propag. Lett.*, Vol. 10, 1068–1071, 2011.
18. Martínez-Ros, A. J., J. L. Gomez-Tornero, and G. Goussetis, “Conformal tapered microstrip leaky-wave antennas,” *Proc. of the 6th European Conf. on Antennas and Propag. (EUCAP)*, 154–158, Prague, Czech Republic, Mar. 2012.
19. Martínez-Ros, A. J., J. L. Gomez-Tornero, and G. Goussetis, “Broadside radiation from radial arrays of substrate integrated leaky-wave antennas,” *Proc. of the 6th European Conf. on Antennas and Propag. (EUCAP)*, 252–254, Prague, Czech Republic, Mar. 2012.

20. Martínez-Ros, A. J. and J. L. Gomez-Tornero “Free-space near-field focusing from an array of sinusoidally modulated half-mode LWAs,” *Proc. of the IEEE Int. Antennas and Propag. Symp. (APS)*, Orlando, FL, USA, July 2013 (in press).
21. Gómez-Tornero, J. L., A. J. Martínez-Ros, A. Alvarez-Melcón, F. Mesa, and F. Medina, “Substrate integrated waveguide leaky-wave antenna with reduced beam squint,” *Proc. of the IEEE European Microwave Week (EuMW)*, Nuremberg, Germany, Oct. 2013 (in press).
22. Martínez-Ros, A. J., J. L. Gomez-Tornero, and G. Goussetis, “Frequency scanning leaky-wave antenna for positioning and identification of RFID tags,” *2011 IEEE International Conference on RFID-Technologies and Applications (RFID-TA)*, 451–456, Barcelona, Sept. 2011.
23. Guzmán-Quirós, R., J. L. Gomez-Tornero, M. García-Vigueras, A. R. Weily, and Y. J. Guo, “Advances in electronically reconfigurable LWAs in Fabry-Pérot and SIW technologies,” *Proceedings of the 7th European Conference on Antennas and Propag. (EUCAP)*, 1–4, Gothenburg, Sweden, Apr. 2013.
24. DuFort, E. C., “Pattern synthesis based on adaptive array theory,” *IEEE Trans. Antennas Propag.*, Vol. 37, No. 8, 1011–1018, Aug. 1989.
25. Robustillo, P., J. Zapata, J. A. Encinar, and J. Rubio, “ANN characterization of multi-layer reflectarray elements for contoured-beam space antennas in the Ku-band,” *IEEE Trans. Antennas Propag.*, Vol. 60, No. 7, 3205–3214, Jul. 2012.
26. Padilla, P., A. Munoz-Acevedo, M. Sierra-Castaner, and M. Sierra-Perez, “Electronically reconfigurable transmitarray at Ku band for microwave applications,” *IEEE Trans. Antennas Propag.*, Vol. 58, No. 8, 2571–2579, 2010.

Continuous Transition of Heat Transport across a Closing Vacuum Gap from Thermal Radiation to Thermal Conduction

B. V. Budaev and D. B. Bogy
University of California, Berkeley, USA

Abstract— The paper discusses the common physical origin of heat conductance and heat radiation, proposes a simple model of micro and nanoscale heat transport caused by Casimir/van der Waals intermolecular forces, and addresses some difficulties arising in common applications of the conventional theory of heat radiation to micro and nano scale systems.

1. INTRODUCTION

For a long time conductance and radiation have been viewed as different mechanisms of heat transport based on different physical phenomena. Conduction dominates heat transport in solids, it is governed by the Fourier law and is described by the heat/diffusion equation. Radiation appears as the only means for heat exchange across a vacuum gap, it is governed by the laws of electromagnetism and, therefore, is described in terms of Maxwell’s equations.

However, in early studies of heat transport at micro and nano scales it was noticed that the conventional theory of radiative heat transport provides calculations that differ from experimental observations. Thus, while the classical theory predicts the heat transport coefficient between two half-spaces separated by a vacuum gap should be independent of the gap’s width h , it is observed [1] that this coefficient starts increasing when h decreases below a few microns, and more recent experiments [2] suggest that the increase is proportional to $1/h^2$ as the gap narrows. Although these observations contradict classical theories of radiative heat transport, they agree with common sense which suggests that as h vanishes the radiation heat transport across it should evolve to heat conductance across the resulting interface. In particular, the thermal resistance of a vacuum gap should gradually evolve to the interface thermal resistance, known as Kapitsa resistance, as the gap closes.

It is shown in [3] that when the distance h between separated bodies is comparable to the interatomic distance, then the molecules of these bodies interact via the short-range electric forces, as if they belong to a single body, and this results in heat transport by a mechanism similar to conductance. Conversely as h increases, short-range intermolecular interactions gradually diminish compared to the long-range interactions responsible for radiation, and this provides a gradual transition from heat conductance to heat radiation. It is expected that the heat conductance due to interatomic forces starts exceeding heat radiation at gaps of the order ~ 5 nm and below, which suggests that this phenomenon may find application in the design of Heat Assistance Magnetic Recording (HAMR) systems where the spacing between the head and disk is less than 5 nm.

It is shown in [3] that for intermediate-range separations, between $h \sim 10$ nm and $h \sim 3000$ nm, the heat transport coefficient follows the rate $1/h^2$ observed in [2], but not predicted by the conventional theory of thermal radiation. It is demonstrated that this failure of the conventional theory is caused by unjustified applications of some theoretical concepts, such as of the Fluctuation-Dissipation theorem, which, as stated in [4], is not applicable for the analysis of heat transport in nanoscale systems. A foundation for a self-consistent approach to radiative heat transport in nanoscale systems was proposed in [5], and a procedure for its implementation is discussed in the end of this paper.

2. A MODEL OF HEAT CONDUCTANCE DUE TO CASIMIR EFFECT

The electric field of a moving particle with the charge q is described by the expression [6]

$$\mathbf{E} = \frac{-\mathbf{q}}{4\pi\epsilon} \left\{ \underbrace{\frac{\mathbf{e}_r}{|\mathbf{r}|^2} + \frac{\mathbf{r}}{c} \frac{d}{dt} \left(\frac{\mathbf{e}_r}{|\mathbf{r}|^2} \right)}_{\text{Coulomb term}} + \underbrace{\frac{1}{c^2} \frac{d^2 \mathbf{e}_r}{dt^2}}_{\text{Radiation term}} \right\}, \quad \mathbf{e}_r = \frac{\mathbf{r}}{|\mathbf{r}|}, \quad (1)$$

where c is the speed of light, and \mathbf{r} is the “retarded” vector connecting the observer with the particle’s position at the time $\Delta t = -|\mathbf{r}|/c$, when this field was “radiated”. The first two terms in (1)

describe the static Coulomb field and its correction due to the finite speed of light propagation. These terms decay proportionally to $1/|\mathbf{r}|^2$, while the third term describing electromagnetic radiation decays proportionally to $1/|\mathbf{r}|$. Therefore, the first two terms dominate at short distances and make a major contribution to intermolecular forces which are responsible for heat conductance, and the third term dominates at long distances and is responsible for heat radiation.

Since matter consists of many particles, it may be impractical and idealistic to analyze heat transport by computing the superpositions of the fields (1) generated by all atoms. However, as in the case of conventional conductance, it is possible to derive some simple macroscopic models of heat conductance between slightly separated bodies.

It is generally accepted that heat in solid dielectrics is due to elastic waves of lattice vibrations. To model heat conductance between dielectrics separated by a vacuum gap we consider a one-dimensional chain shown in Fig. 1 where the masses $m = \rho_{\mp}a$ are located at $x_n = an < 0$ and at $x_n = h + an$, where $n \geq 0$ and $h > 0$. Assume that the springs inside the half-chains $x < 0$ and $x > h$ have the elastic moduli γ_{\mp} and the spring connecting x_{-1} and x_0 has the modulus γ_h .

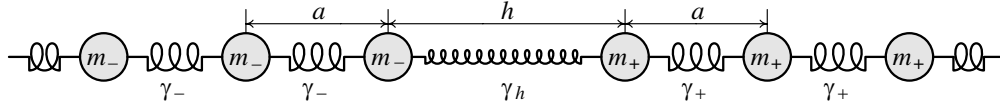


Figure 1: Two separated chains.

The motion of this chain is described by the equations

$$\rho_{\pm}a^2\ddot{\xi}(x_n) = \gamma_{\pm} [\xi(x_{n+1}) + \xi(x_{n-1}) - 2\xi(x_n)]. \quad (2)$$

valid for all particles with $n \neq -1$, $n \neq 0$, and by two additional equations

$$\begin{aligned} \rho_-a\ddot{\xi}(0) &= \frac{\gamma_h}{h} [\xi(h) - \xi(0)] + \frac{\gamma_-}{a} [\xi(-a) - \xi(0)], \\ \rho_+a\ddot{\xi}(h) &= \frac{\gamma_h}{h} [\xi(0) - \xi(h)] + \frac{\gamma_+}{a} [\xi(h+a) - \xi(h)], \end{aligned} \quad (3)$$

for the particles at the boundaries of the half-chains. If $a \rightarrow 0$ then (2) converges to the wave equation $\ddot{\xi}(x) = c_{\pm}^2 \nabla^2 \xi(x)$, where x is a continuous coordinate, $c_{\pm} = \sqrt{\gamma_{\pm}/\rho_{\pm}}$ are the sound speeds, and (3) reduces to the interface conditions

$$\gamma_- \xi'(0) = \gamma_+ \xi'(h) = \gamma_h [\xi(h) - \xi(0)] / h, \quad (4)$$

which compliment the wave equation describing the motions in the homogenous domains.

If the interconnection $0 < x < h$ is very strong in the sense that $\gamma_h/h \rightarrow \infty$ then (4) reduce to the condition $\xi(t, h) = \xi(t, 0)$, which implies that the boundaries $x = 0$ and $x = h$ are firmly connected. In the opposite case of a very weak interconnection, (4) reduces to the Neumann conditions $\xi'(0) = \xi'(h) = 0$, which imply that the domains $x < 0$ and $x > h$ move independently of each other.

To describe a realistic three dimensional medium this model can be combined with Debye's theory, which assumes that heat is carried by acoustic waves described in terms of a pressure $p(\mathbf{r})$ related to the displacement vector field $\boldsymbol{\xi}(\mathbf{r})$ by $\rho \ddot{\boldsymbol{\xi}} = -\nabla p$. The pressure satisfies the equations $\ddot{p} = c_{\pm}^2 \nabla^2 p$, describing the motions in the homogenous half-spaces $x > h$ and $x < 0$, and it also obeys the interface conditions

$$\gamma_- p''(0) = \gamma_+ p''(h) = \gamma_h [p'(h) - p'(0)] / h, \quad (5)$$

which generalize (4). To use these interface conditions it is necessary to know the elastic modulus γ_h of the vacuum gap of width h . If h is large compared to the intermolecular distance then γ_h can be estimated as $\gamma_h = h|F'(h)|$, where $F(h)$ is the force of interaction between the half-spaces. As shown in [7], for $h \ll \lambda_0$, where λ_0 is the dominant wavelength of electromagnetic radiation, $F(h) \approx C/h^3$. Correspondingly, the modulus γ_h has the asymptote $\gamma_h \approx \gamma_0 a^3/h^3$, where the factor $\gamma_0 a^3$ is determined by an assumption that as h reduces to the interatomic distance a , then γ_h should approach the average γ_0 of the elastic moduli of the half-spaces.

Properties of the wave motion imply that the capability of acoustic waves to carry heat between two interacting half-spaces $x < 0$ and $x > h$ is proportional to the square $|K|^2$ of the transmission coefficient K of the gap. If the width of the gap h is bigger than the wave length Λ of thermally excited acoustic waves, which is of the order of ~ 1 nm at room temperature, then

$$|K| \approx \gamma_h \Lambda / \gamma_0 h, \quad (h \gg \Lambda \approx 1 \text{ nm}), \quad (6)$$

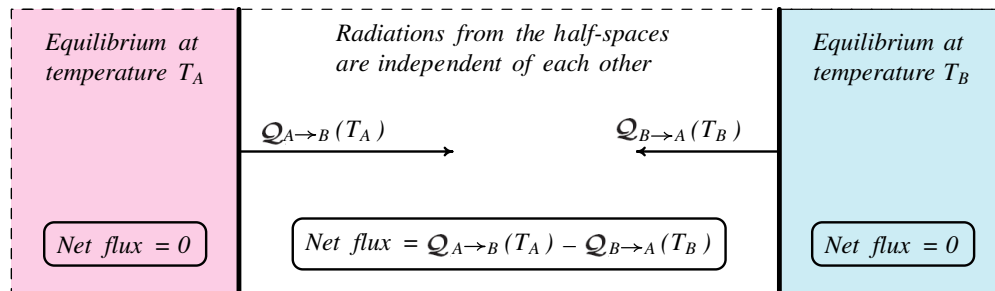
where γ_0 is the average of the elastic moduli of the half-spaces, and γ_h is the elastic modulus of the vacuum gap. This estimate conforms with the expectations that a vanishingly narrow gap between identical media has full transmission and that a wide gap has no transmission.

The above implies that acoustic waves can penetrate the vacuum gap separating material half-spaces and that, therefore, these waves can carry heat across a vacuum gap. However, the importance of this channel of heat transfer strongly depends on the gap's width. In particular, for gaps wider than ~ 10 nm electromagnetic radiation remains the sole heat carrier, but for gaps narrower than ~ 5 nm acoustic waves become the dominant heat carriers.

3. RADIATIVE HEAT TRANSPORT IN NANOSCALE

Although intermolecular Casimir/van der Waals forces cause intensive heat transport across a gap narrower than ~ 5 nm, these forces do not provide a noticeable contribution to heat exchange across gaps wider than 30 nm, such as those experimentally studied in [2]. Correspondingly, these forces cannot explain why the heat transport coefficients of 30 nm–1000 nm gaps measured in [2] appear to be up to two orders of magnitude higher than predicted by the conventional theory.

Since experiments suggest that the theory of heat radiation has flaws which show up in the nanoscale, these flaws must be identified because otherwise any modification of the theory will mislead further studies in a wrong direction. To discuss the main flaw of this theory it suffices to consider two homogeneous half-spaces separated by a vacuum gap, as shown in Fig. 2.



- 1) This scheme does not comply with conservation of energy
- 2) As the gap vanishes, the net flux remains finite instead of diverging

Figure 2: Inconsistency of the conventional approach to radiative heat transport.

The conventional approach to radiative heat transport is based on the assumptions that objects that emit or absorb heat may be treated as if they remain in thermal equilibrium at constant temperatures, and that the ensembles of thermally excited electromagnetic fields in heat exchanging objects are statistically uncorrelated. If these assumptions were correct then the net heat flux between bodies A and B could be represented by the difference

$$\mathbf{Q} = \mathbf{Q}_{A \rightarrow B}(T_A) - \mathbf{Q}_{B \rightarrow A}(T_B), \quad (7)$$

where the symbol $\mathbf{Q}_{X \rightarrow Y}(T_X)$ represents the flux, which would be radiated from the body X at temperature T_X to the place occupied by the body Y , under the assumption that Y does not exist.

In larger scale systems this approach, combined with Planck's law used to determine $\mathbf{Q}_{X \rightarrow Y}(T_X)$, leads to many well-tested results, such as the Stefan-Boltzmann law, and hence its assumptions appear to be well-justified by experiments. Nevertheless, a closer look suggests that these assumptions cannot be valid in cases when the distance h between heat exchanging bodies is smaller than or even comparable with the dominant wavelength λ of the thermal radiation. Indeed, the uncertainty principle implies that a field with the dominant wavelength λ cannot be localized in a domain smaller or even comparable with λ . Therefore, if two bodies are separated by the distance $h < \lambda$, then it is impossible to distinguish which of them radiates an electromagnetic field with

the dominant wavelength λ . Correspondingly, if $h < \lambda$, then the assumption of the conventional approach regarding the statistical independence of the radiations from different bodies is violated and, therefore, this approach can not be reliably used.

Despite the transparency of the above reasoning explaining why thermal radiations from bodies separated by a nanoscale gap are correlated, there is a common misconception that the statistical independence of such radiations follows from the Fluctuation-Dissipation theorem [8], which states that under certain conditions the cross-correlation of different components of thermally excited currents can be computed as

$$\langle J_l(\mathbf{r}, \omega) \mathbf{J}_m^*(\mathbf{r}_1, \omega) \rangle = \frac{1}{\pi} \text{Im}(\epsilon_\omega) \omega \Theta(\omega, T) \delta_{lm} \delta(\mathbf{r} - \mathbf{r}_1), \quad (8)$$

where the brackets $\langle \cdot \rangle$ denote averaging, \mathbf{r} and \mathbf{r}_1 are two points, ω is the frequency, J_l and J_m are the components of the thermally excited current, ϵ_ω is the permittivity, and $\Theta(\omega, T)$ is the average energy of an oscillator at frequency ω in an *equilibrium* ensemble at temperature T .

If (8) is valid, then the currents at $\mathbf{r} \neq \mathbf{r}_1$ and, correspondingly, the radiations from these points are uncorrelated. Then, assigning \mathbf{r} and \mathbf{r}_1 to different bodies A and B , as in Fig. 2, one can show that the radiations from A and B are statistically independent. This suggests that the flow of energy from A to B is determined solely by the properties A , and the flow from B to A is determined solely by B . So, the theorem (8) seems to imply that the radiative heat transfer between two bodies is described by formula (7), regardless of the scale of separation. However, this theorem is clearly restricted to equilibrium systems where the heat flux vanishes *a priori* and, thus, does not need to be computed by any method.

To further illustrate that (8) is not applicable to the analysis of heat transport in layered nanoscale systems it suffices to notice that it involves only one temperature T , that it involves the spectrum $\Theta(\omega, T)$ of an equilibrium ensemble, and that “... *the use of equilibrium laws (including FDT) is no longer quite rigorous, but still justified if, as often the case, the role of the transport phenomena is as yet insignificant*”, as stated in [?, Page 112].

This reasoning eliminates any practical need to discuss applications of the Fluctuation-Dissipation Theorem to the analysis of radiative heat transport. Nevertheless, it is worth mentioning that this theorem is not applicable even to nanoscale systems in equilibrium.

Indeed, since (8) is applied to physical systems, then the δ -function should be considered as a distribution over a finite domain of a correlation radius $\epsilon \ll 1$, which has “... *same order of magnitude as size of the nonlocality region in the material equations*”, [?, Pages 122-123]. But the system of two half-spaces from Fig. 1 has an inhomogeneity of width h , and this means that the carrier of the δ -function in (8) is spread over the domain of the size $\sim h$ comparable with the distance between the half-spaces. Therefore, the Fluctuation-Dissipation theorem implies that the electric currents on different sides of the gap of the width H are correlated, which violates the main assumption of the conventional theory of radiative heat transport.

The above discussion implies that to understand heat transport across sub-micron gaps it is necessary to admit that thermal radiations from closely spaced objects are correlated and to include their correlation into the analysis, which radically distinguishes the theory of radiative heat transport in nano and low-micro scale structures from the conventional macroscopic theory.

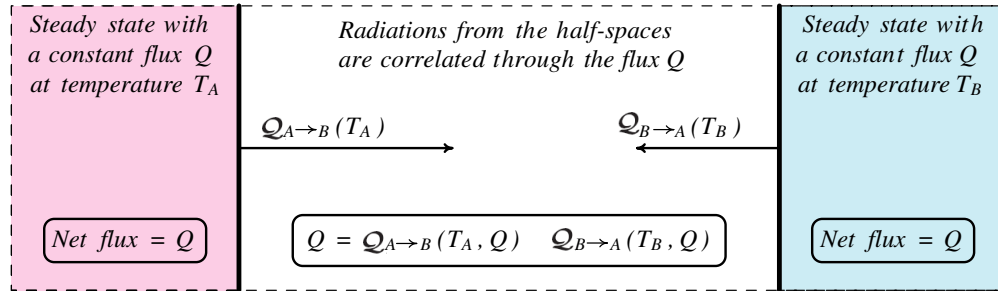
In [4] the correlation between the radiations from closely spaced objects is taken into account by a modification of the Fluctuation-Dissipation theorem. This paper considers systems of stationary non-equilibrium, which make it possible to represent the correlations between radiations from different bodies in terms of the scattering operators associated with these bodies.

Another approach to nanoscale radiative heat transport is based on representation of the energy flows from A to B and from B to A as functions $\mathcal{Q}_{A \rightarrow B}(T_A, \mathbf{Q})$ and $\mathcal{Q}_{B \rightarrow A}(T_B, \mathbf{Q})$ depending not only on the temperatures of the corresponding bodies, but also on the net flux \mathbf{Q} . Then, \mathbf{Q} satisfies the equation

$$\mathbf{Q} = \mathcal{Q}_{A \rightarrow B}(T_A, \mathbf{Q}) - \mathcal{Q}_{B \rightarrow A}(T_B, \mathbf{Q}), \quad (9)$$

which couples the radiations from A and B through the unknown net flux \mathbf{Q} and connects this flux with the temperatures T_A and T_B . This approach, illustrated in Fig. 3, is based on the extension of Planck’s law of equilibrium thermal radiation to systems with a steady heat flux [5]. The technique developed on these ideas makes it possible to compute the terms in the right-hand side of (9) and, thus, to reduce the problem to the numerical analysis of that equation. Therefore, this approach appears as a straightforward generalization of the conventional theory of radiative heat transport

obtained by the elimination of the invalid assumptions regarding the independence of thermal radiations from closely separated bodies.



This scheme establishes the relationship between the temperatures and the net flux

Figure 3: Self-consistency of the proposed approach to radiative heat transport.

4. CONCLUSION

The outlined method can be applied to the analysis of heat transport by any mechanism related with wave propagation, including radiative transport by electromagnetic waves and heat conduction due to acoustic waves, including that caused by Casimir/van der Waals forces. Its application to radiative transport easily explains the $\sim 1/h^2$ dependence of the heat transport coefficient of a gap of intermediate h between ~ 10 nm and ~ 3000 nm. Its application to the heat conduction across the interface made it possible to get correct order-of-magnitude estimates to Kapitza resistance, without any data fitting and with use of very crude zero-order approximations of the obtained equations. Finally, since this method describes heat radiation and heat conduction in similar terms, it provides a smooth transition between radiation and conduction across a closing vacuum gap.

ACKNOWLEDGMENT

This work was supported by the William S. Floyd, Jr., Distinguished Chair, held by D. Bogy.

REFERENCES

1. Cravalho, E. G., C. L. Tien, and R. P. Caren, "Effect of small spacing on radiative transfer between two dielectrics," *J. Heat Transport*, Vol. 89, No. 3, 351–358, 1967.
2. Narayanaswamy, A., S. Shen, L. Hu, X. Chen, and G. Chen, "Breakdown of the Planck blackbody radiation law at nanoscale gaps," *Appl. Phys. A*, Vol. 96, 357–362, 2009.
3. Budaev, B. V. and D. B. Bogy, "On the role of acoustic waves (phonons) in equilibrium heat exchange across a vacuum gap," *APL*, Vol. 99, No. 5, 053109-1–053109-3, 2011.
4. Messina, R. and M. Antezza, "Casimir-Lifshitz force out of thermal equilibrium and heat transfer between arbitrary bodies," *EPL (Europhysics Letters)*, Vol. 95, No. 6, 61002-p1–61002-p6, 2011.
5. Budaev, B. V. and D. B. Bogy, "Extension of Planck's law of thermal radiation to systems with a steady heat flux," *Ann. der Physik*, Vol. 523, No. 10, 791–804, 2011.
6. Feynman, R. P., R. B. Leighton, and M. L. Sands, *The Feynman Lectures on Physics, V. 1.*, 3rd Edition, Pearson/Addison-Wesley, 2006.
7. Israelachvili, J., *Intermolecular and Surface Forces*, 2nd Edition, Academic Press, London, 1991.
8. Rytov, S. M., Yu. A. Kravtsov, and V. I. Tatarskii, *Principles of Statistical Radiophysics. 3. Elements of Random Fields*, Springer-Verlag, Berlin, 1987.
9. Budaev, B. V. and D. B. Bogy, "An extension of Khalatnikov's theory of Kapitza thermal resistance," *Ann. der Physik*, Vol. 523, No. 3, 208–225, 2011.

A Triple-mode Passband Filter with Improved Upper-stopband Performance

Huan Yang¹, Shao-Bin Liu¹, Xiang-Kun Kong^{1,2}, Hui-Chao Zhao¹, and Bo-Rui Bian¹

¹College of Electronic and Information Engineering

Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

²Jiangsu Key Laboratory of Meteorological Observation and Information Processing
Nanjing University of Information Science and Technology, Nanjing 210044, China

Abstract— In this letter, a high selectivity microstrip bandpass filter (BPF) with loaded stubs is proposed. This filter consists of a half-wavelength transmission line with two different open-ended stubs at the line center. Compared with conventional uniform stub, the circular impedance-stepped stub can improve the performance of the upper-stopband. This filter can generate three operation modes, which can be easily adjusted individually. Owing to the symmetrical structure of it, we can use the odd-even-mode method to analysis it. When the coupling is weak, the even-modes frequency can be flexibly controlled by the open-ended stubs. Due to the two unsymmetrical stubs, two transmission zeros are brought out at low and high rejection bands, which is leads to high selectivity. Simulated results show that central frequency is 6.3 GHz with 3-dB fractional bandwidth of 49%. Its upper-stopband is extended 8.4–16.7 GHz with a rejection level of about -25 dB.

1. INTRODUCTION

With the rapid development of modern wireless and mobile communication systems, filters, which play an important and essential role in systems, are facing ever more stringent requirements—higher performance, smaller size, light weight, and lower cost. Many kinds of BPF with good stopband performance have been proposed [1–6]. In [1], a high selectivity quadruple-mode BPF with source-load coupling is proposed. It uses SIR to adjust harmonic frequency, and results to get upper stopband performance. In [2], a novel compact quadruplet BPF based on alternative J/K inverters and quarter-wavelength resonators is presented. In the structure, the transmission zeros produced by the cross-coupled via magnetic coupling between the first and fourth resonators, and the source-load coupling, contribute to obtain good out-of-band performance. In [3, 4], by introducing open stubs, the filter can get extra transmission zeros, and good stopband performance. Furthermore, EBG structures are also introduced to the design of filters [5, 6], which combines with coupled line can get excellent stopband performance.

In this letter, a compact BPF with improved upper stopband performance is presented. This filter is perfectly symmetric in structure, so the odd-even-mode method can be used to analysis it. Based on [7, 8], we can learned that antenuation poles could be controlled by the load stubs. The open circular stub is introduced to get better upper stopband performance. The filter is designed using HFSS and it is implemented on the substrate with a relative dielectric constant of 2.2 and a thickness of 0.508 mm.

2. FILTER DESIGN

Figure 1(a) shows the geometry of the proposed filter. The resonator consists of a half-wavelength transmission line with two different open-ended stubs at the line center. Since the resonator is perfectly symmetrical to the centre plane, the odd-even-mode method can be used to it. For even-mode excitation, there will be no current flow through the symmetrical plane. The structure will be equivalent to another structure of Fig. 1(b). Because of two current paths in this circuit, two even-mode frequencies come out. The frequency of path 1 is f_{even}^1 , and the other one is f_{even}^2 . For odd-mode excitation, there will be a voltage null in the middle of the resonator. The resonator can be equivalent to a structure of Fig. 1(c), and its corresponding frequency is named as f_{odd} .

From [9], we can conclude that, the change of even frequency depends on the loaded stub, while the odd frequency is constant. So, we can use the radius of circular stub (R) to change f_{even}^1 . Fig. 2(a) shows the relationship between R and f_{even}^1 under the weak coupling case. It can be obviously seen that, when R is selected as follows: 3.6 mm, 3.9 mm and 4.2 mm, the f_{even}^1 moves towards the upper frequency, whereas the resonant frequencies f_{odd} and f_{even}^2 are hardly changed. While we can change the other stub length (L_3) to effect f_{even}^2 . Fig. 2(b) shows the relationship

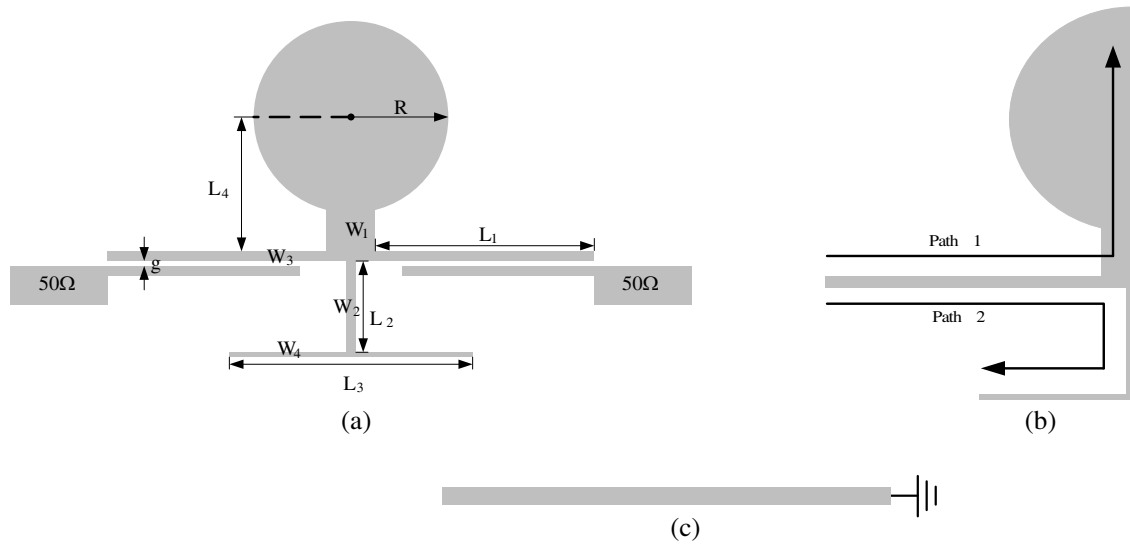


Figure 1: (a) Schematic of the triple-mode BPF. (b) Even mode structure of the triple-mode BPF. (c) Odd mode structure of the triple-mode BPF.

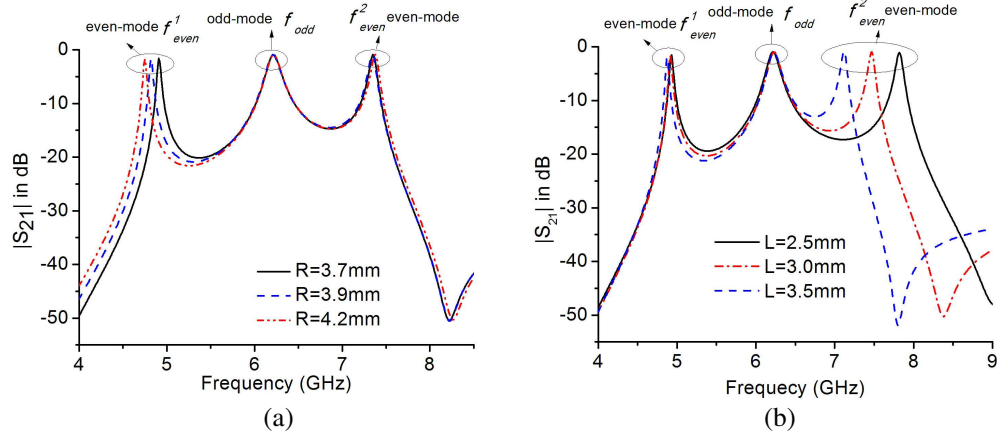


Figure 2: Frequency-dependent responses of the resonator under weak coupling. (a) Versus the radius of circular stub (R). (b) Versus the stub length (L_3).

between L_3 and f_{even}^2 under the weak coupling case. It can be easily known that the f_{even}^2 moves towards the lower frequency with increased L_3 , whereas the resonant frequencies f_{odd} and f_{even}^1 are hardly changed. It is reported in [7], the length of central tapped open stub can change the transmission zero location. And from [8], the length proportion of the half-wavelength resonator and the center circular open stub can change the transmission zero location. So, based on [7, 8, 10], we can learn that, the transmission zero in low-stopband is mainly controlled by the circular stub, and the other transmission zero in up-stopband mainly relates to the other stub. The upper-stopband performance is improved by the loaded circular stubs. So we also use circular stub to get better upper-stopband performance. Consequently, a compact triple-mode BPF with improved upper-stopband performance can be obtained. And the optimized parameters of the proposed filter are shown in Table 1.

The proposed filter is fabricated on a thin dielectric substrate with a low relative permittivity of

Table 1: Dimensions of the proposed filter (UNIT: mm).

Parameters of the proposed filter							
L_1	8.6	L_2	3	L_3	6	L_4	5.5
W_1	0.5	W_2	0.2	W_3	0.2	W_4	0.1
R	3.8	g	0.2				

2.2, a loss $\tan \delta$ of 0.0009 and a thickness of 0.508 mm. The 50 ohm feed line is terminated with a standard SMA connector to facilitate the measurement and connect with other standard microwave modules.

3. RESULTS AND DISCUSSION

The filter is fabricated on the RT/Duriod 5880 substrate and its photograph is shown in Fig. 4. The S -parameters are measured using an N5230A network analyzer, and are plotted in Fig. 5 together with simulated results. Simulated results show that the central frequency is 6.75 GHz with 3-dB fractional bandwidth of 45.3%. Two transmission zeros near the cut-off frequency are located at 3.59 GHz and 8.5 GHz resulting in sharp skirt. The measured upper stopband with 25 dB attenuation level is extended to 16.7 GHz. In summary, the measured and simulated results are well complied with each other.

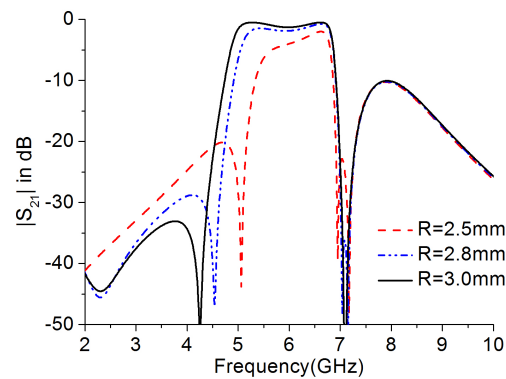


Figure 3: The transmission zero in low-stopband versus the radius of circular stub (R).

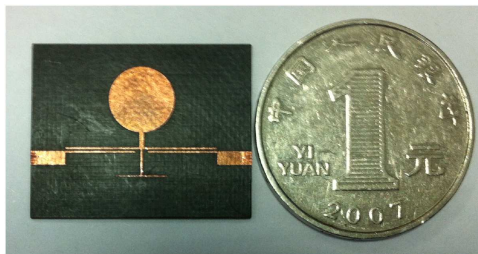


Figure 4: Photograph of the fabricated filter.

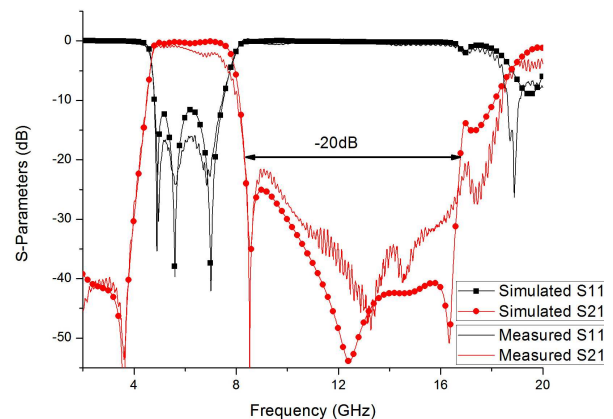


Figure 5: Comparison between simulated and results of proposed filter.

4. CONCLUSIONS

The application of loaded open/short stub in microstrip filter design has been studied very intensively. It revealed that a wideband triple-mode BPF with two transmission zeros at both skirts and good upper stopband performance can be obtained easily. The filter, whose fractional bandwidth of 45.3%, is fabricated to demonstrate the predicted performances in experiment.

ACKNOWLEDGMENT

This work was supported by the supports from Chinese Specialized Research Fund for the Doctoral Program of Higher Education (grant No. 20123218110017), the Jiangsu Province Science Foundation (Grant No. BK2011727), the Foundation of Aeronautical Science (No. 20121852030), the Foundation of Outstanding Doctoral Dissertation in NUAU (Grant No. BCXJ11-05) and the Open Research Program in Jiangsu Key Laboratory of Meteorological Observation and Information Processing (Grant No. KDXS1207).

REFERENCES

1. Zhou, L., S.-B. Liu, Y.-N. Guo, X.-K. Kong, and H.-F. Zhang, “A high selectivity quadruple-mode BPF with two short-circuited stub-loaded SIRs,” *Progress In Electromagnetics Research Letters*, Vol. 24, 43–50, 2011.
2. Zhang, S., L. Zhu, and R. Li, “Compact quadruplet bandpass filter based on alternative J/K inverters and resonators,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 22, No. 5, 224–226, May 2012.
3. Zhang, X.-Y., Q. Xue, and B.-J. Hu, “Novel bandpass filter with size reduction and harmonic suppression,” *Microwave Opt. Technol. Lett.*, Vol. 49, No. 5, 914–916, Apr. 2007.
4. Wu, X.-H., Q.-X. Chu, X.-K. Tian, and X. OuYang, “Quintuple-mode UWB bandpass filter with sharp roll-off and super-wide upper stopband,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 21, No. 12, 661–663, Dec. 2011.
5. Wong, S. W. and L. Zhu, “EBG-embedded multiple-mode resonator for UWB bandpass filter with improved upper-stopband performance,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 17, No. 6, 421–423, Jun. 2007.
6. Yao, B., Y. Zhou, Q. Cao, and Y. Chen, “Compact uwb bandpass filter with improved upper-stopband performance,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 19, No. 1, 27–29, Jan. 2009.
7. Lee, J.-R., J.-H. Cho, and S.-W. Yun, “New compact bandpass filter using microstrip resonator with open stub inverter,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 10, No. 12, 526–527, Dec. 2000.
8. Zhang, X.-C., Z.-Y. Yu, and J. Xu, “Design of microstrip dual-mode filters based on source-load coupling,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 18, No. 10, 677–679, Oct. 2008.
9. Zhang, X. Y., J.-X. Chen, Q. Xue, and S.-M. Li, “Dual-band bandpass filter using stub-loaded resonators,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 17, No. 8, 583–585, Nov. 2007.
10. Zhu, L. and W. Menzel, “Compact microstrip bandpass filter with two transmission zeros using a stub-tapped half-wavelength line resonator,” *IEEE Microw. Wireless Compon. Lett.*, Vol. 13, No. 1, 16–18, Jan. 2003.

Automated Design of Frequency Selective Surfaces with the Application to Wi-Fi Band-stop Filter

P. Tomasek and S. Gona

Tomas Bata University in Zlin, The Czech Republic

Abstract— This article presents a technique for analysis and automated design of frequency selective surfaces. In this work the method of moments is used to analyse a planar periodic structure. The approach allows to automate the whole process of the filter design and frees the users from the detailed knowledge of the filter design theory.

For the practical part of the paper, the algorithm of Levenberg-Marquardt is chosen as a local optimisation method. Whole process of automation is implemented in Matlab. An optimisation of a band-stop filter for Wi-Fi signals serves as a practical example. A Wi-Fi device communicating under standard 802.11b or 802.11g uses a specific channel which has frequency between 2.412 and 2.484 GHz. Therefore the goal is to design a band-stop filter which ideally does not transmit mentioned band of frequencies. The geometry of Jerusalem-cross serves as a structure to be optimized. Four design variables are defined for optimisation: the width and the height of a periodic cell (a square), the width of an arm and two other variables which influence the length of an outer arm and the length of the whole structure of the Jerusalem-cross. This filter consists of two dielectrically separated conductive layers which make the filter more narrow-band. S-parameters of the final optimized geometry of the Wi-Fi filter are presented in the paper.

1. INTRODUCTION

Frequency Selective Surfaces (FSSs) are important spatial filters which can efficiently filter desired band of frequencies. Therefore they can play a significant role in electromagnetic related problems.

To briefly sketch the history, the beginning of FSS relates to Ben A. Munk which was the guru of this approach [1]. In the last decade, the idea of FSS has spread out into many applications. Example of a band-pass FSS is in [2, 3] where the goal was to transmit GSM signals through energy efficient windows. The first FSS absorber was presented by Salisbury and Jaumann [4, 5]. Ghaffer et al. [6] and Umair et al. [7] proposed a novel and compact design to obtain stable frequency response by absorbing 5 GHz Wi-Fi signals.

In our paper we investigate the possibility of filtering of 2.4 GHz Wi-Fi signal, and we present a process of automation of filter design which frees the users from the detailed knowledge of the filter design theory.

2. STATEMENT OF THE PROBLEM

Assume that there is a need to prevent transmission of Wi-Fi signal so that it cannot spread out of a given room.

A Wi-Fi device communicating under standard 802.11b or 802.11g uses a specific channel which has frequency between 2.412 and 2.484 GHz [8]. Therefore the goal is to design a band-stop filter which ideally does not transmit mentioned band of frequencies.

3. DESIGN OF AN APPROPRIATE FSS

A double layer Jerusalem-cross is chosen as the schema to be optimized, see Fig. 1. The first reason for this choice was in potentially better reflection in comparison with a simple cross. The second reason was the relative simplicity of the model which can be modelled by rectangular elements. In Fig. 1, a represents the width and height of a cell, l is the total width and height of the Jerusalem-cross, w is the width of an arm and l_e represents the length of the bar connected to the end of an arm.

The electrical conductivity of the metallization is 56 MS/m and the thickness is 17 μm . The relative permittivity of the dielectric layer is 1.0 and the thickness is 1.57 mm.

4. OPTIMIZATION

A frequency range, an initial geometry with design variables (e.g., width and height of the arms of the cross) and optimization goals must be set before performing the optimization of an FSS filter.

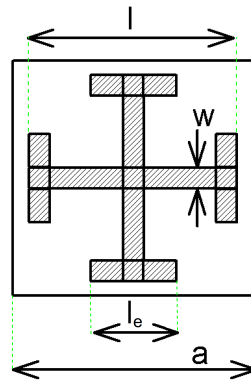


Figure 1: Schema of a cell containing the Jerusalem-cross.

The transmission coefficient depends on frequency and other parameters forming the parameter vector of the filter which specifies the geometry (defined by design variables). An optimization method searches for the set of parameters which satisfies the given objectives, at least approximately, being thus in a certain sense optimal.

An optimization goal is defined by a frequency range where the transmission coefficient must be lower or greater than a threshold value set by the user.

In our experiment three optimization goals were modelled (also presented in Fig. 2 together with results of initial configuration):

1. To pass frequencies from 1.0 to 2.2032 GHz (threshold: -2.5 dB)
2. To stop frequencies from 2.3256 to 2.5704 GHz (threshold: -20.0 dB, this range relates to the Q factor equal to 10)
3. To pass frequencies from 2.6928 to 5.0 GHz (threshold: -2.5 dB)

The initial values of design parameters with lower and upper bounds are mentioned in Table 1 where $l = k_1 a$ and $l_e = k_2 l$.

Table 1: Description of design parameters (LB and UB stands for lower and upper bound).

Parameter	Description	Initial Value	LB	UB
a	The width and height of a cell [m]	0.05	0.03	0.07
w	The width of an arm [m]	0.002	0.001	0.003
k_1	The width parameter ($k_1 = l/a$)	0.85	0.7	1.0
k_2	The length parameter ($k_2 = l_e/l$)	0.35	0.2	0.5

In our work, optimization was performed numerically using a local optimizer Levenberg-Marquardt (a possible alternative is *fmincon* [14] or *fminsearchbnd* [13] which can be directly used in Matlab). The method of moments [1, 9–11] was chosen for analysis and estimation of the FSS transmission coefficients. All computations were based on perpendicular angle of incidence only. In this study, we used FSSMR software [12] which was developed at Tomas Bata University in Zlin and which analyses the planar periodic structures and tries to optimize them with respect to the optimization goals.

5. RESULTS

The optimization procedure results in an well-performing filter which suppresses the desired band of frequencies (from 2.412 to 2.484 GHz). The transmission for these frequencies is less than -20 dB. The final transmission coefficients are presented in Fig. 3. The process of optimization in Matlab took at about 30 hours using an average computer.

The optimized values of design parameters are presented in the list below:

- $a = 0.042554$ m
- $w = 0.002991$ m

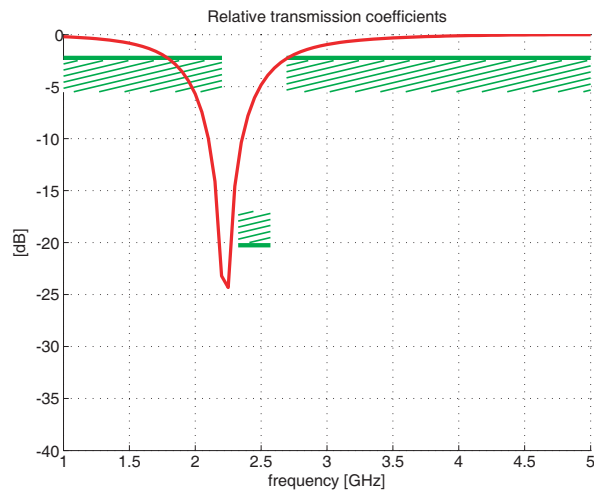


Figure 2: Transmission coefficients of the initial FSS Wi-Fi filter.

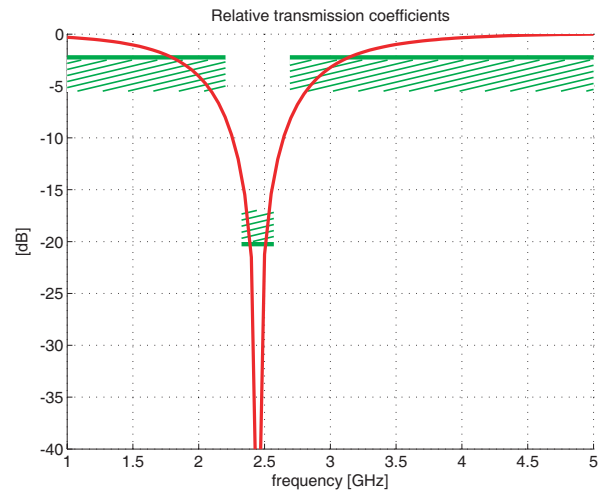


Figure 3: Transmission coefficients of the optimized FSS Wi-Fi filter.

- $k_1 = 0.900435$
- $k_2 = 0.403906$

Furthermore, from the design parameters we can compute the lengths l and l_e in the following way:

- $l = k_1 a$, $l = 0.038317111$ m (the total width and height of the Jerusalem-cross)
- $l_e = k_2 l$, $l_e = 0.015476511$ m (the total length of an outer arm)

6. CONCLUSION

A method of optimization of an FSS filter was proposed and tested on a problem of filtering of Wi-Fi signal. An initial proposed solution was optimized and corresponding transmission coefficients of optimized filter were shown.

The results presented in this paper are very promising. Our method could help to find solutions of other complicated electromagnetic problems. Anyway, further work in this direction should prove this theoretical study by results of real measurements.

ACKNOWLEDGMENT

This work was supported by the internal grant of Tomas Bata University in Zlin: IGA/FAI/2013/007.

REFERENCES

1. Munk, B., *Frequency Selective Surfaces — Theory and Design*, Wiley & Sons, New York, NY, USA, 2000.
2. Kiani, G. I., L. G. Olsson, A. Karlsson, K. P. Esselle, and M. Nilsson, “Cross-dipole bandpass frequency selective surface for energy-saving glass used in buildings,” *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 2, 520–525, 2011.
3. Rafique, U., M. M. Ahmed, M. A. Haq, and M. T. Rana, “Transmission of RF signals through energy efficient window using FSS,” *Proceedings of the 7th International Conference on Emerging Technologies*, 1–4, 2011.
4. Haupt, R. L., “Scattering from small salisbury screens,” *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 6, 1807–1810, 2006.
5. Knott, E. F. and C. D. Lunden, “The two-sheet capacitive jaumann absorber,” *IEEE Transactions on Antennas and Propagation*, Vol. 43, No. 11, 1339–1343, 1995.
6. Kiani, G. I., K. L. Ford, K. P. Esselle, A. R. Wiley, and C. L. Panagamuwa, “Oblique incidence performance of a novel frequency selective surface absorber,” *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 10, 2931–2934, 2007.

7. Rafique, U., G. I. Kiani, M. M. Ahmed, and S. Habib, “Frequency selective surface absorber for WLAN security,” *Proceedings of the 5th European Conference on Antennas and Propagation*, 872–875, 2011.
8. Thomas, K. G. and M. Sreenivasan, “A simple dual-band microstrip-fed printed antenna for WLAN applications,” *IET Microwaves, Antennas & Propagation*, Vol. 3, No. 4, 687–694, 2009, ISSN 1751-8725.
9. Chan, R. and R. Mittra, “Techniques for analyzing frequency selective surfaces: A review,” *Proceedings of the IEEE*, Vol. 76, No. 12, 1593–1615, 1988.
10. Wu, T. K., *Frequency Selective Surfaces and Grid Arrays*, Wiley & Sons, New York, NY, USA, 1995.
11. Wan, C. and J. A. Encinar, “Efficient computation of generalized scattering matrix for analyzing multilayered periodic structures,” *IEEE Transactions on Antennas and Propagation*, Vol. 43, No. 11, 1233–1242, 1995.
12. Gona, S. and V. Kresalek, “Development of a versatile planar periodic structure simulator in MATLAB,” *Conference on Microwave Techniques (COMITE)*, Vol. 14, No. 1, 2008.
13. D’Errico, J., “Fminsearchbnd, fminsearchcon — file exchange — MATLAB central [online],” <http://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd>, 2012-02-06 [cit. 2012-12-20].
14. The MathWorks Inc. Mathworks nordic, “Find minimum of constrained nonlinear multivariable function — MATLAB [online],” <http://www.mathworks.se/help/optim/ug/fmincon.html>, [cit. 2012-12-20].

Reflectance Spectroscopy of $\text{Ba}_{3+x}\text{Zn}_{1+y}\text{Nb}_2\text{O}_9$ Perovskites

W. De La Cruz¹, S. Moodie¹, J. Manson¹,
D. A. Crandles¹, D. Grebennikov^{2,3}, and P. Mascher²

¹Brock University, St. Catharines, ON, Canada

²McMaster University, Hamilton, ON, Canada

³Western University, London, ON, Canada

Abstract— Complex perovskites in the $\text{Ba}_{3+x}\text{Zn}_{1+y}\text{Nb}_2\text{O}_9$ family were studied via reflectance spectroscopy for photon energies between 0.006 and 1 eV. These materials are of interest as potential dielectric resonator materials which require large ϵ_1 (to enable device miniaturization), large $Q \approx \epsilon_1/\epsilon_2$ (for selectivity) and small temperature dependence of optical functions (for device stability). The dielectric functions were modeled by fitting the reflectance spectra to both Lorentz oscillator and factorized dielectric functions in order to get a sense of the uncertainty in extrapolating the measured far-infrared dielectric function to the microwave (MW) region (300 GHz). Both models suggest that for the stoichiometric composition $\epsilon_1 \approx 40$ while the extrapolated value of ϵ_2 has much more uncertainty. The extrapolated value of $Q \approx 2000$ at room temperature at MW frequency for $\text{Ba}_3\text{ZnNb}_2\text{O}_9$ and Q is largest near stoichiometric composition. ϵ_1 is only weakly composition dependent, except for the sample furthest from stoichiometric composition ($\text{Ba}_{2.7}\text{ZnNb}_2\text{O}_9$). Comparison of the present data for $\text{Ba}_3\text{ZnNb}_2\text{O}_9$ with previous work reveals that the phonon scattering rates and low frequency ϵ_2 are much higher in the present samples yielding lower Q values, which were prepared at somewhat higher temperature than previous workers. It is possible that microstructure — which depends strongly on sample preparation temperature — is influencing ϵ_2 in the microwave region more strongly than deviation from ideal stoichiometry $\text{Ba}_3\text{ZnNb}_2\text{O}_9$.

1. INTRODUCTION

The development of inexpensive dielectric materials for microwave resonators and filters is of continuing interest to the wireless communication industry. The performance criteria include large ϵ_1 (to enable device miniaturization), large $Q \approx \epsilon_1/\epsilon_2$ (for selectivity) and small temperature dependence of optical functions (for device stability) [1]. One of the families of possible advanced dielectric materials are the complex perovskites based on $\text{Ba}(\text{B}'_{1/3}\text{B}''_{2/3})\text{O}_3$ where $\text{B}' = \text{Mg}, \text{Co}, \text{or Zn}$ and $\text{B}'' = \text{Nb or Ta}$. In particular there has been recent interest in perovskites based on $\text{Ba}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3$ (BZN) because of comparatively cheaper cost of niobium as compared to tantalum ore. It has been shown that, depending on processing temperature, sintering off-stoichiometric mixtures of starting materials can produce a greater amount of the 1:2 ordered phase, potentially increasing the Q value, while introducing secondary phases into the ceramics [2–5].

2. EXPERIMENTAL DETAILS

The samples were made in a two step process [2]. First, columbites ($\text{Zn}_{1+y}\text{Nb}_2\text{O}_6$) were prepared by combining the appropriate amount of commercially available ZnO and Nb_2O_5 (99.9%) and then calcining at 1000°C. In the next stage, the appropriate amount of BaCO_3 was mixed with the columbite and the powder pressed into 12.5 mm diameter disks, buried in powder of the same composition (to minimize Zn loss) and then sintered in air at 1440°C for 8 h. The phase composition of the samples was checked using x-ray diffraction (XRD) and scanning electron microscopy (SEM) [2]. The ceramic disks were polished to optical smoothness (final polishing grit 1 μm). The absolute reflectance of the samples was determined near normal incidence using the *in situ* gold evaporation technique [6].

3. RESULTS AND DISCUSSION

Figure 1 compares the room temperature reflectance of the BZN prepared in this work with the literature data [7]. Observe that there are 11 peaks and/or shoulders observable in the spectrum whereas Kamba et al. used 14 TO modes in a factorized model of the dielectric function, 7 of which had $\omega_{\text{TO}} \leq 182 \text{ cm}^{-1}$. In the present work there appear to be only 4 modes in the same frequency range. It is possible that the discrepancy is due to lower measurement resolution. Group theory predicts 16 infrared active modes for the $P\bar{m}n1$ structure adopted by 1:2 ordered BZN [7].

The lower number of modes can be explained by overlapping modes and/or low mode oscillator strengths. In Fig. 1 note that the reflectance peaks of the present sample are slightly lower, the edges of the bands are more rounded and the minima are not as deep as in the literature data. This is an indication that the phonon modes are more strongly scattered in the present material which is associated with higher dielectric loss ϵ_2 which could be due to the different methods of preparation.

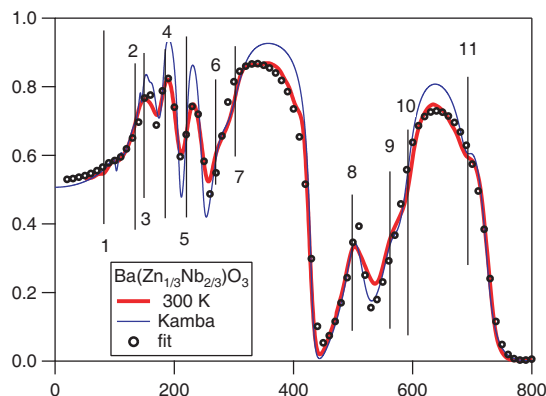


Figure 1: Room temperature reflectance of BZN compared with literature data (Ref. [7]) and a sum oscillator fit.

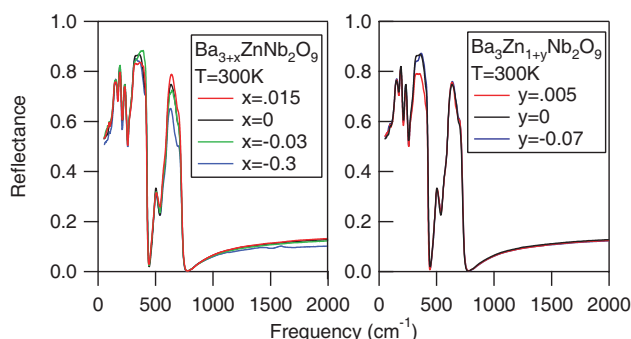


Figure 2: Room temperature reflectance of $\text{Ba}_{3+x}\text{Zn}_{1+y}\text{Nb}_2\text{O}_9$.

Figure 2 presents the raw data over most of the measured frequency range. The reflectance for $\omega \geq 2000 \text{ cm}^{-1}$ is essentially constant for all materials. Note that the peaks and shoulders for all of the samples are in essentially the identical positions. Thus it is interesting to compare the infrared reflectance (IR) spectra with both x-ray diffraction spectra taken on these samples [2] and the IR reflectance spectrum of $\text{Ba}_3\text{MgNb}_2\text{O}_9$ (BMN) [9] which has a similar Goldschmidt tolerance factor [1]. There are no superlattice peaks associated with 1:2 order in the x-ray diffraction spectrum of $\text{Ba}_3\text{Zn}_{0.93}\text{Nb}_2\text{O}_9$ whatsoever. Yet there is more structure in the IR reflectance spectrum of $\text{Ba}_3\text{Zn}_{0.93}\text{Nb}_2\text{O}_9$ than that of BMN, where it was shown the short-range 1:2 order exists [9]. It has been demonstrated [9] that IR reflectance spectroscopy can be more sensitive to 1:2 ordering if the ordering is short-range. It is also interesting to note that the samples here were sintered at 1450°C . Note that there is some disagreement in the order-disorder transition temperature with some authors placing it at $\approx 1375^\circ\text{C}$ [10], while others [11] place it between 1400 and 1500°C . The XRD patterns indicate that small amounts of Ba vacancy promotes long-range 1:2 order, a large amount of either Ba or Zn vacancies cause the superlattice peaks associated with 1:2 order to disappear [2]. However, the present IR spectra suggest that 1:2 order exists in all of the samples, and hence must be at short-range in those samples where XRD superlattice reflections cannot be seen. XRD spectra of $\text{Ba}_3\text{Zn}_{0.93}\text{Nb}_2\text{O}_9$ indicated the presence of Zn-poor phases $\text{Ba}_5\text{Nb}_4\text{O}_{15}$ and $\text{Ba}_8\text{ZnNb}_6\text{O}_{24}$. However there are no extra peaks in the reflectance spectrum of this sample.

The sample with the most different spectrum is that of grossly non-stoichiometric $\text{Ba}_{2.7}\text{ZnNb}_2\text{O}_9$. It presents the only unique feature ($\approx 1600 \text{ cm}^{-1}$) in the infrared reflectance spectrum which is most

likely associated with $\text{Ba}_6\text{ZnNb}_9\text{O}_{30}$ which XRD detects in sizeable amounts [2]. In addition the high frequency dielectric constant is lower.

In order to estimate the dielectric properties of these materials, the reflectance was fit using the both the classical Lorentz model and the factorized form of the dielectric function [7]. For lack of space we discuss only the classical fits here. Only six of the 11 peaks/shoulders gave robust (i.e., independent of the starting values) Lorentz oscillator parameters (ω_{oj} -position, γ_j -scattering rate, ω_{pj} , strength). It was decided to use seven oscillators (I-VII) and it was necessary to fix the parameters of the lowest frequency oscillator in order to maintain any feature in the fitted reflectance spectrum near 100 cm^{-1} . The final parameters for BZN are listed in Table 1. And a sample fit for BZN is shown in Fig. 1.

$$\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega) = \epsilon_\infty + \sum_{j=I}^{VII} \frac{\omega_{pj}^2}{\omega_{oj}^2 - \omega^2 - i\omega\gamma_j} \quad (1)$$

Table 1: Parameters found for BZN using the sum model of the dielectric function. Only seven oscillators were used; several closely spaced shoulders were represented by a single oscillator in some cases. $\epsilon_\infty \approx 4.9$. Peak I parameters were fixed during the least squares fitting. The peak and/or shoulder numbers refer to Fig. 1. The units for ω_{oj} , γ_j and ω_{pj} are cm^{-1} .

j	Peak and/or shoulders	$\omega_{oj} \pm 0.2$	$\gamma_j \pm 0.5$	$\omega_{pj} \pm 5$
I	1	102	20	100
II	2, 3	153.2	13.8	588
III	4	179.2	16.2	538
IV	5	223.3	22.4	438
V	6, 7	284.5	29.3	625
VI	8	503.6	19.8	302
VII	9, 10, 11	588.0	34.3	632

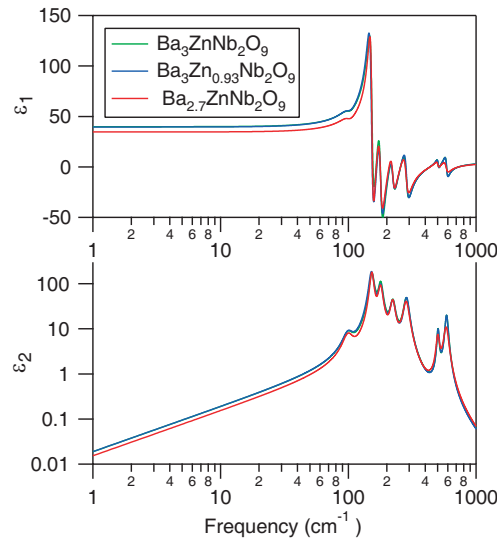


Figure 3: Room temperature sum model dielectric functions for $\text{Ba}_{3+x}\text{Zn}_{1+y}\text{Nb}_2\text{O}_9$.

The real and imaginary parts of the dielectric function determined by these fits are shown in Fig. 3. Note that the low frequency limit of the real part of the dielectric function of all the materials except $\text{Ba}_{2.7}\text{ZnNb}_2\text{O}_9$ is ≈ 40 in agreement with the literature [3, 7, 10]. The $Q \approx \frac{\epsilon_1}{\epsilon_2}$ values at $0.1\text{ cm}^{-1} = 3\text{ GHz}$ obtained from the extrapolated dielectric functions are admittedly very low ($\approx 40/0.02 = 2000$), but there is huge uncertainty (a factor of 10 at least). Lower frequency data are

necessary to measure the low frequency quality factor. In terms of trends, Belous et al. measured a maximum Q value for $y = 0$ in the $\text{Ba}_3\text{Zn}_{1+y}\text{Nb}_2\text{O}_9$ series and that small amounts of Ba-vacancies improved the Q -value in the $\text{Ba}_{3+x}\text{ZnNb}_2\text{O}_9$ series. The lack of lower frequency IR data makes it difficult to confirm these findings.

4. CONCLUSION

The IR data confirm the presence of short-range 1:2 order in all of the prepared ceramics in the $\text{Ba}_{3+x}\text{Zn}_{1+y}\text{Nb}_2\text{O}_9$ system, despite sintering at high temperatures 1440° . The far-infrared dielectric properties are not particularly sensitive to stoichiometry except for grossly non-stoichiometric $\text{Ba}_{2.7}\text{ZnNb}_2\text{O}_9$.

ACKNOWLEDGMENT

The authors acknowledge financial support from the National Science and Engineering Research Council of Canada (NSERC).

REFERENCES

1. Cava, R. J., "Dielectric materials for applications in microwave communications," *J. Mater. Chem.*, Vol. 11, 54–62, 2001.
2. Grebbernikov, D. and P. Mascher, "Structural properties of near-stoichiometric composition of $\text{Ba}(\text{B}'_{1/3}\text{B}''_{2/3})\text{O}_3$ ($\text{B}' = \text{Mg, Co, or Zn}$ and $\text{B}'' = \text{Nb or Ta}$) perovskites," *J. Mater. Res.*, Vol. 26, 1116–1125, 2011.
3. Belous, A. G., O. V. Ovchar, O. V. Kramarenko, J. Bezjak, B. Jancar, D. Suvorov, and G. Annino, "Low-loss microwave ceramics based on non-stoichiometric perovskites $\text{Ba}(\text{Co}_{1/3}\text{Nb}_{2/3})\text{O}_3$ and $\text{Ba}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3$," *Ferroelectrics*, Vol. 367, 149–162, 2008.
4. Wu, H. and P. K. Davies, "Influence of non-stoichiometry on the structure and properties of $\text{Ba}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3$ microwave dielectrics. II. Compositional Variations in Pure BZN," *J. Am. Ceram. Soc.*, Vol. 89, 2250–2263, 2006.
5. Wu, H. and P. K. Davies, "Influence of non-stoichiometry on the structure and properties of $\text{Ba}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3$ microwave dielectrics. IV. Tuning τ_f and the part size dependence of $Q \times f$," *J. Am. Ceram. Soc.*, Vol. 89, 2271–2278, 2006.
6. Homes, C. C., M. Reedyk, D. A. Crandles, and T. Timusk, "Technique for measuring the reflectance of irregular, submillimeter-sized samples," *Applied Optics*, Vol. 32, 2976–2983, 1993.
7. Kamba, S., H. Hughes, D. Noujni, S. Surendran, R. C. Pullar, P. Samoukhina, J. Petzelt, R. Freer, N. McNalford, and D. M. Idles, "Relationship between microwave and lattice vibration properties in $\text{Ba}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3$ -based microwave dielectric ceramics," *J. Phys D: Appl. Phys.*, Vol. 27, 1980–1986, 2004.
8. Chen, Y., H. Cheng, H. Liu, and C. Chia, "Correlation of microwave dielectric properties and normal vibration modes of $x\text{Ba}(\text{Mg}_{1/3}\text{Ta}_{2/3})\text{O}_3 - (1-x)\text{Ba}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3$ ceramics: Infrared spectroscopy," *J. Appl. Phys.*, Vol. 94, 3365–3370, 2003.
9. Reaney, I. M., J. Petzelt, V. V. Voitsekhovshii, F. Chu, and N. Setter, "B site order and infrared reflectivity in $\text{A}(\text{B}'\text{B}'')\text{O}_3$ complex perovskites," *J. Appl. Phys.*, Vol. 76, 2086–2092, 1994.
10. Lee, C.-T., Y. C. Lin, C. Y. Huang, C. Y. Su, and C. L. Hu, "Cation ordering and dielectric characteristics in Barium Zinc Niobate," *J. Am. Cer. Soc.*, Vol. 90, 483–489, 2007.
11. Yoshioka, H., "Ordering of cations in $\text{Ba}(\text{Mg}_{1/3}\text{Nb}_{2/3})\text{O}_3$ and $\text{Ba}(\text{Zn}_{1/3}\text{Nb}_{2/3})\text{O}_3$," *Bull. Chem. Soc. Jpn.*, Vol. 60, 3433–3434, 1987.

Modular Antenna Array Concept for Millimeter-wave Beam-steering Applications

U. Johannsen, M. I. Kazim, A. B. Smolders, and M. H. A. J. Herben
Eindhoven University of Technology, the Netherlands

Abstract— Steerable antenna arrays are of particular importance for consumer electronics applications in the millimeter-wave band. However, the exact array size depends on the specific application scenario. Therefore, a modular array concept is proposed in this paper that meets the cost requirements of the consumer electronics market. Moreover, since circular polarization is desired due to the line-of-sight nature of the communication channel at those frequencies, special attention is placed on the construction of circularly-polarized antenna arrays. Finally, a deflector design is proposed that allows to extend the scan-range of the proposed modular antenna array.

1. INTRODUCTION

The release of the unlicensed 60 GHz band (57 GHz–66 GHz in Europe) tackles the increasing demand for more bandwidth in the wireless consumer electronics market. In order to meet the cost requirements of this market, many research groups have worked towards 60 GHz front-end electronics in mainstream silicon technologies, i.e., CMOS and BiCMOS. Moreover, there is a wide consensus that the antenna should be integrated in the same package with the front-end integrated circuit (IC). A single antenna element, however, does not satisfy the gain requirements of many 60 GHz applications, e.g., for high definition (HD) video streaming with a communication distance of up to 10 m. Therefore, antenna arrays are of particular interest in the millimeter-wave (mm-wave) band since they offer sufficiently high antenna gain. For this, arrays with a fixed amplitude and phase distribution are often considered in the literature. Especially for mobile applications, however, the antenna array must exhibit the capability to align its narrow beam in the direction of the strongest signal. Hence, an antenna array with beam-steering capability is imperative, see [1–5].

In addition to the required high gain, circular polarization is seen to be a key requirement for 60 GHz wireless applications. The use of circular polarization can significantly increase the robustness of a wireless communication link due to the line-of-sight nature of the communication channel at mm-waves. Otherwise, linearly-polarized solutions may require an accurate polarization alignment, which is not very practical to implement for portable applications.

With respect to beam-steerable antenna arrays, the authors of [4] propose to integrate all array electronics in one single chip in order to minimize the total required chip area as well as the differentiation between antenna branch properties. A complete 60 GHz receiver module with integrated antenna array that uses this kind of approach was published in [6]. Such a solution, however, is not very flexible since the number of antenna elements and their arrangement is fixed. As a result, the range of suitable applications is limited, which ultimately results in high costs for the module due to low production volumes. In order to achieve a high flexibility that can target a larger application range and, therefore, lower the costs by increasing the production volume, a modular array concept is proposed here. This concept is suitable for creating linearly-polarized arrays, as briefly illustrated in Section 2.1, as well as for creating circularly-polarized arrays, as outlined in more detail in Section 2.2. In both cases, the mm-wave module from [7] was used as basic array element. In order to increase the scan-range of the circularly-polarized array, the deflector design from [8] is proposed, which is described in Section 3. Finally, the findings of this approach are discussed in Section 4.

2. MODULAR ARRAY CONCEPT

A conceptual drawing of the proposed modular array approach is provided in Figure 1(a). It consists of mm-wave front-end modules like the one shown in Figure 1(b), which is based on the design presented in [7]. The desired overall array size and antenna element arrangement is then composed out of several such modules. As shown in Figure 1(a), a low-frequency reference oscillator for phase synchronization can be supplied to the modules from an external source using a distribution network. On-chip phase shifters and variable gain amplifiers can then be used to apply the desired amplitude and phase distribution to each individual antenna element. As for

the reference oscillator signal, the communication path can be distributed to the modules in IF, base-band, or even digital domain.

Two examples of possible 2×2 array configurations with the module from Figure 1(b) are depicted in Figure 2, which will be further discussed in the following sections.

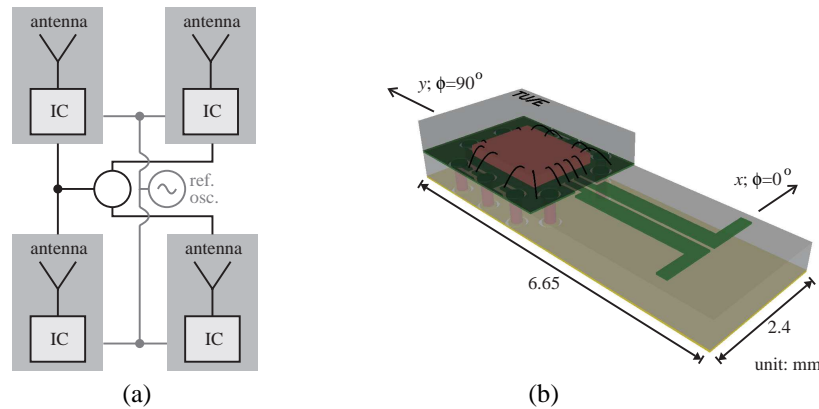


Figure 1: Conceptual drawings of the modular array approach and the proposed AiP design. (a) Modular array approach. (b) Conceptual drawing of the AiP design for use in a modular array.

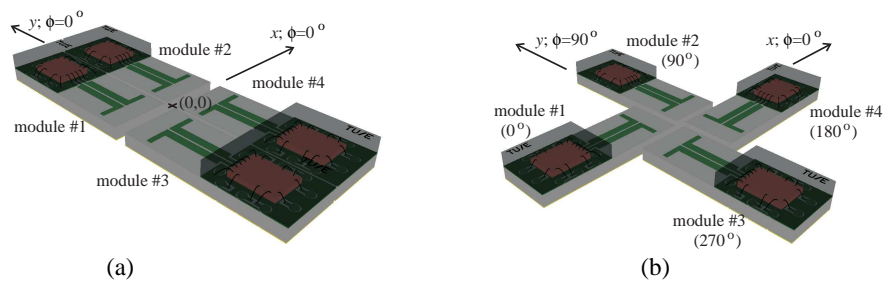


Figure 2: Examples of 2×2 antenna arrays consisting of integrated mm-wave front-end modules as shown in Figure 1(b). (a) Linearly-polarized array. (b) Circularly-polarized array.

2.1. Linearly-polarized Array

As a first example, a linearly-polarized array is demonstrated, which can be constructed by arranging several mm-wave modules from Figure 1(b) as shown in Figure 2(a). It shows a 2×2 array with a spacing of $\lambda_0/2$ between the printed dipoles, where λ_0 is the free space wavelength at 60 GHz, i.e., 5 mm. If a time delay is now applied to, e.g., modules #1 and #3, the main beam can be steered in the $\phi = 0^\circ$ plane.

If a larger array is required, it can be built in the same manner by simply using more integrated mm-wave modules. Here, the particular module choice in this example exhibits a limitation in this respect. Its size does namely not permit a $\lambda_0/2$ spacing in y -direction, see Figure 1(b). However, a $2 \times N$ -array can easily be constructed, with N being any number of elements in x -direction. By this, a high gain fan-beam can be generated that exhibits, for example, a large coverage in elevation while being steerable in azimuth, or vice versa.

2.2. Circularly-polarized Array

As another example, the flexibility of the modular array approach is used to construct a circularly-polarized array with the same modules. An elegant way to create circular polarization with linearly-polarized elements in an array environment is by using the sequential rotation technique as introduced by John Huang in 1986 [9]. Figure 2(b) shows the basic antenna arrangement for this method. Here, circularly-polarized (CP) radiation is achieved by applying proper phase shifts to sequentially rotated linearly-polarized antenna elements, as also illustrated in the figure. Using this configuration for a transmitting array, a right-hand circular polarization is achieved. The resulting active input reflection coefficients of the sequentially rotated integrated antennas with an antenna spacing of $\lambda_0/2$ and for simultaneous excitation with the same amplitude for each antenna element are shown in Figure 3(a).

The overall gain pattern for the $\phi = 0^\circ$ and $\phi = 90^\circ$ planes are provided in Figure 3(b). The achieved maximum gain was computed to range from 9.9 dBi to 10.6 dBi with a 3 dB beam-width of 42° at 60 GHz. Furthermore, the axial ratio (AR) at 60 GHz of this circularly-polarized antenna array in the $\phi = 0^\circ$ and $\phi = 90^\circ$ planes is shown in Figure 3(c). The AR is defined as

$$AR = \frac{|E_\theta \vec{u}_\theta + E_\phi \vec{u}_\phi|_{\max}}{|E_\theta \vec{u}_\theta + E_\phi \vec{u}_\phi|_{\min}}, \quad (1)$$

with \vec{u}_θ the unit vector in θ -direction, \vec{u}_ϕ the unit vector in ϕ -direction, and E_θ and E_ϕ the electric field in θ - and ϕ -direction, respectively. It is evident from Figure 3(c) that for $\theta = 0^\circ$ practically perfect circular polarization is achieved. The AR exceeds 3 dB for angles beyond $\pm 19^\circ$ and, hence, nearly the full 3 dB beam-width exhibits an acceptable degree of circular polarization.

When the beam is steered away from the broad-side direction ($\theta_0 \neq 0^\circ$), however, the axial ratio at the desired scan angle degrades, as can be seen from Figure 4. It shows the gain pattern in the $\phi = 0^\circ$ plane (Figure 4(b)) when the beam is steered towards $\theta_0 = -15^\circ$. All elements are still well matched (Figure 4(a)) and a maximum gain of 9.2 dBi to 9.9 dBi is achieved. The simulated axial ratio at the scan angle, however, is already degraded to 1 dB, see Figure 4(c), which is basically a result of the unequal gain patterns of the single mm-wave module in its $\phi = 0^\circ$ and $\phi = 90^\circ$ planes. Hence, for even larger scan angles, this degradation can be expected to increase. A possible method to improve the axial ratio over a wide scan range can be found in [10].

However, the scan-range of such a solution is limited, e.g., by the allowed maximum side-lobe-level. In order to extend the scan range of the antenna array, a deflector design similar to the one presented in [8] can be used. The proposed design and basic operation principle of such a deflector

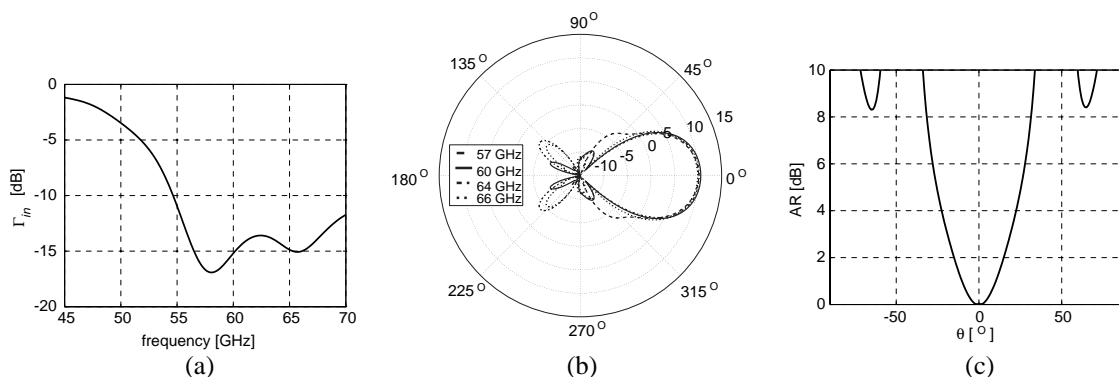


Figure 3: Antenna parameters of the antenna array shown in Figure 2(b) for simultaneous excitation of all four elements with the same amplitude and a phase-delay distribution as shown in the figure. (a) Active input reflection coefficient with respect to $Z_0 = 100 \Omega$. (b) Gain pattern at 60 GHz for the $\phi = 0^\circ$ and $\phi = 90^\circ$ planes when steering the beam towards $\theta = 0^\circ$. (c) Axial ratio at 60 GHz for the $\phi = 0^\circ$ and $\phi = 90^\circ$ planes.

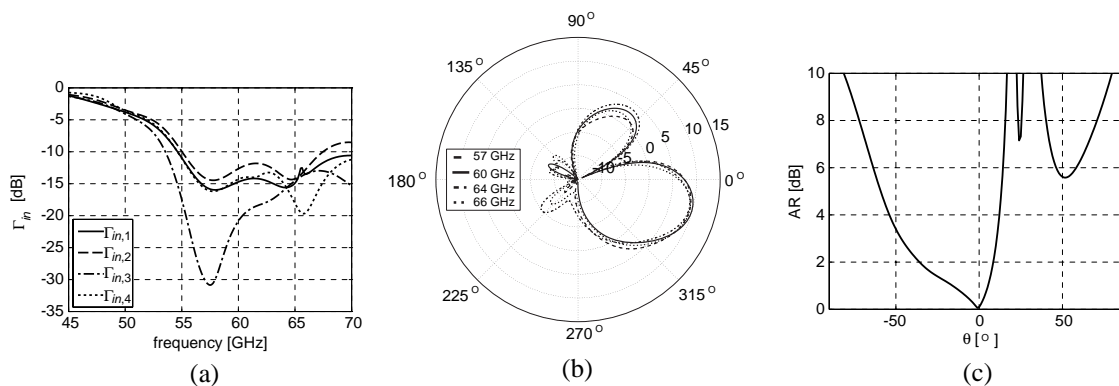


Figure 4: Antenna parameters of the antenna array shown in Figure 2(b) for a scan angle of -15° in the $\phi = 0^\circ$ plane. (a) Active input reflection coefficient with respect to $Z_0 = 100 \Omega$. (b) Gain pattern at $\phi = 0^\circ$. (c) Axial ratio at 60 GHz for the $\phi = 0^\circ$ plane.

is described in the next section.

3. SCAN-RANGE EXTENSION BY A SPHERICAL DEFLECTOR

In order to extend the scan-range of the afore-mentioned CP antenna array, a 3D spherical passive electromagnetic deflector is proposed here, which is shown in Figure 5(a). It uses a spherical surface consisting of passive circularly-polarized deflecting elements, i.e., with fixed phase-shift. In principle, by defining fixed phase-shifts on the spherical deflector surface, a beam can be produced that exhibits a modified scan-angle (θ_S) relative to the scan-angle of the emitted beam from the source array (θ_P). By this, the scan-range of the planar array is extended while the minimum link-budget requirement is still satisfied.

Since the suggested deflector configuration is very large in terms of wavelength, it is difficult to analyze it with commercially-available 3D electromagnetic tools due to the computationally-intensive nature of the problem. A generalized scalar analytical formulation of the parametric deflector configuration with the required phase-shift determination was, therefore, implemented in MATLAB. The results obtained from this model for a lossless concentric circular array (CCA) source and a hemispherical deflector at 60 GHz in the $\phi = 0^\circ$ -plane are shown in Figures 5(b)–(c). The design challenges involved in the realization of a 60 GHz CP deflector element that can achieve the maximum phase-shift range with minimal losses over the whole band of operation have been investigated. The results of this proposed solution, based on the principle presented in [11], will be published shortly.

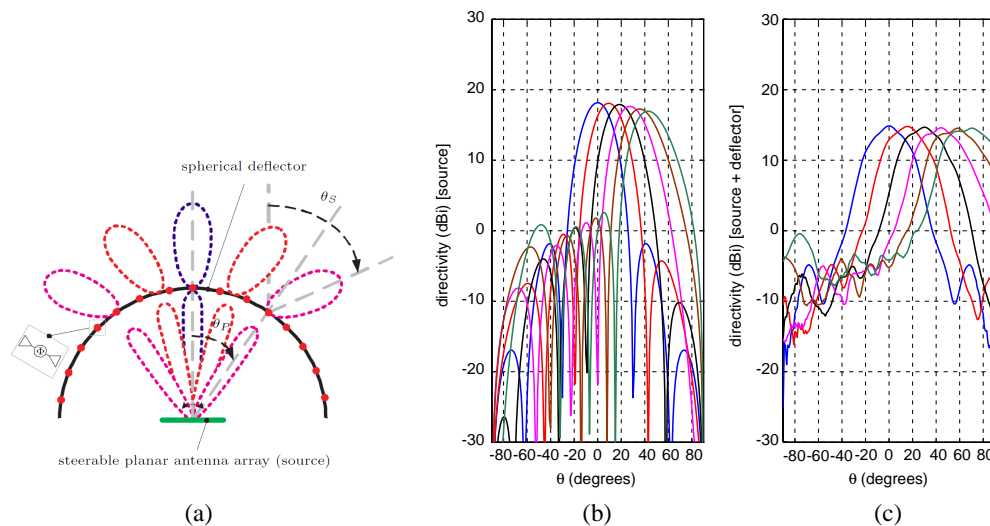


Figure 5: Scan-range extension of a CP source antenna array by a spherical deflector. (a) Proposed topology of 3D spherical passive electromagnetic deflector. Directivity patterns when a source antenna array is scanned to 0° (blue), 10° (red), 20° (black), 30° (magenta), 40° (brown) & 50° (dark-green); (b) source only, (c) source-spherical deflector configuration

4. CONCLUSIONS AND RECOMMENDATIONS

In this paper a flexible antenna array concept is presented that tackles the stringent cost requirements of the consumer electronics market by a modular approach. The first simulation results are promising, but still need to be verified by experiments. Moreover, different array configurations should be investigated, in order to increase the number of antenna elements and, therefore, achieve a larger gain with the presented front-end modules.

Furthermore proposed in this paper is the use of a spherical deflector in order to increase the scan-range of the source array. Also for this design an experimental validation remains to be conducted.

Therefore, a full-scale demonstrator consisting of modular antenna array and deflector is currently under planning.

ACKNOWLEDGMENT

This work was funded by the European Catrene project PANAMA, the FP7 project Par4CR, as well as IOP-GenCom's Sigi-Spot project.

REFERENCES

1. Yong, S. K. and C.-C. Chong, "An overview of multigigabit wireless through millimeter wave technology: Potentials and technical challenges," *EURASIP Journal on Wireless Communications and Networking*, Vol. 2007, No. 1, 50–50, 2007.
2. Guo, N., R. C. Qiu, S. S. Mo, and K. Takahashi, "60-GHz millimeter-wave radio: Principle, technology, and new results," *EURASIP Journal on Wireless Communications and Networking*, Vol. 2007, No. 1, 2007.
3. Daniels, R., J. Murdock, T. Rappaport, and R. Heath, "60 GHz wireless: Up close and personal," *IEEE Microwave Magazine*, Vol. 11, No. 7, 44–50, Dec. 2010.
4. Smolders, P., H. Yang, and I. Akkermans, "On the design of low-cost 60-GHz radios for multigigabit-per-second transmission over short distances [topics in radio communications]," *IEEE Communications Magazine*, Vol. 45, No. 12, 44–51, Dec. 2007.
5. Daniels, R. and R. Heath, "60 GHz wireless communications: Emerging requirements and design recommendations," *IEEE Vehicular Technology Magazine*, Vol. 2, No. 3, 41–50, Sep. 2007.
6. Liu, D., I. Akkermans, H.-C. Chen, and B. Floyd, "Packages with integrated 60-GHz aperture-coupled patch antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 10, 3607–3616, 2011.
7. Johannsen, U., A. B. Smolders, A. C. F. Reniers, A. R. V. Dommele, and M. D. Huang, "Integrated antenna concept for millimeter-wave front-end modules in proven technologies," *6th European Conference on Antennas and Propagation (EUCAP)*, 2560–2563, 2012.
8. Kazim, M. I. and M. H. A. J. Herben, "Trade-offs in multifaceted passive electromagnetic deflector for the 60 GHz frequency band," *5th European Conference on Antennas and Propagation (EUCAP)*, 1044–1048, Rome, Italy, Apr. 2011.
9. Huang, J., "A technique for an array to generate circular polarization with linearly polarized elements," *IEEE Transactions on Antennas and Propagation*, Vol. 34, No. 9, 1113–1124, Sep. 1986.
10. Smolders, A. B. and U. Johannsen, "Axial ratio enhancement for circularly-polarized millimeter-wave phased-arrays using a sequential rotation technique," *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 9, 3465–3469, Sep. 2011.
11. Fox, A., "An adjustable wave-guide phase changer," *Proceedings of the IRE*, Vol. 35, No. 12, 1489–1498, Dec. 1947.

Analysis and Design of Multi-mode Dielectric Waveguide Interconnect with Planar Excitation

Nemat Dolatsha^{1,2} and Amin Arbabian³

¹EE Department, ETHZ University, Zurich, Switzerland

²EE Department, Stanford University, Stanford, CA, USA

³Center of Integrated Systems (CIS), EE Department, Stanford University, Stanford, CA, USA

Abstract— Design of an all-electrical and easy to package wideband chip-to-chip solution on a multi-mode dielectric waveguide is discussed. Different parameters such as bandwidth-per-pitch, range, as well as crosstalk in aggregated lines are analyzed. Using a Rogers RO3006 material, a bandwidth-per-pitch of 16 GHz/mm (and an absolute bandwidth-per-line of 2×40 GHz) at the center frequency of 100 GHz is achieved while maintaining a range of 1 m with crosstalk below 15 dB. The sensitivity of the line to the curvature is also examined. The signal is coupled from the silicon chip to the degenerate fundamental and polarization-orthogonal E_x^{11} and E_y^{11} waveguide modes using planar electric and slot dipole antennas, respectively. The large available bandwidth will be channelized in frequency for optimal overall efficiency and throughput with a CMOS transceiver. The performance sensitivity of the structure to possible fabrication imperfections is examined and discussed. The proposed waveguide offers a solution for Tera bit-per-second (Tbps) fully-electrical wireline links.

1. INTRODUCTION

There is a large demand to substantially increase overall system bandwidth as well as connectivity at all levels- from individual ICs to backplanes. As aggregate link data rate requirements are approaching Tbps levels, current physical channels are nearing their limits. To this end, a power and data-rate “wall” is emerging on the horizon. This is in terms of maximum “data flux density” (bandwidth per unit width) and link efficiency (in terms of energy per bit) when the bandwidth of conventional planar on-board transmission lines is pushed to tens of GHz.

In assessing communication channels, several figures of merit exist: bandwidth and data-rate (considering width and length of the channel), overall energy efficiency, and the associated cost (e.g., with extra components or need for precision/complex packaging). In improving data bandwidth, main culprits are the high-loss, dispersion, crosstalk and reflections in the conventional planar TEM electrical channel. In these lines, the capacity of the link is primarily limited by skin effect and impedance mismatches [1]. For medium to large distances and/or to achieve high data rates, the cost of equalization in electrical lines becomes excessive.

Optical fiber links have been extensively used in long-haul communication and are now also slowly making their way into the short-range arena. However, the associated power and cost overhead (due to precision packaging and the addition of electro-optical modules) pose limitations on adaptability for ubiquitous board-level interconnects.

In this paper, we propose a structure that addresses these concerns. An all-electrical, easy to package and planar structure on a dielectric waveguide is presented and different aspects (as an interconnect) are analyzed. This has applications for chip-to-chip interconnects as well as for short active cables. Dielectric waveguides made of simple dielectric strips are known for very low transmission loss at mm-wave and sub-THz frequencies [2, 3]. The challenge lies in designing an efficient, wideband and low-cost coupling structure that makes appropriate use of the waveguide modes. Here, we propose a novel planar feed structure to excite two *polarization-orthogonal* modes of the waveguide, namely the E_x^{11} and E_y^{11} modes. This doubles the applicable bandwidth of the dielectric waveguide without compromising performance. Planar electric dipole and slot dipole antennas can launch the E_x^{11} and E_y^{11} modes, respectively. Excitation with planar feed antennas simplifies the assembly and enables the use of conventional packaging with CMOS transceivers. Fig. 1 shows a conceptual schematic of the proposed structure. As depicted in Fig. 1(b), the electric field orientations of the first two modes are perpendicular to each other.

2. DIELECTRIC WAVEGUIDE DESIGN

Dielectric waveguide for data communication has been proposed in [4, 5]. In [5], due to the selected structure and modes, the implementation presents lower coupling efficiency to the waveguide

as well as strong waveguide mode-coupling and corresponding length dependencies that affect robustness. Mode leakage due to discontinuities may also pose limitations. Our approach focuses on fundamental and polarization-orthogonal degenerate modes where the mode-coupling to higher order modes is greatly suppressed and the coupling efficiency (from the chip to the waveguide) is considerably better. It should be noted that another communication channel based on substrate-integrated waveguide (SIW) is reported in [6]. Due to operating at lower mm-wave frequencies, the bandwidth of the structure is limited.

As shown in Fig. 1, the proposed waveguide is simply made of a dielectric strip, dimensions of which are primarily determined by the operating frequency band of the first two fundamental E_x^{11} and E_y^{11} modes. In addition, limiting the cross section avoids mode cross coupling along the waveguide by eliminating higher order modes. The field confinement in the cross section of the waveguide is related to the dielectric strip permittivity as well as waveguide dimensions. Higher permittivity leads to an increased field concentration within the waveguide core, as opposed to fields on the outside of the strip, allowing for a smaller pitch.

In order to generalize the argument, the available bandwidth-per-pitch and the maximum length of waveguides with high and low permittivities are approximated for different frequency bands. It should be noted that dielectric materials applicable for millimeter-waves usually have a permittivity in the range of 2 to 12. In [7] the dispersion diagram of square rectangular dielectric waveguides made of materials with permittivity of 12 and 2.3 are extracted in terms of a normalized frequency parameter ($v = k_o a \sqrt{\epsilon_r - 1}$). The values of v at cut-off frequencies are known. Therefore the bandwidth can be calculated as a function of center operating frequency:

$$BW = \frac{K_c}{2\pi a \sqrt{\epsilon_r - 1}} \quad (1)$$

$$a = \frac{4.5c}{2\pi f_o \sqrt{\epsilon_r - 1}} \quad (2)$$

where c , f_o , and a are the free space wave velocity, center frequency, and waveguide cross section, respectively. K is 1.5 and 2, respectively, for dielectrics with permittivity of 12 and 2.3. Assuming a pitch of $2a$ for the waveguide, the bandwidth-per-pitch parameter is plotted in Fig. 2. This normalized bandwidth for a single mode is proportional to f_o^2 . In addition, with increasing the permittivity the bandwidth-per-pitch increases. This is mainly due to the increasing field confinement. On the other hand, the maximum operating frequency is restricted by the attenuation of the waveguide. An approximate relation for the attenuation constant of a dielectric waveguide is presented in [8]. Assuming a maximum acceptable transmission loss of 50 dB for the link, the maximum length of the waveguide is calculated for dielectric loss tangent of 0.001 and plotted in Fig. 2. As the frequency increases, due to higher attenuation, the maximum available length (L_{\max}) decreases. In addition, for the same loss tangent, the attenuation per unit length increases with higher permittivity. It should be noted that dielectrics with higher permittivity usually also present a higher loss tangent. From this analysis, the required length of the link determines the center frequency, which in turn limits the maximum bandwidth of the link. Utilizing repeaters can partially alleviate this tradeoff.

In a multiplexed lines scenario, the minimum pitch of single wires is determined by the maximum

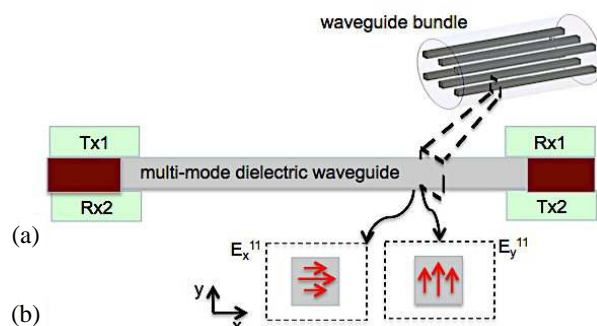


Figure 1: Conceptual schematic of the proposed multi-mode dielectric waveguide for chip-to-chip interconnects or active cables.

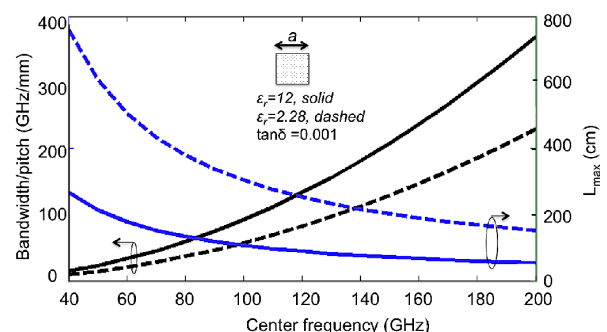


Figure 2: Bandwidth-per-pitch and maximum length of the link (with acceptable transmission loss of 50 dB) in terms of operating frequency.

acceptable crosstalk, which is related to the length of the link as well as the field expansion around the waveguide. Fig. 3 compares the expansion of E_x component of E_x^{11} mode in the cross section of waveguides at the lower edge of the operating frequency band (with lower field confinement) for different permittivities. The discontinuity in the E_x along x -axis is due to discontinuity in permittivity seen by normal electric field. The E_y^{11} mode field distributions are dual of the E_x^{11} mode. As observable, fields of waveguide with lower permittivity propagate with a relatively lower confinement in the cross section. This limits the minimum distance between possible parallel “wires” in a multiplexed link. In order to quantify this argument, the maximum range of a link with a crosstalk of less than -15 dB is calculated and shown in Fig. 4. As expected, for the same length, links composed of waveguides with lower permittivity need a larger pitch.

Taking this analysis into account, a rectangular strip waveguide made of Rogers RO3006 dielectric material (loss tangent 0.002) with a medium permittivity of 6.15 is constructed and designed for a 1m link. Assuming an acceptable transmission loss of 50 dB (including waveguide attenuation and mode excitation losses) the appropriate operating frequency is calculated. Simulations show when operating at frequencies around 100 GHz (waveguide cross section: 0.85×0.85 mm²) an attenuation of 0.044 dB/mm (44 dB/m) is achievable. This waveguide supports a bandwidth of approximately 2×40 GHz each from 80 GHz to 120 GHz. These waveguides can then be bundled together for a larger aggregate throughput. Returning to Fig. 4, to have a minimum crosstalk of 15 dB a pitch of larger than $1.65\lambda_0$ (λ_0 is the free-space wavelength at the center operating frequency) is required. Therefore, a bandwidth-per-pitch of 2×16 GHz/mm is achievable.

The sensitivity of transmission and mode cross coupling to the bending of the “wire” (see Fig. 5) is also investigated. Simulations show an extra loss of less than 0.2 dB at a 90° bend with radius of 5 mm at 100 GHz. The amount of mode cross coupling at such a bend is negligible and less than -33 dB in the whole frequency range from 80 GHz to 120 GHz.

On the transceiver front, several wideband millimeter-wave silicon IC solutions have been realized in CMOS and BiCMOS technologies. We have recently demonstrated a silicon transceiver with > 30 GHz of bandwidth (on a 94 GHz carrier) withon-chip antennas [9]. This chip has sufficient bandwidth to drive one of the orthogonal channels in the proposed dielectric waveguide. A simple QPSK modulation achieves over 120 Gbps throughput on a single waveguide. Frequency channelization and interleaving can enable this bandwidth in baseband. Other high-speed mm-wave CMOS transceivers for wireless communication have been demonstrated [10, 11]. Here, the channel response is significantly better than the wireless case.

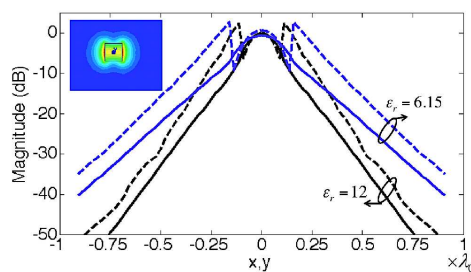


Figure 3: E_x magnitude of the E_x^{11} mode in the waveguide cross section along x -axis (dashed) and y -axis (solid), inset shows the E -field pattern in the cross section.

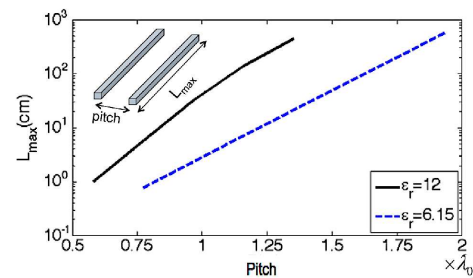


Figure 4: Maximum length of aggregated lines for a crosstalk of less than -15 dB for waveguides with permittivity of 12 (solid line) and 6.15 (dashed line).

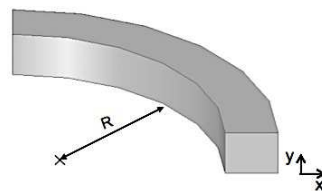


Figure 5: Multi-mode propagation along a bend with a radius of $R = 5$ mm.

3. SENSITIVITY ANALYSIS OF THE E_x^{11} AND E_y^{11} MODE LAUNCHERS

As seen in Fig. 1(b), the E_x^{11} mode is characterized by an electric field primarily in the x direction and thus, can be excited by an electric dipole along the axis. Conversely, the E_y^{11} mode is the dual of the E_x^{11} mode and as slot dipole antenna (dual of electric dipole) can be utilized to launch the mode into the waveguide. Recently we have presented planar structures, which efficiently launch the signal from integrated circuits into dielectric waveguides [12, 13]. A drawing of the structure presented in [13] is depicted in Fig. 6. The electric and slot dipoles are printed on 100 μm thick RO3010 substrates of permittivity 10.2, and placed on both sides of the waveguide. The high permittivity of the substrate reduces the radiation and therefore enhances the coupling efficiency. Fundamental design parameters of the excitation structures are detailed in [13].

Transmission parameters of S_{13} , S_{24} and isolation parameters of S_{14} , S_{23} , and S_{12} of the back-to-back structure (see port numbering in Fig. 6) with a length of 10 mm are plotted in Fig. 7. The obtained results show a 3-dB transmission bandwidth from 88 GHz to 115 GHz for the E_x^{11} mode and from 88 GHz to 120 GHz for the E_y^{11} mode. The coupling efficiency of each excitation structure is better than 3 dB in the operating frequency band. The bandwidth and efficiency of the field coupling can be improved by adding parasitic elements or loading to the dipole or slot antennas (e.g., Yagi-like planar electric and slot dipole antennas). An isolation of better than -50 dB between the two input couplers (S_{12}) is also demonstrated. The amount of cross coupling between input electric dipole and output slot dipole (S_{14}) as well as between input slot dipole and output electric dipole (S_{23}) is also better than 50 dB.

The sensitivity of coupling efficiency to the horizontal and angular misalignment of excitation sections is investigated. Fig. 8 shows the effect of horizontal offset of the E_x^{11} excitation section from the center of the waveguide. The overall performance is not highly perturbed and discrepancies are less than 0.5 dB. The amount of coupling at 100 GHz for different offsets is plotted in Fig. 9. As the offset increases the coupling efficiency decreases. This is mainly due to radiation and partially due to input impedance mismatch.

Beside the horizontal offset, the effect of angular misalignments is also examined. Figs. 10 and 11 demonstrate the effect of this misalignment on E_x^{11} and E_y^{11} mode launchers, respectively. These results show that E_y^{11} mode is relatively more sensitive to angular misalignment. This may be due to partial excitation of surface wave TM_0 mode in the substrate, which results in radiation in θ direction.

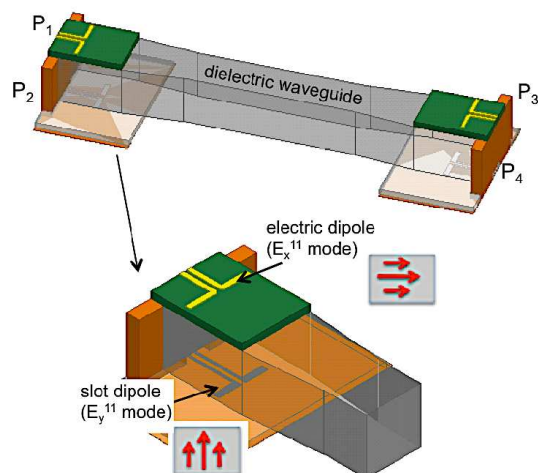


Figure 6: Drawing of the proposed structure for excitation of E_x^{11} (using an electric dipole) and E_y^{11} (using a slot dipole) modes.

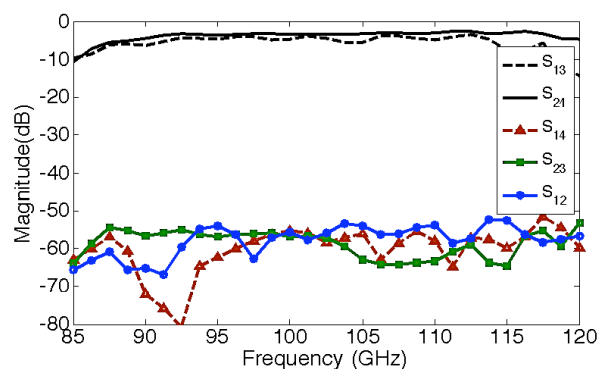


Figure 7: Transmission and isolation parameters of a 10 mm line terminated with E_x^{11} and E_y^{11} mode launchers (see port numbering in Fig. 6).

4. COMPARISON TO OTHER APPROACHES

Table 1 compares the planar mm-wave dielectric waveguide approach to two other common interconnects being used for high-data rate chip/board level communication. In [1] the partition length

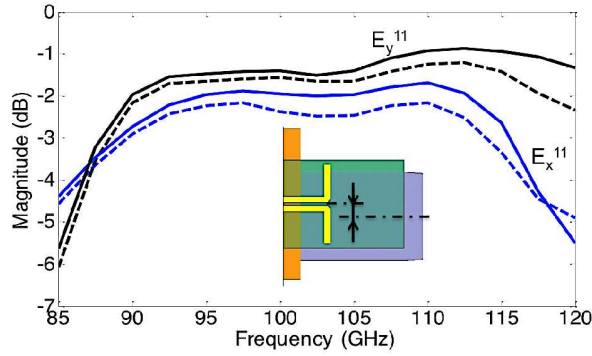


Figure 8: The sensitivity of the coupling efficiency to the misalignment of the waveguide (solid line: original structure, dashed line: 150 μm misalignment).

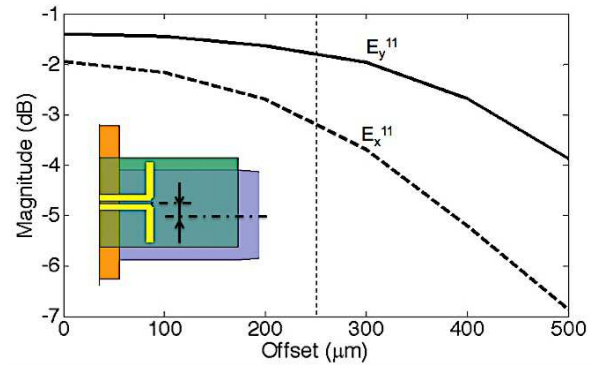


Figure 9: The effect of horizontal misalignment of both excitation sections. E_x^{11} is more sensitive to this offset.

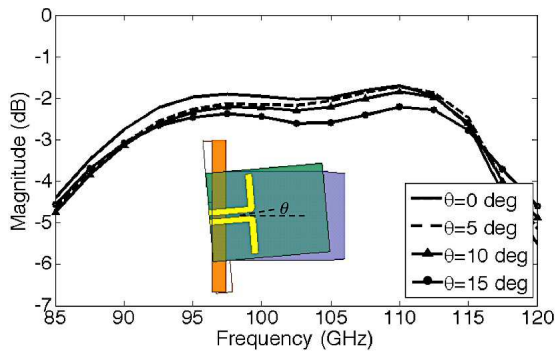


Figure 10: Sensitivity of the coupling efficiency of the E_x^{11} mode launcher to angular mismatches.

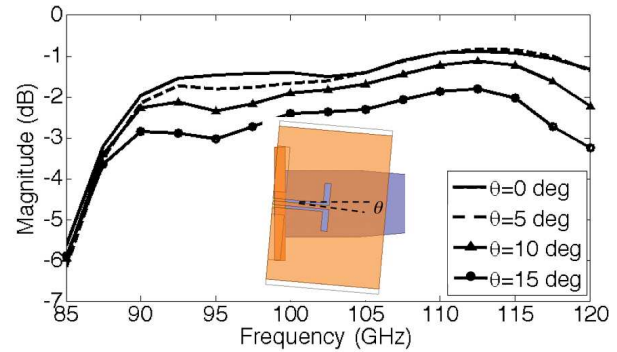


Figure 11: Sensitivity of the coupling efficiency of the E_y^{11} mode launcher to the angular mismatch.

Table 1: Interconnects options for high-data rate links.

Physical Medium:	Common Electrical	Optical	mm-Wave Waveguide
Transmission loss	High	Very low	Low
BW/pitch	Small	Very large	Large
Crosstalk	High	Very low	Low-Medium
Energy/Bit for high data-rates	High	High	Potentially Low
Overall Cost	Low	High	Medium

over which board level optical interconnects outperform conductive electrical wire lines is calculated. Common electrical interconnects such as planar microstrip or two-wire lines suffer from high loss and crosstalk. This limits the maximum applicable length and maximum bandwidth-per-pitch of a link in a single or aggregated lines scenarios. In addition, to achieve high throughput in these electrical links, higher order modulation schemes and complicated equalizers are inevitable. This reduces the power efficiency substantially and increases the cost and complexity of the system. Optical interconnects offer very wide bandwidth, small crosstalk and extremely low transmission loss. Hence they can be used for short/long-range communication. However the cost and complexity due to high precision needed for assembly and electro-optical modules limits the application of such interconnects.

Current approach based on a low-loss dielectric waveguide and its multi-mode excitation presents a large bandwidth over a reasonable length (around 1 m). In addition, the all-electrical and planar construction offers a low-cost implementation option. Superior channel response leads to simpler equalization circuits and can enhance the power efficiency of these high-throughput links.

5. CONCLUSION

Different aspects of an all-electrical, low-loss, and extremely wideband interconnect technology based on dielectric waveguides are analyzed. For a given range and receiver sensitivity various link parameters — the operating frequency, minimum line pitch, and consequently the maximum bandwidth-per-pitch — are calculated and verified with simulations. A reliable performance of the interconnect under physical curvature of the line is confirmed. Millimeter-wave signals are simultaneously launched into a dielectric waveguide using optimized planar dipole and slot antennas. The sensitivity of coupling efficiencies to horizontal and angular misalignment of the excitation structures is investigated. This low-cost interface presents a high-speed and low-dispersion alternative to current electrical and optical links. When combined with appropriate frequency channelization and interleaving, these links can achieve aggregate data rates approaching several Tbps over distances as long as one meter.

ACKNOWLEDGMENT

The authors acknowledge contributions from the sponsors of the Center for Integrated Systems (CIS) and the Rethinking Analog Design (RAD) initiative at Stanford University. Special thanks to Professor David Leeson for valuable suggestions and discussions. The authors thank Mustafa Rangwala for helpful comments.

REFERENCES

1. Naeemi, A., A. V. Mule, and J. D. Meindel, "Partition length between board-level electrical and optical interconnects," *IEEE Intern. Tech. Conf.*, 230–232, Jun. 2003.
2. Shindo, S. and T. Itanami, "Low-loss rectangular dielectric image line for millimeter-wave integrated circuits," *IEEE Trans. Microw. Theory Tech.*, Vol. 26, No. 10, 747–751, Oct. 1978.
3. Dolatsha, N. and J. Hesselbarth, "Millimeter-wave chip-to-chip transmission using an insulated image guide excited by an on-chip dipole antenna at 90 GHz," *IEEE Microw. Wireless Comp. Lett.*, Vol. 22, No. 5, 266–268, May 2012.
4. Haroun, B. S., M. Corsi, S. Akhtar, and N. C. Warke, "Chip-to-dielectric waveguide interface for sub-millimeter-wave communication link," US Patent Application Filed, Sept. 2010.
5. Fukuda, S., Y. Hino, S. Ohashi, T. Takeda, H. Yamagishi, S. Shinke, K. Komori, M. Uno, Y. Akiyama, K. Kawasaki, and A. Hajimiri, "A 12.5 + 12.5 Gb/s full-duplex plastic waveguide interconnect," *IEEE J. Solid-State Circuits*, Vol. 46, No. 12, 3113–3125, Dec. 2011.
6. Suntives, A. and R. Abhari, "Dual-mode high-speed data transmission using substrate integrated waveguide interconnects," *IEEE Electrical Performance Electronic Packaging*, 215–218, Oct. 2007.
7. Bierwirth, K., N. Schulz, and F. Arndt, "Finite-difference analysis of rectangular dielectric waveguide structures," *IEEE Trans. Microw. Propag.*, Vol. 34, No. 11, 1104–1114, Nov. 1986.
8. Koul, S. K., *Millimeter Wave and Optical Dielectric Integrated Guides and Circuits*, J. Wiley Sons, New York, 1997.
9. Arbabian, A., et al., "A 94 GHz mm-wave to baseband pulsed-radar for imaging and gesture recognition," *IEEE Symp. VLSI Circuits (VLSIC)*, 56–57, 2012.
10. Marcu, C., et al., "A 90 nm CMOS low-power 60 GHz transceiver with integrated baseband circuitry," *IEEE J. Solid-State Circuits*, Vol. 44, No. 12, 3434–3447, 2009.
11. Siligaris, A., et al., "A 65-nm CMOS fully integrated transceiver module for 60-GHz wireless HD applications," *IEEE J. Solid-State Circuits*, Vol. 46, No. 12, 3005–3017, Dec. 2011.
12. Dolatsha, N. and J. Hesselbarth, "Power divider based on grounded dielectric slab waveguide operating in E_x^{11} mode," *German Microw. Conf. (GeMIC)*, 1–4, Darmstadt, Germany, May 2011.
13. Dolatsha, N. and A. Arbabian, "Dielectric waveguide with planar multi-mode excitation for high data-rate chip-to-chip interconnects," Accepted in *IEEE International Conf. Ultra-Wideband (ICUWB)*, 2013.

Design Considerations for a Low-frequency Vivaldi Array Element

A. Tibaldi¹, G. Virone², F. Perini³, J. Monari³, M. Z. Farooqui¹, M. Lumia²,
O. A. Peverini², G. Addamo², R. Tascone², and R. Orta¹

¹Dipartimento di Elettronica e Telecomunicazioni (DET), Politecnico di Torino
Corso Duca degli Abruzzi 24, Torino 10129, Italy

²Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni (IEIIT), Consiglio
National Research Council of Italy (CNR) c/o Politecnico di Torino
Corso Duca degli Abruzzi 24, Turin 10129, Italy

³Istituto di Radioastronomia (IRA), Istituto Nazionale di Astrofisica (INAF)
Via Fiorentina 3513, Medicina (BO) 40059, Italy

Abstract— A cavity-backed Vivaldi antenna is suggested as dual-polarization array element for the low-frequency instrument of the Square Kilometer Array (SKA) project. A design strategy aimed at maximizing the sensitivity for such an array element is described. As an example, an antenna was obtained, with a sensitivity higher than $10 \text{ cm}^2/\text{K}$ in the operative bandwidth and in the 45° sky coverage angle, for each polarization.

1. INTRODUCTION

The Square Kilometer Array (SKA) represents one of the most interesting new-generation radio-telescopes owing to its extreme sensitivity performance [1]. One of the main SKA subsystems is the low-frequency Aperture Array (AA-low), which has to operate in the $[70, 450]$ MHz band [2]. Several wide-band antenna systems have been studied within this framework: the spiral antenna [2], the BLU-antenna [3], the log-periodic antenna [4]. Recently, a dual-polarization Vivaldi array element has been proposed as a potential candidate for AA-low [5]. The main advantages of this configuration are a single ended 50Ω matching and a low cross-polarization in the principal planes, owing to the symmetry of the antenna. Furthermore, the antenna is self-standing, does not require either bulky dielectric parts or ground planes and can be manufactured in a gridded version [6].

2. DEVELOPMENT OF A CAVITY-BACKED VIVALDI CONFIGURATION

The first item of this work is the selection of a suitable geometry for the rear part of the antenna, in order to enhance its performance in terms of front-to-back ratio (FBR). The starting structure is the Vivaldi antenna with a circular back stub adopted in [5]; its FBR is presented in Figure 1, as a black dashed line. This parameter can not be improved by introducing a ground plane behind the antenna, because this would introduce an additional ripple in the curve, without significant improvements (red dash-dotted line).

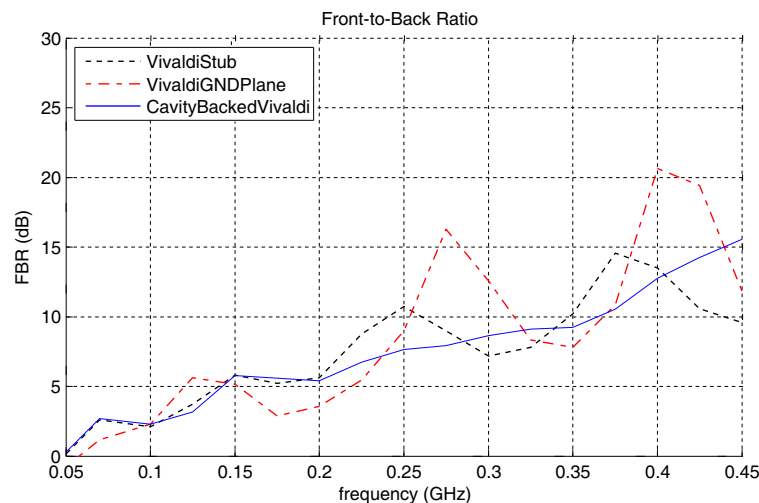


Figure 1: Front-to-back ratio of Vivaldi with three back structures.

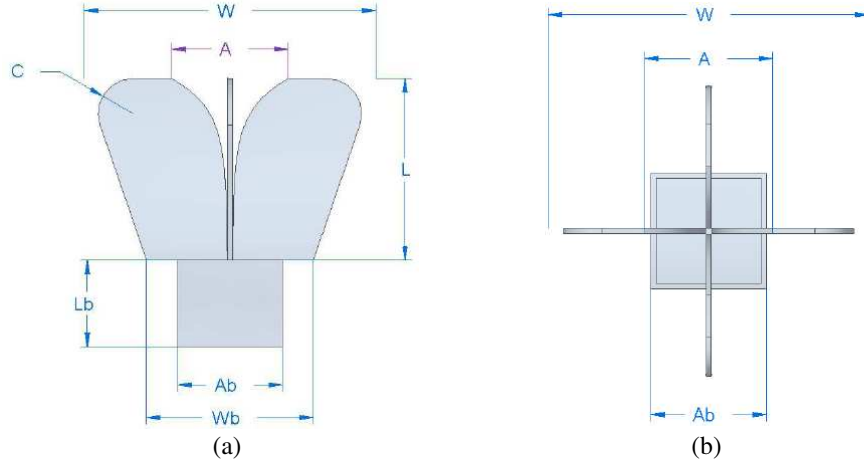


Figure 2: Cavity-backed Vivaldi antenna. (a) Lateral view. (b) Front view.

On the contrary, recalling the TEM horn structures [7], the cavity-backed configuration represented in Figure 2 has been conceived. The feed point of the antenna is located at the junction of the Vivaldi “wings” with the cavity. The design parameters of this structure are: The front width of the non-blended antenna W , the aperture width A , the antenna length L , the blending radius C , the base width W_b , the back cavity length L_b and the back cavity width A_b .

In Figure 1 it is possible to observe that the FBR curve of a non-optimized version of this structure is almost monotone with improved values at higher frequencies. This structure has been used as starting guess for the following design procedure.

3. DESIGN PROCEDURE

The AA-low instrument is a sparse random array, where the average embedded element pattern is, as a first approximation, similar to the pattern of the single (isolated) element [8,9]. In this regard, the element design can be performed focusing on the sensitivity enhancement of the single element [2]. The maximization of the worst-case sensitivity of the array element corresponds to the minimization of the number of antennas needed to satisfy the SKA sensitivity specifications, leading to a reduction of manufacturing costs.

The sensitivity S is defined as the ratio of the element effective area A_{eff} to its noise temperature T_{sys} :

$$S(\hat{r}, f) \triangleq \frac{A_{eff}(\hat{r}, f)}{T_{sys}(f)} \left(\frac{\text{m}^2}{\text{K}} \right) \quad (1)$$

where \hat{r} indicates the observation direction and f is the frequency. A_{eff} is calculated from the radiation patterns, which are obtained using a full-wave simulator.

For what concerns the denominator, T_{sys} can be calculated as the sum of three contributions:

$$T_{sys}(f) = T_{ant,sky}(f) + T_{ant,gnd}(f) + T_{rec}(f) \quad (2)$$

$T_{ant,sky}(f)$ and $T_{ant,gnd}(f)$ quantify the noise contributions coming from the sky and from the ground. These two quantities are evaluated by means of the Cortes model [10]. The measured receiver noise temperature $T_{rec}(f)$ is approximately 30 K in the whole band.

The structure should be designed in order to maximize the sensitivity in the 70–450 MHz band within the 45° sky coverage from zenith (SC), for each polarization. This goal is achieved by exploiting a synthetic representation of the sensitivity as a function of the geometrical parameters. It is useful to define the goal function \tilde{S} :

$$\tilde{S}(f) = \min_{\hat{r} \in \text{SC}} S(\hat{r}, f) \quad (3)$$

As far as frequency is concerned, it should be noted that the operative conditions in the AA-low band are not homogeneous. Indeed, in the lower part of this band, the sky noise contribution is dominant. On the other hand, the most significant high-frequency noise contributions are the

remaining two. Therefore, $\tilde{S}(f)$ has been parametrized by means of its minimum values in three sub-bands: $B_1 = 70\text{--}200$ MHz, which is the sky-noise dominated band; $B_2 = 200\text{--}350$ MHz, which is a transition band; $B_3 = 350\text{--}450$ MHz band, which is dominated by the receiver noise.

This representation of the goal function is very convenient for the evaluation of the effects of the geometrical parameters; for example, $\tilde{S}(f)$ is shown in Figure 3 as a function of the length of the Vivaldi antenna L and of its aperture width A . The white star markers identify the optimal values for the B_2 and B_3 bands. From the figure it is also possible to observe that the design procedure, in this case, is mainly driven by the higher band, where the sensitivity is generally low. Moreover, the sensitivity in B_1 appears to be almost independent of A and L .

The same procedure must be performed varying other couples of parameters, in order to complete the design of the structure.

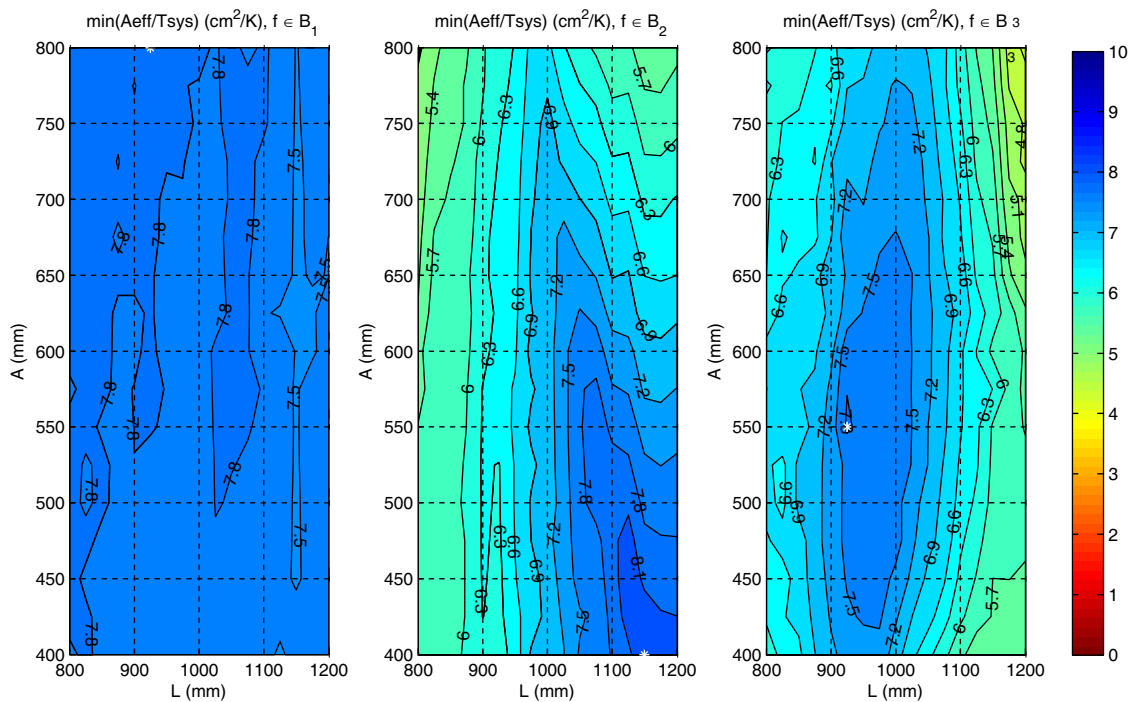


Figure 3: Minima of $\tilde{S}(f)$ in the three sub-bands.

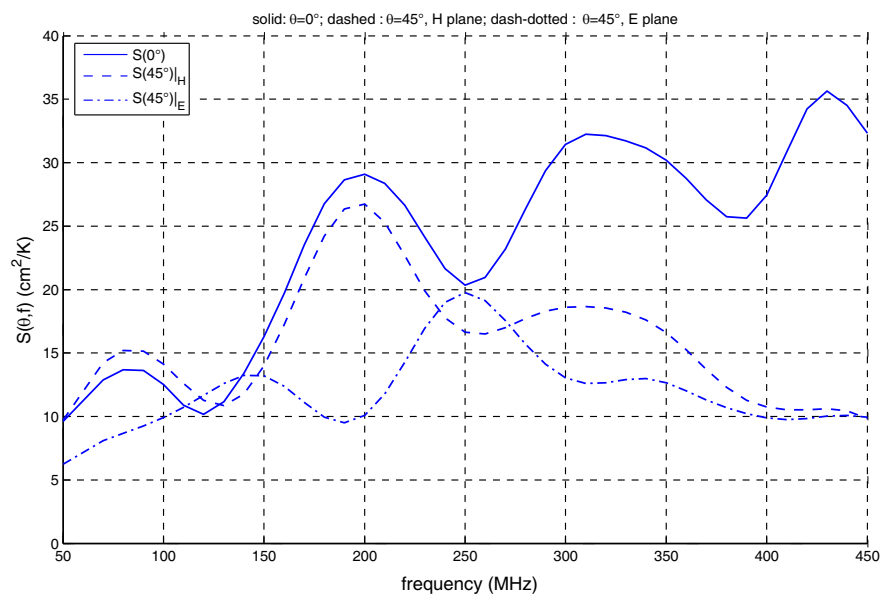


Figure 4: Sensitivity goal function of the designed structure.

4. DESIGN RESULTS

The performance of a significant design example is discussed in this section. The main dimensions of the designed antenna are $1.2 \times 1.2 \text{ m}^2$ footprint and 1.5 m height. Figure 4 shows the sensitivity function $S(\hat{r}, f)$ for $\vartheta = 0^\circ$ and $\vartheta = 45^\circ$ in the H -plane and in the E -plane; it is possible to observe that $S(\hat{r}, f)$ is higher than $10 \text{ cm}^2/\text{K}$ in most situations. The effective area of the designed antenna and the three noise temperature contributions are represented in Figure 5 and in Figure 6. In particular, Figure 6 confirms that the main noise contribution for lower frequencies is the sky one, while the receiver noise is dominant for higher frequencies. Moreover, the new configuration exhibits a good symmetry of the pattern, which leads to high IXR values [11].

In Figure 7 it is possible to observe that the 50Ω reflection coefficient is below -10 dB above 170 MHz.

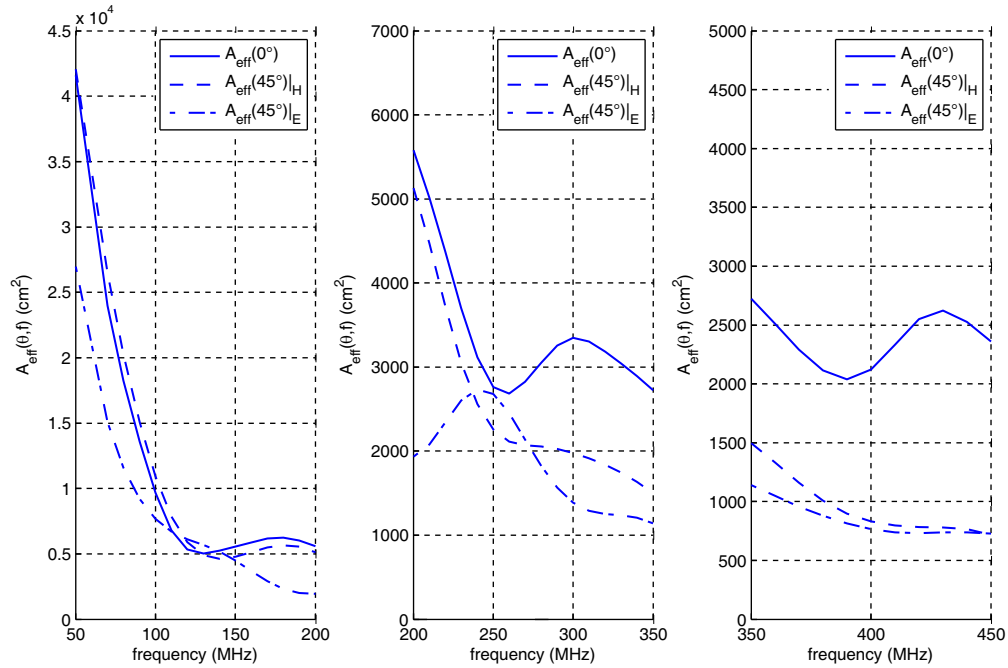


Figure 5: $A_{eff}(\vartheta, f)$ of the structure.

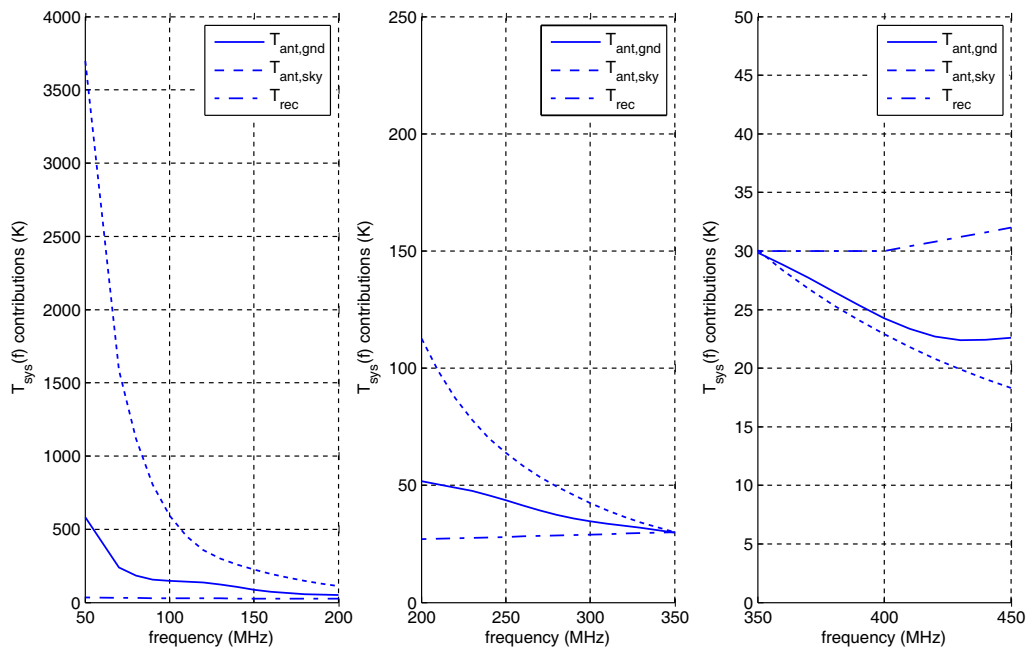


Figure 6: $T_{sys}(f)$ contributions of the designed structure.

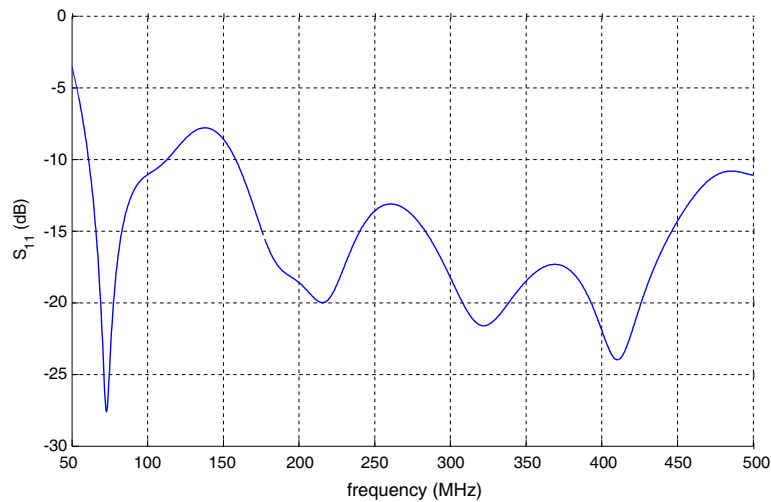


Figure 7: Reflection coefficient of the designed structure.

5. CONCLUSION

A cavity-backed Vivaldi antenna has been proposed as array element for the SKA AA-low project. A design procedure based on a full-wave simulator and the Cortes noise temperature model has been developed and described. An effective organization of the sensitivity parametric analyses has been proposed in order to define a suitable and exhaustive design strategy; it should be noted that the same design procedure can be applied to other antenna structures.

In conclusion, the results obtained demonstrated that good performance can be achieved with the Vivaldi element.

REFERENCES

1. <http://www.skatelescope.org/>.
2. De Vaate, J. G. B., et al., "Low frequency aperture array developments for phase 1 SKA," *XXXth URSI General Assembly and Scientific Symposium*, 1–4, Istanbul, Aug. 13–20, 2011.
3. De Lera Acedo, E., N. Razavi-Ghods, E. Garcia, P. Duffett-Smith, and P. Alexander, "Analysis of an ultra wideband aperture array element for low frequency radio astronomy," *IEEE Antennas and Propagation Society International Symposium, APSURSI '09*, 1–4, Jun. 1–5, 2009.
4. De Lera Acedo, E., "SKALA: A log-periodic antenna for the SKA," *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 353–356, Sep. 2–7, 2012.
5. Virone, G., G. Addamo, O. A. Peverini, and R. Tascone, "Broadband array element for the SKA low-frequency aperture array," *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 560–561, Sep. 12–16, 2011.
6. Virone, G., R. Sarkis, C. Craeye, G. Addamo, and O. A. Peverini, "Gridded vivaldi antenna feed system for the northern cross radio telescope," *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 6, 1963–1971, Jun. 2011.
7. Malherbe, J. A. G., "Hybrid elliptic TEM horn with internal fins," *IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)*, 1009–1011, Sep. 2–7, 2012.
8. Gonzalez-Ovejero, D., E. de Lera Acedo, N. Razavi-Ghods, C. Craeye, and L. E. Garcia Munoz, "Non-periodic arrays for radio-astronomy applications," *IEEE International Symposium on Antennas and Propagation (APSURSI)*, 1762–1765, Jul. 3–8, 2011.
9. De Lera Acedo, E., N. Razavi-Ghods, et al., "SKA AA-low front-end developments (at Cambridge University)," *6th European Conference on Antennas and Propagation (EUCAP)*, 616–620, Prague, Mar. 26–30, 2012.
10. Bolli, P., F. Perini, S. Montebugnoli, G. Pelosi, and S. Poppi, "Basic element for square kilometer array training (BEST): Evaluation of the antenna noise temperature," *IEEE Antennas and Propagation Magazine*, Vol. 50, No. 2, 58–65, Apr. 2008.
11. Fiorelli, B., M. Arts, G. Virone, E. de Lera Acedo, and W. A. van Cappellen, "Polarization analysis and evaluation for radio astronomy aperture array antennas" *7th European Conference on Antennas and Propagation (EUCAP)*, 453–457, Apr. 2013.

Ultra-Wideband Tightly Coupled Phased Array Antenna for Low-Frequency Radio Telescope

E. O. Farhat¹, K. Z. Adami^{1,2}, Y. Zhang³, A. K. Brown³, and C. V. Sammut¹

¹University of Malta, Malta

²University of Oxford, United Kingdom

³The University of Manchester, United Kingdom

Abstract— This paper introduces a novel approach to a broadband array design for a low frequency radio telescope. It presents a low profile ultra-wideband tightly coupled phased array antenna with integrated feedlines. The approach consists of applying broadband techniques to an array of capacitively coupled planar element pairs with an octagonal fractal geometry, backed by a ground plane. Designed as a low cost, low loss, dual-polarized wideband array, this antenna is optimised for operation between 50 and 250 MHz. Simulations have shown that the antenna has a wide-scanning ability with a low cross-polarisation level, over the operational broad frequency range.

1. INTRODUCTION

Nowadays, inherently low profile ultra wideband phased array antennas are significantly in demand especially for low frequency band radio telescope applications. Tightly coupled phased arrays (TCPAs) have a significant feature to imply low profile ultra wideband antennas. This feature is critical to deploy and develop radio telescope in space base station. It is also important for other applications such as advanced defence and commercial communication systems.

Characteristically, the tightly coupled arrays are planar elements with strong mutual coupling. Each element array acts as an aperture array rather than discrete elements. The inter element capacitive coupling is used to counteract the ground plane reactive impedance and maintain a real input impedance over a wide bandwidth. This paper presents a tightly coupled array of a conformal performance while maintaining a low profile of roughly $0.05\lambda_{low}$ where λ_{low} is the wavelength at the lowest operational frequency of the operating bandwidth [1]. However, the Vivaldi antenna is widely used as an ultra-wideband antenna, but it exhibits a profile of $0.5\lambda_{low}$ which is too thick especially in the low frequency range [2]. Additionally, the bunny ear elements also provide an ultra-wideband performance with low profile of about $0.125\lambda_{low}$ but still thick for low frequency applications [3].

This work aims to enhance the bandwidth of a proposed TCPA, operating between 50 and 250 MHz to operate from 50 MHz to 400 MHz, using a resistive frequency selective surface. The Frequency Selective Surfaces (FSSs) were first demonstrated by Ben A. Munk and used in airborne applications [4]. Recently, many researchers proposed FSS designs for different microwave applications, such as radar absorber [5], and applications associated with radome design for radar systems [6]. In [7] an artificial dielectric material was placed on a frequency selective surface to absorb the energy of electromagnetic waves at a particular frequency or range of frequencies. Also, a resistive FSS was used in [8] to enhance the bandwidth of an ultra-thin absorber. A tightly coupled bowtie array with a resistive FSS layer has shown to exhibit a 21 : 1 bandwidth with VSWR < 3 in [9], implemented in an infinite array configuration.

The paper is organised as follows: in Section 2, capacitively coupled phased array design and analysis are presented. Section 3 describes new broadbanding technique used to enhance the bandwidth performance of the proposed capacitively coupled phased array, whilst the analysis of an infinite array is also presented. The conclusions are subsequently driven in Section 4.

2. CAPACITIVELY COUPLED PHASED ARRAY ELEMENT DESIGN

A TCPA is an array of very closely spaced elements, frequently separated via tip capacitors. The tip capacitors are adjusted to provide wideband performance. The aperture elements are Tightly Coupled Fractal Octagonal rings Array (TCFOA) backed by a ground plane. Therefore, a strong mutual coupling is utilised between the elements resulting in a continuous current distribution array. The geometry and characterization of TCFOA element is introduced in an infinite array environment. The radiators or receptors are dual polarised via two orthogonal feeding points, in

each element the centre ring is shared between other two rings. An impedance transformer layer, known as a matching layer, is placed at a certain distance above the radiator layer. The matching layer is a scaled down version of the radiating elements, applied to enhance the bandwidth of the array. The scaling factor used to form the matching layer was 0.9. Figure 1 illustrates the unit cell of the TCFOA element antenna design. The space between the adjacent elements is 800 mm, which is slightly greater than half wavelength at the highest frequency (250 MHz for this unit cell). The unit cell covers frequency range from 60 MHz to 250 MHz. The distance from the radiators to the matching layer is 235 mm while the ground plane is placed at 390 mm from the radiators. To lower the array's operating frequency, capacitors are used in the gap between the neighbouring rings at the tip ends. The gap between the rings is 5 mm. These capacitors are bulk capacitors of 3 pF. The array design parameters are jointly optimized together, using Genetic Algorithm implemented in Matlab to validate a perfect match via commercial software Ansoft HFSS v12. The objective of this optimisation is to reduce the VSWR values over the operating frequency band of the antenna array. An iteration of nine scaled and subdivided octagonal rings are proposed to create an aperture array of fractal geometry patches. The bandwidth performance of the periodic array is sensitive to the outer and inner diameter of the fractal rings. Therefore, the optimised values of both the outermost and the innermost ring radius are 197.5 mm and 170 mm respectively.

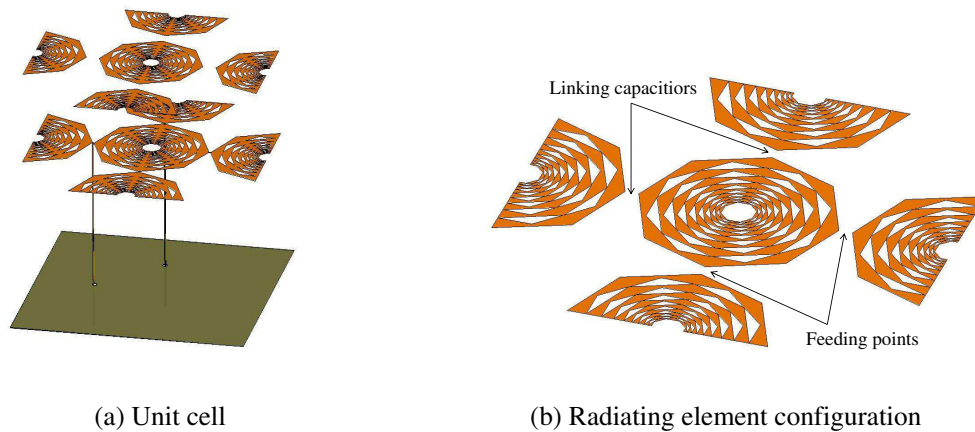


Figure 1: Geometry of TCFOA unit cell with the feeding lines.

A stripline line was introduced to provide a balance feed to the TCFOA, this feeding line was first presented in [10]. To match antennas input impedance with a $50\ \Omega$ input SMA connector, an impedance transformer was implemented with the stripline design. This feeding lines have a significant feature of a low insertion loss. For an ultra wideband performance, the single-ended feeding lines were optimised with the unit cell. As a result, the stripline length is 503.5 mm. Thus, an extra part of the body is reaching out the ground plane. Comparing with the tightly coupled dipole presented by Munk et al. [11], the candidate array design offers a convenient integration with the feeding lines.

To further demonstrate the validity of the proposed design, a full-wave simulation data is presented. The unit cell is modelled using periodic boundary conditions. Although the infinite array approach provides a rapid computational analysis, it does not account for the finite array edge effect. The finite arrays edge and corner diffraction affects the outer periphery elements causing a mismatch. As a result, the impedance of the edge and corner within the finite array size differs significantly from the infinite array. Therefore, the finite array bandwidth is degraded and several redesign steps are required to achieve the intended operational bandwidth. The broadside as well as the scanning active VSWRs of an infinite dual-polarised TCFOA was calculated using HFSS over a unit cell. The VSWR performance shown in Figure 2 suggests that satisfactory scanning properties and wideband performance could be achieved over 4.4 : 1 frequency band. Figure 3 illustrates that the TCFOA antenna design has a wideband performance with a stable low-cross polarisation level across the bandwidth from 60 MHz to 266 MHz.

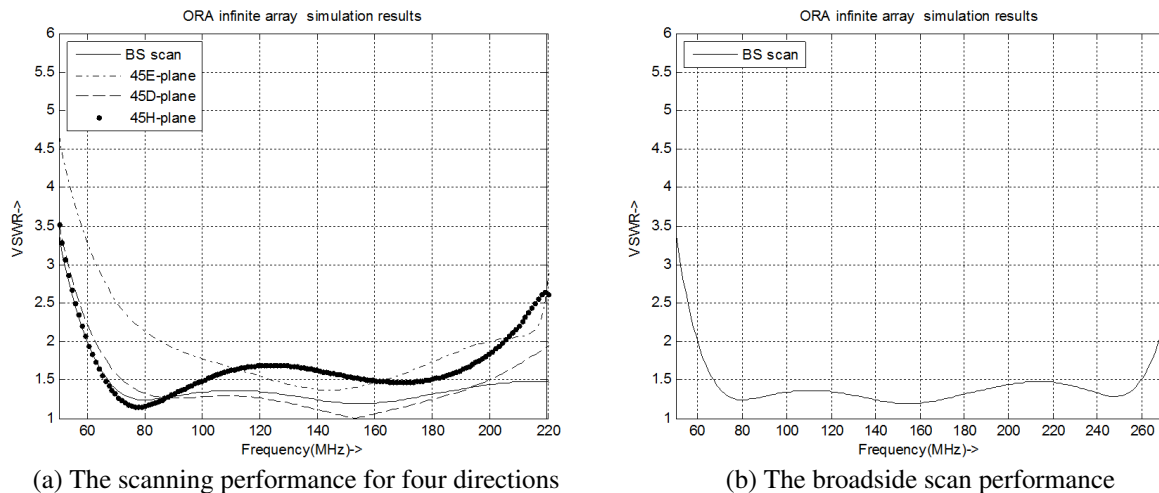


Figure 2: The VSWR for an element in the infinite array.

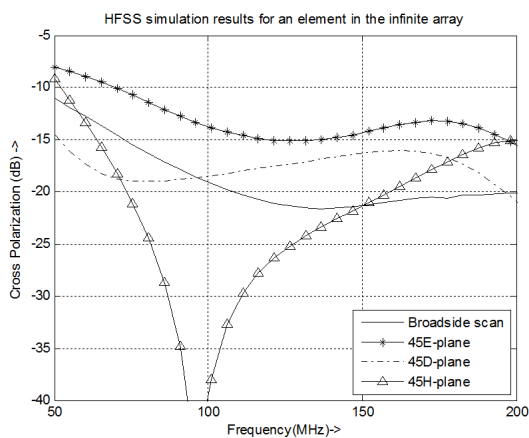


Figure 3: The cross polarization level in dB for an element in the infinite dual polarized array.

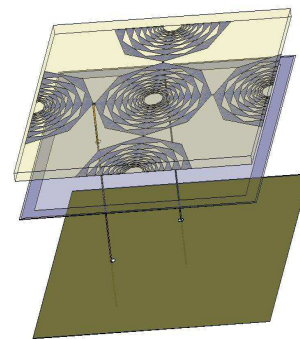


Figure 4: Geometry of the TCFORA unit cell with the resistive FSS.

3. CAPACITIVELY-COUPLED ARRAY WITH RESISTIVE FSS

In general, a tightly coupled array bandwidth is limited by the ground plane. More precisely, an array with a ground plane spacing of h is short circuited at the upper bond of the operational band at $f_{upper} = c/2h$. In other words, the upper bound bandwidth of any TCPAs is limited due to ground plane Z_{GP} given by:

$$Z_{GP} = j\eta_0 \tan(\beta h) \quad (1)$$

where η_0 is the substrate impedance (in this array it is free space), β is the substrate propagation constant, and h is the array spacing above the ground plane. The array depicts a resonance peak at $h = c/2f_{upper}$, because the ground plane impedance becomes $Z_{GP} = 0$.

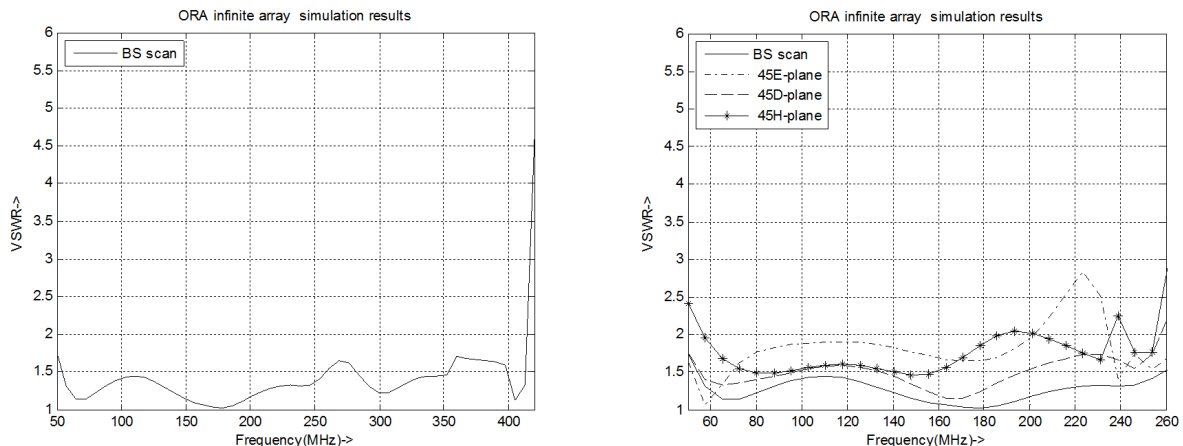
Sating that a conformal array is short circuited when $h = \frac{\lambda_H}{2}$ (where λ_H is the wavelength at the highest/upper operational frequency), a method to alleviate ground plane effect is required. Though Electromagnetic Band Gap structures (EBG) were proposed to overcome the ground plane effect, it operates only over limited bandwidth [12]. Ferrites were also used to improve bandwidth, but their weight limits their applications [13]. In order to avoid boresight radiation cancellation from the ground plane image current, this paper proposes inserting a resistive FSS between the radiating elements and the ground plane. The resistive FSS suppresses the interfering of the ground plane reflection, increasing the bandwidth performance of the capacitively coupled array.

To serve this purpose, a unit cell is loaded with a square ring of a resistive FSS. Figure 4 depicts the configuration of the capacitively coupled array with the resistive FSS. The array unit cell is modelled with the stripline feeding lines. The physical array design parameters are obtained

through the same aforementioned optimisation procedure. The resistive FSS loop is printed on a Polyethylene dielectric substrate of 0.2 mm thickness, this square ring loop is an ohmic sheet of $75\Omega = sq$. However, inserting FSS in the array leads to severe losses. A superstrate layer of a Polyethylene dielectric substrate (60 mm thickness) is mounted over the radiating elements. Both the FSS and the superstrate must be designed together in tandem. The purpose of a dielectric superstrate layer is to alleviate FSS losses. In addition, this layer behaves as a matching impedance transformer. In an infinite array TCFOA configuration, using superstrate in conjunction with FSS leads to a 8 : 1 bandwidth with $VSWR < 2$. Hence, this unique aspect of the array maximizes the bandwidth by a factor of 2 dB, resulting in a very low profile TCFOA.

The main layer is linearly distributed pairs of fractal octagon rings, perpendicular to each other. The element unit cell spacing is 730 mm, and the array is dual polarised. The array is capacitively coupled by inserting bulk capacitors at the end tips of adjacent rings, the value of these capacitors is 3 pF. The array is backed by a ground plane at 390 mm. The overall profile of the TCFOA with FSS is 450 mm (element spacing 730 mm), while it is 625 mm for the previous TCFOA design(element spacing 800 mm). Thus, the advantage of this arrangement presents a very wide bandwidth with a lower profile. This is a considerable feature for large scale arrays applications, and it facilitates low cost for the mass production.

The optimized infinite TCFOA with FSS scan performance is shown in Figure 5. The array with FSS exhibits a wideband performance, exceeding TCFOA without FSS. The Ludwig third definition of cross polarisation is used. And the cross polarisation of the immersed element scanned to four typical directions is illustrated in Figure 6. A stable cross polarisation performance is observed, and also it can be remarked easily that FSS geometry shows it is polarisation intensive.



(a) The broadside scan performance

(b) The scanning performance for four directions

Figure 5: The VSWR for an element in the infinite array with a resistive FSS layer.

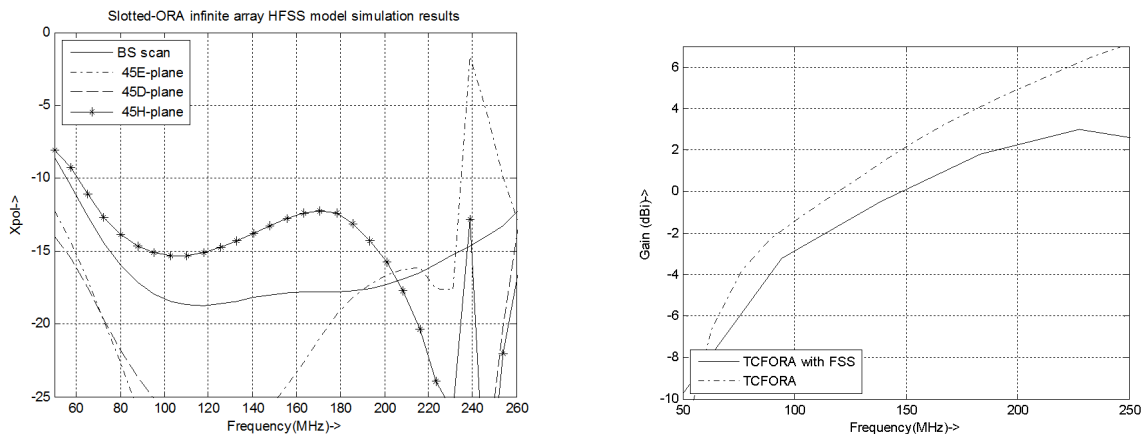


Figure 6: The cross polarization level in dB for an element in the infinite dual-polarized array.

Figure 7: The gain for the immersed element in an infinite array.

Figure 7 illustrates that the presence of FSS cause some reduction in the gain (efficiency), this might be due to the power dissipated in the resistance of FSS.

4. CONCLUSION

In this paper, it has been shown that a properly designed resistive frequency selective surface, in tandem with a dielectric superstrate presents an inherently low-profile ultra-wideband phased array for a low-band radio telescope. The resistive FSS conjugated with a superstrate was demonstrated to significantly increase the array bandwidth. The infinite array environment achieves 8 : 1 bandwidth with VSWR < 2. The array thickness at the lowest operational frequency of 50 MHz is slightly greater than 0.058λ .

ACKNOWLEDGMENT

The authors would like to thank Kristian. Z. Adami for his kind cooperation and help.

REFERENCES

1. Munk, B. A., "Broadband wire arrays," *Finite Antenna Arrays and FSS*, 181–213, Wiley-Interscience, New York, 2003.
2. Lee, J. J., "Ultra wideband arrays," *Antenna Engineering Handbook*, New York, McGraw-Hill, 2007.
3. Lee, J. J., S. Livngston, and R. Koenig, "A low-profile wide-band (5 : 1) dual-pol array," *IEEE Antennas and Wireless Propagation Letters*, Vol. 2, No. 3, 46–49, 2003.
4. Munk, B. A., *Frequency Selective Surface: Theory and Design*, Wiley, New York, 2000.
5. Zadeh, A. K. and A. Karlson, "Capacitive circuit method for fast and efficient design of wideband radar absorbers," *IEEE Transactions on Antennas and Propagation*, Vol. 57, No. 8, 2307–2314, 2009.
6. Virone, G., R. Tascone, G. Addamo, and O. A. Peverini, "A design strategy for large dielectric radome compensated joints," *IEEE Antennas and Wireless Propagation Letters*, Vol. 8, 546–549, 2009.
7. Akins, R. D., "Absorptive/transmissive random," U. S. Patent 5,400,043 A, 1995.
8. Virone, G., R. Tascone, G. Addamo, and O. A. Peverini, "An ultrathin and broadband radar absorber using resistive FSS," *IEEE Antennas and Wireless Propagation Letters*, Vol. 11, 748–751, 2012.
9. Moulder, W. F., K. Sertel, and J. L. Volakis, "Superstrate-enhanced ultrawideband tightly coupled array with resistive FSS," *IEEE Transactions on Antennas and Propagation*, Vol. 60, No. 9, 4166–4172, 2012.
10. Zhang, Y. and A. K. Brown, "Octagonal ring antenna for a compact dual-polarized aperture array," *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 10, 3927–3931, 2011.
11. Munk, B. A., et al., "A low-profile broadband phased arrays," *Proc. Antenna Propagation Soc. Int. Symp.*, 448–451, Jun. 2003.
12. Sievenpiper, D., R. F. Broas, N. G. Alexopolous, and E. Yablonovitch, "High-impedance electromagnetic surfaces with a forbidden frequency band," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 47, No. 11, 2059–2074, 1999.
13. Bell, J. M., M. F. Iskander, and J. J. Lee, "Ultrawideband hybrid EBG/ferrite ground plane for low-profile array antennas," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 1, 2–12, Jan. 2007.

A Compact Triple-band Filtering Microstrip Patch Antenna with the Same Polarization Planes

Y.-J. Lee and S.-J. Chung

Institute of Communications Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan

Abstract— A novel compact filtering microstrip patch antenna with triple-band operation is presented for WLAN/TD-LTE wireless applications. The proposed design is composed of a patch with slot-loaded parasitic elements and a stepped-impedance resonator which shares the same apertured ground. Using full-wave simulator HFSS, the broadside coupling coefficient can be extracted to determine the dimension of the aperture. The antenna has similar broadside radiation patterns in both E - and H -planes in the three frequency bands. Meanwhile, the antenna provides good selectivity as well as the same polarization planes.

1. INTRODUCTION

Nowadays, in the RF front end of wireless communication systems it is usually needed to process different frequency signals. By designing a filtering antenna, it not only improves the performance of the front end but also fulfills the requirements for compact, low cost, and low profile.

In the previous works [1–3], we have studied for synthesis a filtering microstrip antenna in a single band or dual bands. In this paper, a triple-band filtering antenna is presented for applying to WLAN/TD-LTE wireless applications, i.e., 2.4, 3.5, and 5.2 GHz bands. By using the similar ideas of [4] and [5], the patch antenna with slot-loaded parasitic elements has almost the same polarization plane feature in the three bands. It is advantageous to utilize stepped-impedance resonators (SIRs) for designing bandpass filters cause its versatile resonance characteristics. The antenna with filtering function can be developed by the synthesis process of bandpass filters [6].

2. ANTENNA DESIGN

Figure 1 shows the proposed triple-band microstrip patch antenna composed of a patch with slot-loaded parasitic elements and a step-impedance resonator. The slot-loaded parasitic elements are incorporated into the non-radiating edges of a microstrip patch which is operated in the second band as depicted in Fig. 1. The parasitic elements are worked for the first and third bands. The coupling between the main patch and slot-loaded parasitic elements is realized by controlling the vertical and horizontal gaps. No matter which band the patch works in, the maximum magnetic field density occurs near the center section on it. The resonator is also designed to resonate in the same three bands. When it works in different bands, the distribution of electric/magnetic field density is different on it. For example, in the second band, there is the maximum magnetic field density near one/three-fourths section on it. Thus, if there is a thin slit on the common ground for broadside coupling the magnetic coupling can be obtained. So, we tune the size of slit and the position of resonator, the corresponding coupling coefficients in the design can be obtained.

The proposed triple-band antenna is fabricated on two FR4 substrates with common used ground. The patch and its slot-loaded parasitic elements are printed on top side of upper FR4 substrate with dielectric constant 4.4 and thickness 0.8 mm. The resonator is patterned on the bottom side of lower FR4 substrate with the same dielectric constant but thickness 0.4 mm. The etched slit on the ground is used for broadside coupling. The dimensions of the antenna are listed in Table 1.

The antenna was designed at WLAN/TD-LTE bands. The three bands are at 2.45, 3.5, and 5.25 GHz with 3.4%, 5.7%, and 5.7% fractional bandwidth, respectively. The ripples are all 0.3 dB. The corresponding coupling coefficients and external quality factor are found to be

$$\begin{aligned}
 M_{12}^I &= 0.044 \\
 M_{12}^{II} &= 0.063 \\
 M_{12}^{III} &= 0.063 \\
 Q_e^I &= 29.51 \\
 Q_e^{II} &= 20.71 \\
 Q_e^{III} &= 20.71
 \end{aligned} \tag{1}$$

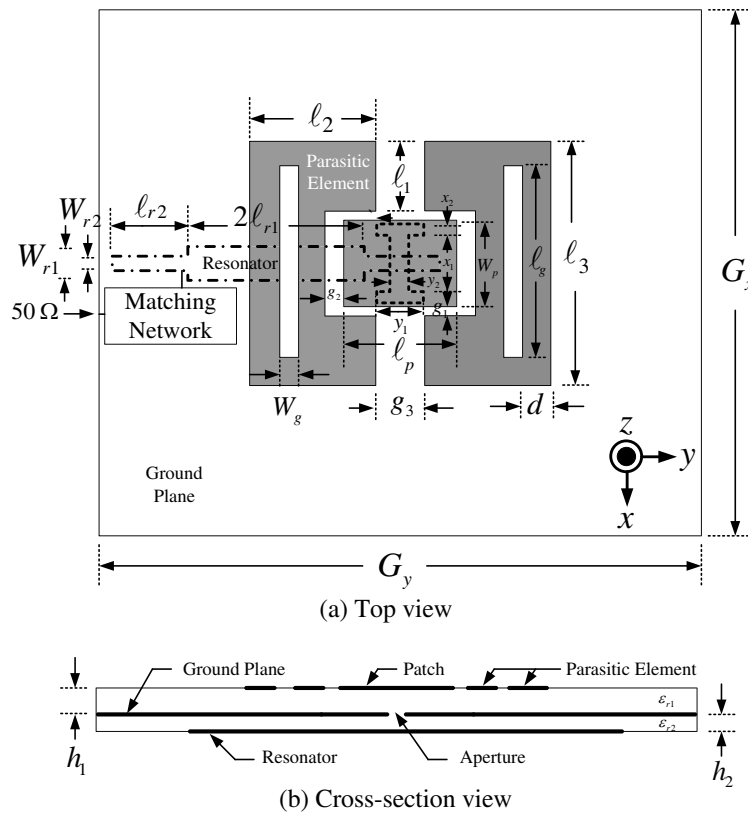


Figure 1: Geometry of the proposed antenna.

Table 1: Dimensions of the proposed antenna.

Parameter	l_p	W_p	l_g	W_g	l_1	l_2	l_3	d	g_1	g_2
(mm)	26	8	18	0.8	5.8	9.4	20	1	0.2	1.2
Parameter	g_3	x_1	x_2	y_1	y_2	l_{r1}	l_{r2}	l	W_{r1}	W_{r2}
(mm)	2	6.6	0.2	4	0.2	7.6	11.7	5	2.6	0.8
Parameter	G_x	G_y	h_1	h_2	-	-	-	-	-	-
(mm)	60	80	0.8	0.4	-	-	-	-	-	-

where I, II and III represent the first, second and third band, respectively. All the above parameters can be extracted by using full-wave simulator HFSS.

3. SIMULATION RESULTS AND DISCUSSION

The simulated return loss and peak gain are shown in Fig. 2. Impedance bandwidths reach 12% (2.19–2.48 GHz), 8.6% (3.41–3.71 GHz), and 4.4% (5.13–5.36 GHz) for 6 dB return loss, respectively. The peak gains in $+z$ -direction are about -4 dBi, -2 dBi, and 4 dBi. There is good selectivity at the band edges in the three bands.

Figure 3 illustrates the simulated far-field radiation pattern at 2.45 GHz, 3.5 GHz, and 5.25 GHz. The broadside radiation patterns in both $E(yz)$ - and $H(xz)$ -planes are shown in the three bands. That is, pure polarization in the same planes. Due to the distance between the two radiating edges of the slot-loaded parasitic elements larger than $\lambda/2$ in the third band, the grating lobes appear in the E -plane. The antenna peak gains in $\theta = 0^\circ$ ($+z$ -direction) are -4 dBi, -2 dBi, and 4 dBi as the same depicted in Fig. 2. The backward radiation is caused by the finite ground effect.

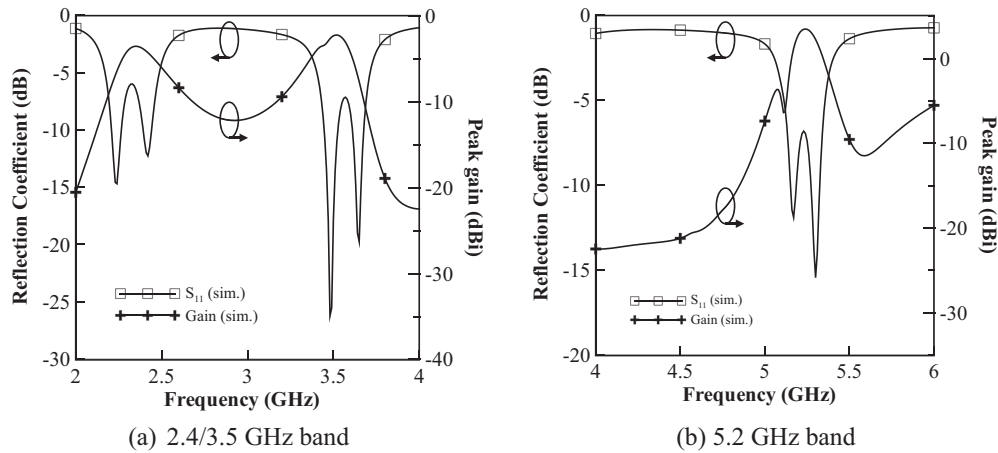


Figure 2: The simulated S_{11} and peak gain in (a) 2.4/3.5 GHz band and (b) 5.2 GHz band.

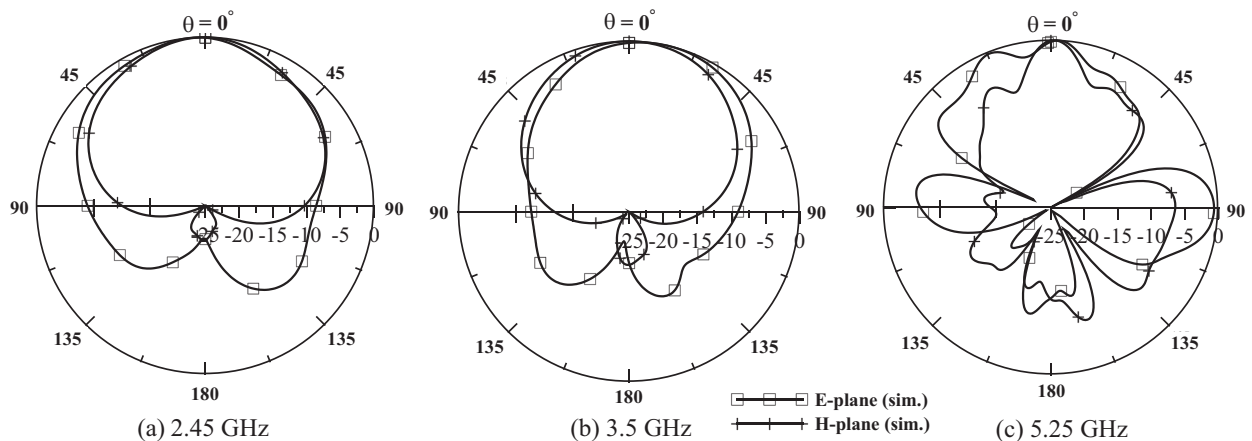


Figure 3: The simulated radiation patterns at (a) 2.45 GHz, (b) 3.5 GHz, and (c) 5.25 GHz.

4. CONCLUSION

A compact triple-band filtering antenna with the same polarization planes is presented. The antenna is fabricated on two different thickness FR4 substrates with the same apertured ground for broadside coupling. The antenna is designed at WLAN/TD-LTE 2.4 GHz, 3.5 GHz, and 5.2 GHz bands with 12%, 8.6%, and 4.4% bandwidth, respectively. There are good selectivity and pure polarization in the same planes in this antenna.

ACKNOWLEDGMENT

This work was supported in part by the National Science Council of Taiwan, under Contract NSC 101-2221-E-009-140-MY3, and by the National Space Organization (NSPO), Taiwan, under Contract NSPO-S-102054.

REFERENCES

1. Chuang C.-T. and S.-J. Chung, "A compact printed filtering antenna using a ground-intruded coupled line resonator," *IEEE Trans. Antennas Propag.*, Vol. 59, No. 10, 3630–3637, Oct. 2011.
2. Lin C.-K. and S.-J. Chung, "A compact filtering microstrip antenna with quasi-elliptic broadside antenna gain response," *IEEE Antennas Wireless Propag. Lett.*, Vol. 10, 381–384, 2011.
3. Lee, Y.-J. and S.-J. Chung, "A compact dual-band filtering microstrip antenna with the same polarization planes," *Proc. Asia-Pacific Microwave Conference (APMC)*, 1178–1180, Kaohsiung, Taiwan, Dec. 2012.
4. Maci, S., G. B. Gentili, P. Piazzesi, and C. Salvador, "Dual-band slot-loaded patch antenna," *IEE Proc. Microw. Antennas Propag.*, Vol. 142, No. 3, 225–232, Jun. 1995.
5. Wi, S.-H., Y.-S. Lee, and J.-G. Yook, "Wideband microstrip patch antenna with U-shaped parasitic elements," *IEEE Trans. Antennas Propag.*, Vol. 55, No. 4, 1196–1199, Apr. 2007.
6. Hong, J. S. and M. J. Lancaster, *Microstrip Filter for RF/Microwave Application*, J. Wiley & Sons, New York, 2001.

A Study of Partial Resonance Control for Edge Elements in a Finite Array

C. I. Kolitsidas and B. L. G. Jonsson

Division of Electromagnetic Engineering, School of Electrical Engineering
KTH Royal Institute of Technology, Stockholm SE-100 44, Sweden

Abstract— This work is focused on developing a method for compensating the edge effects in a mid-sized antenna array. Initially the resonance shift for the E -plane with respect to element position and scan angle was studied in finite \times infinite array of strongly coupled dipoles. We show that geometrical tuning on the outermost elements of the array results in a 40% bandwidth improvement at the edge elements for a 40° scan angle. This improvement is compared with the original edge element behavior and it is achieved with a VSWR < 3 for the worst element. These modifications were based into two observations. Primarily, it was shown that by electrically connecting the edge elements we obtain an improvement in the interelement coupling. By adding stripes to the arms of the edge dipoles we obtained an additional degree of freedom that was used to improve the impedance of the edge elements. This procedure resulted in almost uniform VSWR behavior for all elements up to the edge for the finite \times infinite approach.

1. INTRODUCTION

Future wireless base station are expected to have the ability to support all existing commercial frequency bands as well as new emerging wireless technologies. This implies that the corresponding RF-air interface should be able to cover all these frequency bands simultaneously. Furthermore on top of the classic multiplexing techniques spatial division (SDMA-Spatial Division Multiple Access) needs to be employed in order to achieve maximum network capacity, low interference and power efficiency at a low visual impact device. The latter two are very important since the majority of base stations are used in urban environments.

An implementation possibility for the radiating part of the RF-air interface is a mid-sized, wideband, wide scan angle antenna array. This is an expansion for the “all-in-one antenna” concept as it also employees phased array characteristics. The classic array design procedure starts with the unit cell analysis where the element is considered in an infinitely periodic structure. This design procedure is well known to be efficient for large array antenna systems where most elements are approximately in an infinite periodic environment. The outermost elements in a finite array suffer from “edge effects” which are caused by the finiteness of the array. Applying the same design procedure in the mid-sized arrays is a challenge as most of the elements of the array can be considered to be influenced by the array’s edge effects. The edge effects are mainly caused by the multiple reflections from the edges of the array. The reflections cause variations at the element currents which, in turn, can be translated as a resonance shift at the antenna impedance. The latter was indicated in [1] where an eigencurrent approach was used to characterize the elements in a finite array. The magnitude of the frequency resonance shift also depends on the scanning angle of the array and consequently for large steering the operational bandwidth of the array is shrunked.

Over the last decade it has been shown that strongly coupled elements, [2] are able to provide wideband (5 : 1) performance while keeping a relatively low profile ($\lambda_{high\ freq}/2.5 - \lambda_{high\ freq}/3$). These properties are derived from the fact that a strongly coupled element array can emulate the infinite continuous current sheet concept [3]. The truncated geometry of a finite mid-sized array results in different active reflection coefficients and embedded element patterns for the elements that are closer to the edges. This is due to the difference of the coupling coefficients as the edge elements are subject to reduced coupling contribution.

The circumferential elements in small antenna arrays are susceptible to impedance mismatches that diminishes the total array bandwidth. Methods to compensate “edge behavior” and the associated impedance mismatch include two different approaches. The first approach is to terminate the edge elements to matched resistors. This approach degrades the efficiency of the total array but improves the total bandwidth compared to the simple open circuited structure. Alternatively, tapered excitation can be applied to the structure. Based on this latter approach, a method was developed in [4] using tapered excitation which followed the characteristic mode of the array structure. This method requires apriori knowledge of the characteristic mode of the structure. A

study of edge effects through the coupling coefficients in tapered slot antenna arrays placed in a triangular grid was made in [5]. Finally in [6] a parametric study was performed in order to find the size of the array that must be taken into account so the behavior of inner and edge elements is captured simultaneously.

The present effort is focused on a numerical study of edge effects in a strongly coupled dipole geometry and to develop a compensation procedure by alternating the geometry at the edges of the structure. These geometrical modifications can be optimized in order to keep all elements in the finite array with approximately the same performance as the unit cell.

2. ANTENNA ARRAY DESIGN — INFINITE CASE

Consider closely spaced dipoles as radiating elements. The geometry of the initial unit cell design is depicted in Fig. 1(a). The small interelement gap 2δ , where $\delta = 0.6$ mm, in between the elements serves a dual role. Primarily, it significantly increases the interelement coupling and also introduces a capacitance between the adjacent dipole arms. The latter is required to counteract the inductive behavior of the ground plane as is indicated in [7]. Furthermore the height between the dipoles and the ground plane is chosen to be $\lambda_{high\ freq.}/2.5$ in order to avoid a zero in the broadside direction at the high end of the frequency band. The dipole is fed by a balanced co-planar strip line ($s = 0.3$ mm & $C_\ell = 2$ mm) which is tapered to a wide microstrip line ($w = 11$ mm & $\ell = 9.52$ mm). The unit cell dimensions are chosen $d = \lambda_{high\ freq.}/2$ in order to be below at the grating lobe limit for the fundamental Floquet mode. In order to make an evaluation for the edge element performance a finite \times infinite approach is used as depicted in Fig. 1(b) for the E -edge evaluation. This approach will isolate the behavior along the E -plane of the array. Similarly we can evaluate the behavior of edge elements for the H -plane of the array.

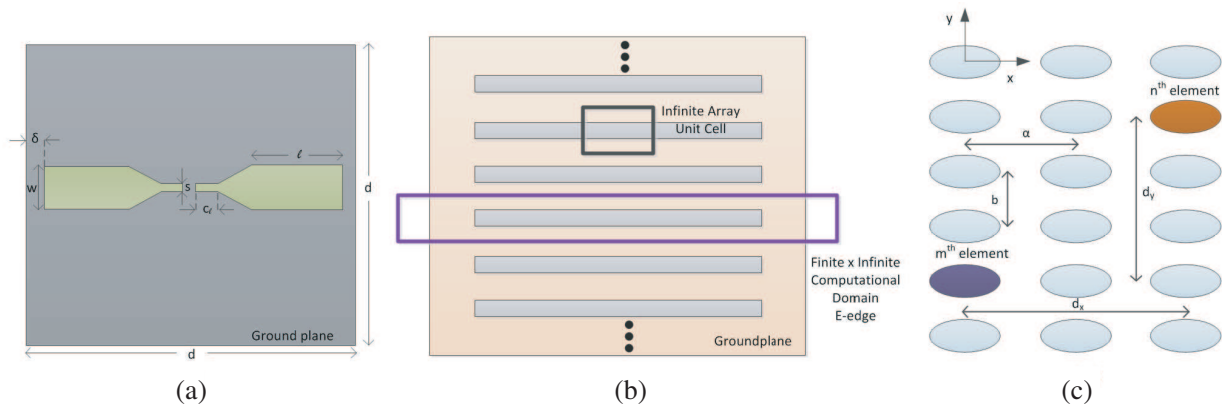


Figure 1: (a) The infinite array unit cell plan. (b) Computational domain for the E -plane in the finite \times infinite approach. (c) Rectangular array lattice.

The active reflection coefficient $\Gamma(\psi_x, \psi_y)$ of the unit cell takes into account all coupling phenomena that occur in the array. For a given planar rectangular lattice under Floquet excitation is given as in [8]:

$$\Gamma(\psi_x, \psi_y) = \sum_{n=-\infty}^{\infty} S_{mn} \exp \left[-j \left(\frac{d_x(m, n)\psi_x}{a} + \frac{d_y(m, n)\psi_y}{b} \right) \right] \quad (1)$$

where all scattering parameters S_{mn} contribute to the final active reflection coefficient. The spatial parameters of the active reflection coefficient are with respect to the Fig. 1(c). In detail, ψ_x and ψ_y are the phase differences between the adjacent elements, in our case $\alpha = b = d$ and $d_x(m, n)$, $d_y(m, n)$ are the relative distances along the x and y direction between the m th and n th element as depicted in Fig. 1(c). Also, as is commonly known, strong mutual coupling can significantly improve the performance of the array as it can suppress undesired resonances in the structure. In Fig. 2(a) is depicted the VSWR for three scan angles at the E -plane for the infinite array. A slight upwards resonance shift is observed for the unit cell behavior as expected. The realized gain behavior in Fig. 2(b) shows the embedded element pattern variation with frequency for the 40° scan angle. All simulations in this study have been carried out in CST Microwave Studio [9].

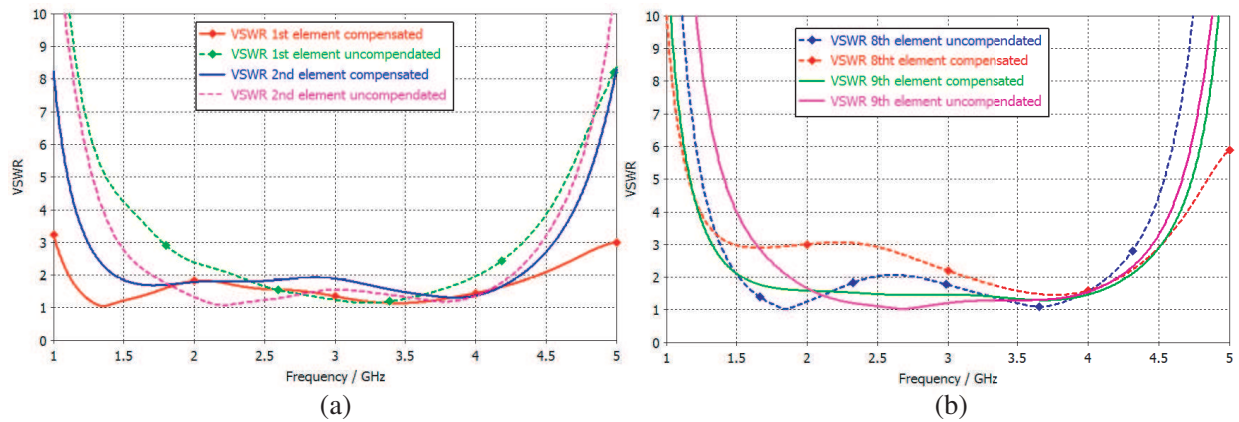


Figure 2: (a) Unit cell VSWR for $\phi = 0^\circ$ & $\theta = 0^\circ$, $\phi = 0^\circ$ & $\theta = 30^\circ$ and $\phi = 0^\circ$ & $\theta = 40^\circ$. (b) Realized gain — embedded element pattern for $\phi = 0^\circ$ & $\theta = 40^\circ$ at $f = 2.5$ GHz and $f = 4$ GHz.

3. EDGE EFFECT COMPENSATION

The finite \times infinite approach reduces the computational domain and enables a study of the edge behavior. The effects that are occurring along the E -plane are more severe when compared to the H -plane edge effects. Thus, only E -plane effects are considered in this study. Toward this end we consider a linear array of 9 elements created by repeating the unit cell as depicted in Fig. 3(a). We impose periodic boundary conditions (PBC) along the y -direction and perfect matching layer (PML) in every other direction. Scanning the beam off-broadside up to 40° we study the variations with respect to the active reflection coefficient for the edge elements. As was observed from Equation (1) the mutual coupling plays a significant role in the array's element behavior. As we move towards the edges, the coupling is weaker for the edge elements. This behavior is visible as a variation in the results for the active reflection coefficient. Based on that fact the first step in order to compensate the edge effects is to increase the coupling between the edge adjacent elements. This can be achieved by electrically connecting the last two elements of the linear array as depicted in Fig. 3(b).

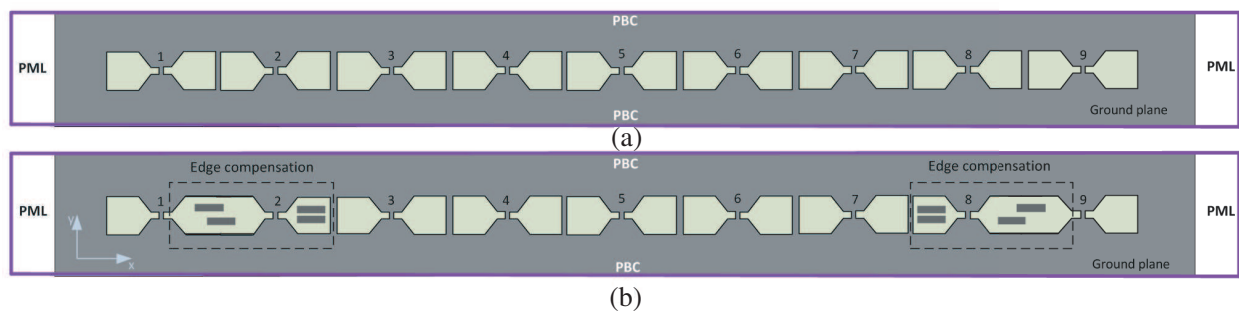


Figure 3: (a) 9 element array. (b) 9 element array with edge compensation.

In addition to the electrical connection, strips are placed in the connected long dipole arm and to the other arm from the second to last element. This introduces a partially controllable second order resonance at the edge elements of the structure. The dimensions and the placement of the strips are optimized in respect to usable bandwidth of the edge elements of the array using a genetic algorithm. The simulations also indicated that the elements 3–7 have a stable performance.

The final results after the optimization procedure for the outermost elements are depicted in Fig. 4 and a comparison is made between the uncompensated (Fig. 3(a)) and the compensated (Fig. 3(b)) array. In Fig. 4(a) the VSWR for the elements 1 & 2 is presented. The results indicate that the usable bandwidth of these elements is increased by 40%. In Fig. 4(b) the VSWR for the elements 8 & 9 is depicted. In this case the usable bandwidth of the edge element is improved by 40% keeping a VSWR < 2 while the 8th element has degraded its performance for the lower half of the bandwidth where the VSWR is < 3 . This is considered within the acceptable limit.

Finally it is important to note that this geometry modifications did not affected the radiation

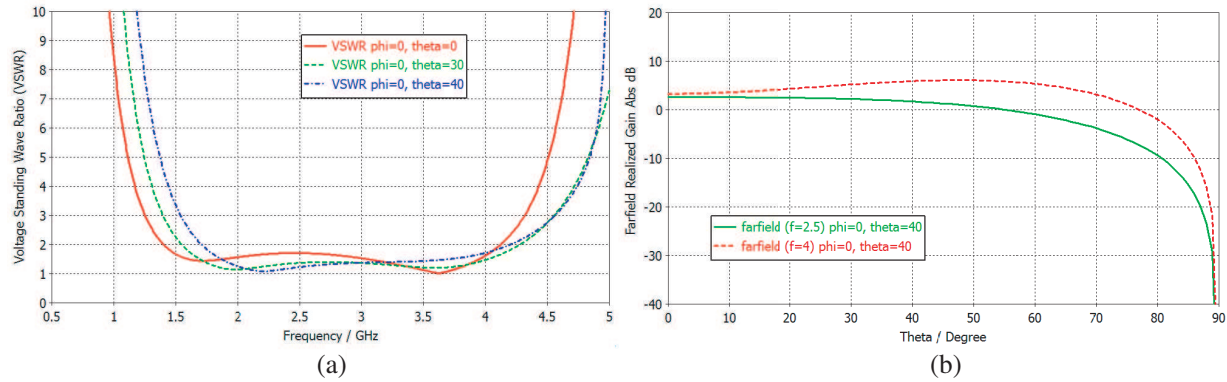


Figure 4: Edge element VSWR comparison between the uncompensated and compensated array. (a) elements 1 & 2. (b) elements 8 & 9.

pattern of the array as the spacial sampling was kept intact. Only a small increase (1 dB) was observed at the second side lobe of the array.

4. CONCLUSION

A method was introduced to compensate the edge effects in a mid-sized array. This study modeled the resonance behavior for a finite \times infinite array of strongly coupled dipoles for different scanning angles to provide an insight in the element behavior. It was shown that geometrical modifications at the outermost elements in an array comprised by 9 strongly coupled elements can improve the bandwidth when the array is scanned up to 40° . This approach increased the usable bandwidth $VSWR < 3$ at the edge elements by 40%. These modifications are divided into two categories. First, in order to improve the interelement coupling at the edge elements of the array an electrical connection at the outermost elements was implemented. Second, as the purpose is to maintain the same impedance at the elements a partial controllable second order resonance was added in the structure by adding stripes at the edge dipoles. It is expected that further improvement can be achieved when matching a matching network is included at the radiating structure. This approach can be also be implemented to other planar structures.

ACKNOWLEDGMENT

This work was supported by The Swedish Governmental Agency for Information Systems (VINNOVA) within the VINN Excellence center, Chase stage III through the research project Next Generation Antenna Arrays.

REFERENCES

1. Bekers, D. J., J. L. Eijndhoven, and A. G. Tijhuis, "An eigencurrent approach of finite arrays of electromagnetically characterized elements," *Radio Sci.*, Vol. 44, RS2S90, 2009.
2. Munk, B. A., *Finite Arrays and FSS*, Wiley-Inter-Science, New York, 2003.
3. Wheeler, H., "Simple relations derived from a phased array antenna made of an infinite current sheet," *IEEE Trans. on Antennas and Propag.*, Vol. 13, No. 4, 506–514, July 1965.
4. Tzanidis, I., K. Sertel, and J. Volakis, "Characteristic excitation taper for ultra wideband tightly coupled antenna arrays," *IEEE Trans. on Antennas and Propag.*, Vol. 60, 1777–1784, 2012.
5. Ellgardt, A. and M. Norgren, "A study of edge effects in triangular grid tapered slot arrays using coupling coefficients," *Radio Sci.*, Vol. 4, RS2005, 2010.
6. Holter, H. and H. Steyskal, "On the size requirement for finite phased array models," *IEEE Trans. on Antennas and Propag.*, Vol. 50, No. 6, 836–840, June 2002.
7. Kolitsidas, C. I and B. L. G. Jonsson, "Investigation of compensating the ground plane effect through array's interelement coupling," *Proceedings of 7th European Conference on Antennas and Propagation (EuCAP)*, 1221–1224, Gothenburg, Sweden, April 2013.
8. Bhattacharyya, A. K., *Phased Array Antennas*, Wiley-Inter-Science, New York, 2005.
9. CST *Microwave studio* [®] 2012.

Broadband, High Gain, See-through Antenna for WiFi, WiMax and LTE 2600 Radio Networks

P.-A. Garcia^{1,2}, T. Razban¹, A. Chousseaud¹, and E. Motta Cruz²

¹LUNAM, Université de Nantes, IETR UMR 6164, rue Christian Pauc, Nantes 44306, France

²Bouygues Telecom, 76 rue des Français Libres, Nantes 44263, France

Abstract— This paper demonstrates the feasibility of broadband, high gain, see-through antennas for radio networks where low visual impact is needed. A single linear polarization, proximity-coupled stacked patch antenna is proposed. Borosilicate glass is used for substrates, but its high permittivity ($\epsilon_r = 4.7$) leads to inefficient unilayer patch antennas, with a gain lower than 5 dBi. However, adding two stacked superstrates of this material allows a substantial increase of the gain, up to 7.3 dBi. The latter configuration offers the opportunity to add resonant patches, so the bandwidth can be broadened. A 2.1–2.8 GHz bandwidth (30% relative bandwidth) is proposed, with VSWR < 1.5 , covering WiMax (around 2.3 GHz and 2.5 GHz), WiFi (2.4–2.483 GHz) and LTE 2600 (2.5–2.69 GHz) frequency bands. We use a simple, low cost process, based on Printed Circuit Board (PCB) technology, to create thick (17 μm) meshed conductors on glass substrate. The gridded mesh is made of 50 μm wide horizontal and vertical strips, with a 450 μm interspacing. This resolution is high enough to make the conductive patterns quite discreet with respect to human eye acuity, with an estimated optical transparency greater than 80%.

1. INTRODUCTION

The increasing number of coexisting standards in cellular networks (GSM, GPRS, UMTS and most recently LTE) leads to a higher number of base-station antennas. Adding antennas is becoming a complex task, especially in urban zones, due to the resulting visual pollution, which is less and less accepted.

Optically transparent antennas are a seductive solution to further reduce visual impact, as they can be integrated on structures that have not been used yet, such as glass windows. Meshed conductors printed on flat substrates can be used to achieve efficient antennas [1], but they remain visible, due to the heterogeneous nature of the mesh. It explains why they are commonly called see-through antennas as opposed to antennas manufactured with homogeneous, transparent conductors [2]. Also, simplest implementation of printed antennas can't exhibit both high gain and large bandwidth, so a tradeoff must be found [3].

Recent advances in manufacturing technologies allow mesh geometries that are dense enough to make them almost homogeneous with respect to human eye acuity. Mesh lines with widths thinner than 20 μm can be made with microelectronics technology [4, 5], but this solution is costly, with limited antenna sizes (approximately 400 mm \times 400 mm, which is the current maximum size of wafers used in microelectronics industry). Also, given reference don't show whether it is possible to print on both faces of the substrate, which can be useful to implement compact or multilayer antennas.

This paper presents a broadband, high gain, see-through antenna, designed with the commercial software HFSS, and manufactured with optically transparent materials, with Printed Circuit Board (PCB) technology, whose availability allows to reduce costs (85% lower than microelectronics technology for a given printed surface) especially thanks to the ability to simultaneously print on both sides of the substrate with a very good alignment between layers ($\pm 50 \mu\text{m}$ as a standard value).

A first section presents antenna design and some simulation results obtained with HFSS. Then, antenna manufacturing process is presented, and finally, measured results are discussed in comparison with simulation results.

2. ANTENNA DESIGN AND SIMULATION RESULTS

The design starts from an unilayer patch antenna. Borosilicate glass is chosen for its low loss factor ($\tan \delta = 0.0068 @ 1.5 \text{ GHz}$ [5]) and for its thermal resistance, which is needed for any metal deposition process. A 3.3 mm thickness is chosen to ensure mechanical resistance. However, Borosilicate high permittivity ($\epsilon_r = 4.7 @ 1.5 \text{ GHz}$ [5]) leads to low isotropic gains ($< 5 \text{ dBi}$). However, it has been noticed that adding two superstrates of Borosilicate glass allows a substantial

increase of the gain, up to 8 dBi, thanks to the radome effect [6]. The latter configuration gives the opportunity to add resonant patches, so the bandwidth can be broadened.

The next step consists of finding patches dimensions, and an initial configuration, illustrated in Figure 1 that allows to observe a “loop combination” in the Smith Chart [7]. The resulting impedance, shown in Figure 2(a) is then matched to a $50\ \Omega$ coaxial cable with a quarter-wave section (Figure 2(b)). The coaxial cable is connected via a drilled piece of brass, since SMA connectors are not suited to 3.3 mm thick substrates. The resulting impedance is shown in Figure 2(c). A 2.1–2.8 GHz bandwidth with $VSWR < 1.5$ is expected, with up to 8 dBi gain in this frequency range (additional results will be shown with the measured ones). It is important to note that the transparent conductor material is simulated as an homogeneous one, like any unmeshed antenna simulation setup, which enables a fair simulation time.

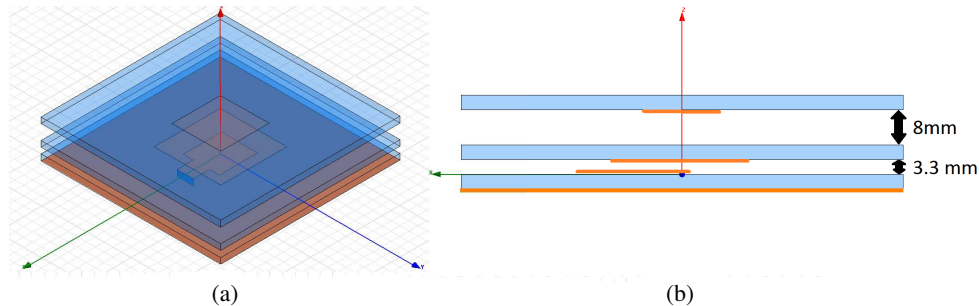


Figure 1: Overview of the unmatched stacked patch antenna designed with HFSS, in (a) isometric view and (b) side view.

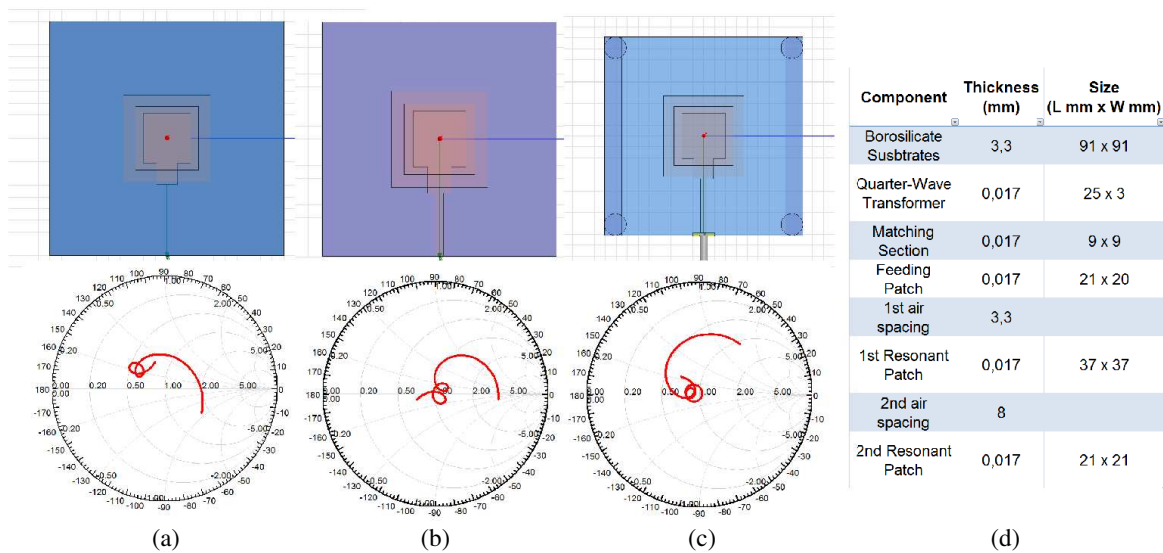


Figure 2: Matching process: (a) initial case with the “loop combination”, (b) adding the quarter wave-transformer, (c) connecting the coaxial cable and adding dielectric spacers, (d) antenna final dimensions.

3. ANTENNA MANUFACTURING

We investigated the use of glass in PCB technology to create glass PCBs, that are etched to create see-through antennas. The process can't be fully described here due to patent pending. Figure 3(a) shows the glass panel after copper etching. We made unmeshed, opaque antennas in order to measure them as references. For meshed antennas, the squared mesh has $50\ \mu\text{m}$ wide lines, with a $450\ \mu\text{m}$ pitch, which yields an estimated transparency greater than 80%. To minimize transmission losses, the meshed feed line pitch is reduced to $250\ \mu\text{m}$. The glass panel is then cut to extract the useful parts using waterjet cutting technology (Figure 3(b)).

We can observe that the panel is no longer transparent after PCB etching, due to additional surface roughness. A coating, such as some varnish layers, can suppress this roughness, as long as protecting copper layers. Figure 4 shows the results obtained in one substrate (a) and for the whole see-through antenna (b), where optical transparency is altered due to a poor coating of the meshed ground plane.

The main advantage of this solution is its cost, which is about 85% lower than microelectronics technology, thanks to the ability to handle larger surfaces ($650\text{ mm} \times 460\text{ mm}$ and even beyond), and the ability to simultaneously print on both sides of the substrate.

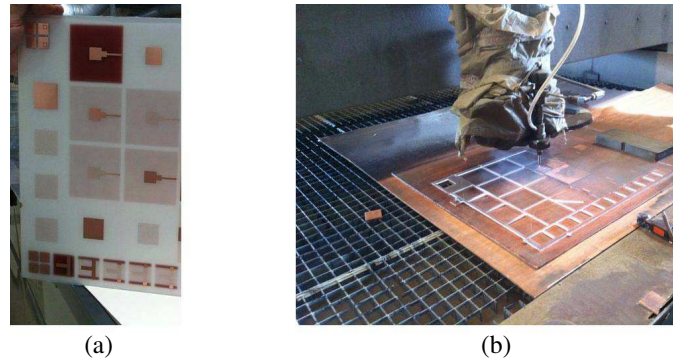


Figure 3: View of a rough PCB glass panel (a) after copper etching, and (b) during waterjet cutting process.



Figure 4: (a) Optical rendering of a varnished glass substrate, and (b) the resulting see-through antenna.

4. MEASUREMENT RESULTS AND DISCUSSION

Impedance, Gain and Radiation Pattern measurements have been done, in comparison with an unmeshed antenna as a reference. Figure 5 shows the impedance results: A good agreement is obtained with simulation, and there is no substantial difference between measurements. Additional ripples are found due to additional components, such as coaxial connectors for cables, that are not included during simulation, and add some electrical delay.

Gain measurement and simulation results are presented in Figure 6. Although the see-through antenna gain is nearly the same as the reference antenna gain, measured values are lower than expected initially in simulation.

However, these losses are not due to unefficient substrates or conductors, but due to the effective width of the main lobe, which is larger than expected, especially in the E plane. That causes a 1 dB mean loss of directivity, close to the ratio [Simulated E plane radiation pattern integral/Measured E plane radiation pattern integral] (Figures 7 and 8). The E plane radiation pattern may have been affected by the presence of the coaxial cable during the measurement. It is also important to note that the chosen mesh geometry, with lines parallel to E and H planes, has no effect on the antenna polarization. The second source of loss is the impedance mismatch, which is not negligible at 2 GHz according to Figure 5, where 8 dB return loss yields 0.8 dB transmission losses.

The described losses can be avoided to achieve higher gains, as high as an opaque antenna, but special care must be taken regarding the radiation pattern control, especially the influence of the cable position. However, we successfully achieved up to 7.3 dBi of Gain along with 30% relative bandwidth for $S_{11} < 14\text{ dB}$ (or $VSWR < 1.5$), that makes one of the first manufactured low-cost broadband high-gain see-through antenna.

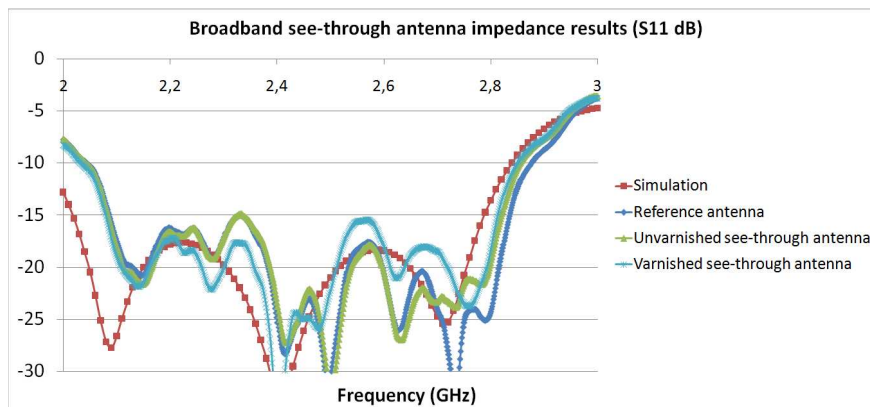


Figure 5: Impedance measurement results for reference antenna, unvarnished and varnished see-through antennas, compared to simulation results.

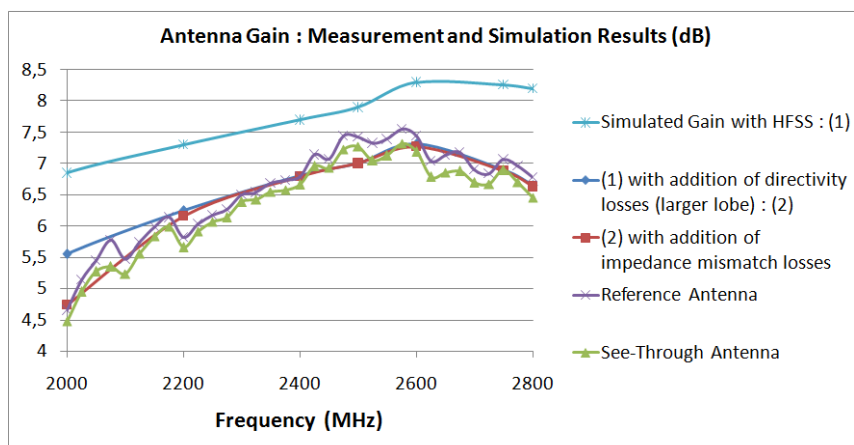


Figure 6: Gain measurement results for reference and see-through antennas, compared to simulation results.

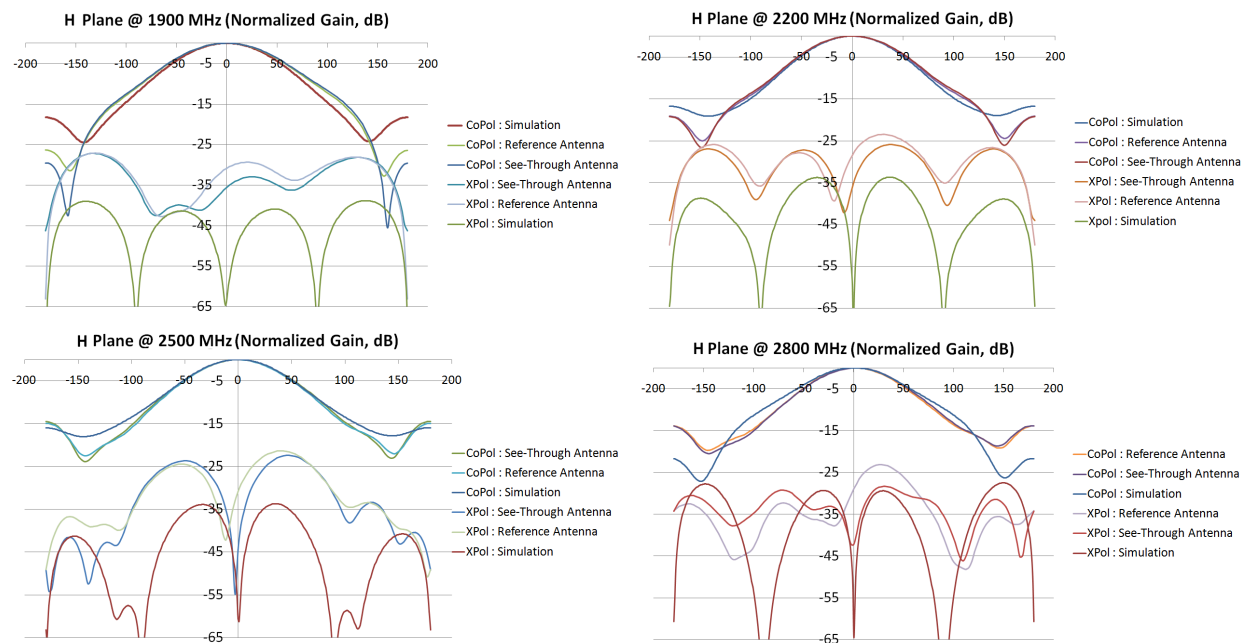


Figure 7: Antennas radiation patterns in H plane (sorted by decreasing levels at Theta = ± 180).

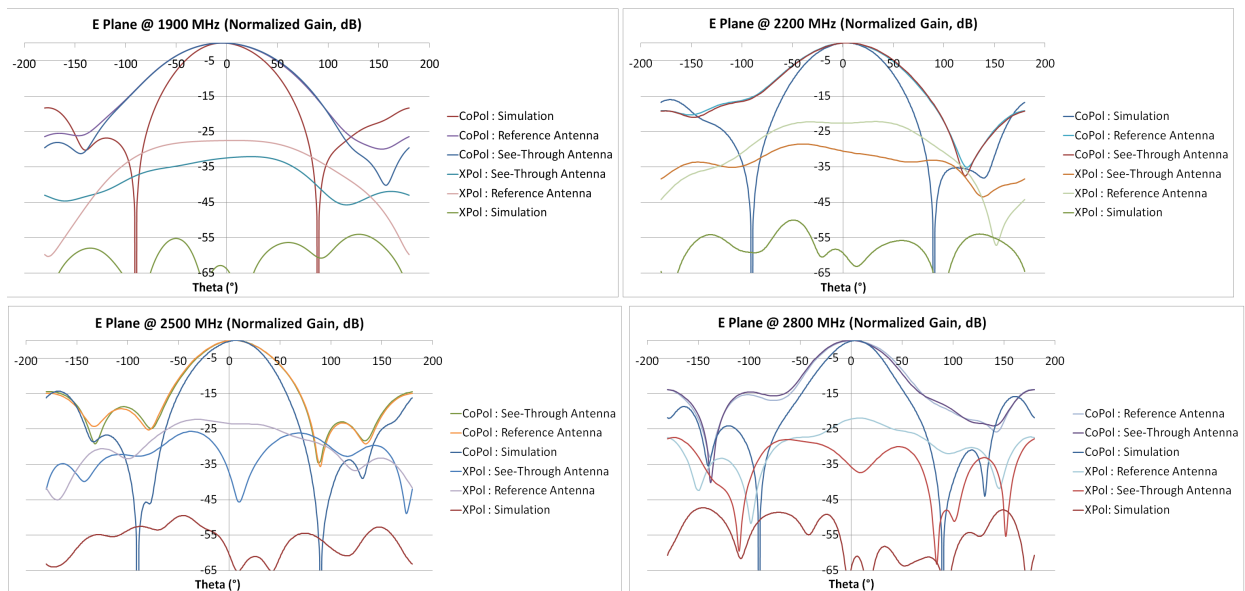


Figure 8: Antennas radiation patterns in E plane (sorted by decreasing levels at $\Theta = \pm 180$).

5. CONCLUSION

We demonstrated the feasibility of high performance, low-cost, see-through antennas. It has been shown that meshing doesn't affect the bandwidth, nor the gain, nor the polarization in a substantial way. The proposed manufacturing process can be used to make more complex antennas. Measurement results highlight some mechanisms that affect the gain, such as lobe broadening, or mismatch losses, where special care has to be taken to minimize them.

REFERENCES

1. Ito, K. and M. Wu, "See-through microstrip antennas constructed on a transparent substrate," *IEEE Seventh International Conference on Antennas and Propagation (ICAP)*, 1991.
2. Chopra, K. L., S. Majora, and D. K. Pandya, "Transparent conductors — A status review," *Thin Solid Films*, Vol. 102, No. 1, 1–46, 1983.
3. Balanis, C. A., *Antenna Theory, Analysis and Design*, 2nd Edition, John Wiley & Sons Inc., New York, 1997.
4. Hautcoeur, J., F. Colombel, X. Castel, M. Himdi, and E. Motta Cruz, "Radiofrequency performances of transparent ultra-wideband antennas," *Progress In Electromagnetics Research C*, Vol. 22, 259–271, 2011.
5. Hautcoeur, J., "Conception d'un matériau transparent et conducteur efficace. Application aux antennes-panneaux transparentes pour les réseaux radio cellulaires mobiles," Ph.D. Thesis, University of Rennes 1, France, 2011.
6. Alexopoulos, N. G. and D. R. Jackson, "Fundamental superstrate (cover) effects on printed circuit antennas," *IEEE Trans. on Ant. & Prop.*, Vol. 32, No. 8, 807–816, 1984.
7. Targonski, K. L., R. B. Waterhouse, and D. M. Pozar, "Design of wide-band aperture-stacked patch microstrip antennas," *IEEE Trans. on Ant. & Prop.*, Vol. 46, No. 9, 1245–1251, 1998.

A Fractal Microstrip Array Antenna with Slots Feeding Network for DTV Reception

Jianhua Zhou, Bingyang Liang, Baiqiang You, Qing Liu, and Xingjing Yan
 Department of Electronic Engineering, Xiamen University, Xiamen, Fujian 361005, China

Abstract—Based on FR4 board, a Cantor fractal microstrip array antenna of multi-ports with slots feeding network is designed for DTV reception. The performances of the proposed antenna are optimized with different feeding networks and array matrices. The simulated results show that applying Cantor fractal elements, the mutual coupling strength between the array elements is decreased effectively and the antenna is miniaturized. In further, the antenna samples with designed $4 \times 4/8 \times 8$ planar arrays and slot coupling array were fabricated. The measurement agreed very well with the simulation, indicating a relatively high efficiency and relative bandwidth of 3.08% and 3.81% respectively.

1. INTRODUCTION

For receiving antennas, as an important part of satellite digital TV (DTV) systems, the parabolic structure is more mature, with two kinds of being fed and partial feed. Along with the development of wireless communication technology, better performances are required for receiving antennas, like light weight, small size, simple produce, low cost, easy to conformal and wide band. Microstrip array antennas are developed on the basis of microstrip antennas [1], with the advantages of small size, simple structure and easy to conformal over the traditional dipole array. But the mutual coupling is easily produced due to so small spacing between array elements, and also the application of microstrip array antennas in satellite communication is limited by some features, like small bandwidth and relatively lower radiation efficiency. In high frequency transmission system, the dielectric loss in microstrip transmission line and the internal coupling and radiation in feed network all will lead to a sharp decline in the efficiency [2]. Such a loss in the transfer efficiency will also be increasingly severe along with the expansion of feed network [3, 4]. In 2005, KoHan Lu et al. proposed a power divider network for the array feeding, which improved the antenna gain and efficiency in a certain extent [5]. In 2008, Mohd Shah et al designed a dual-polarized microstrip array antenna using rectangular patch of 45° tilt angle, which achieved a high gain [6]. In 2009, Mohamed H. Awida designed a Ku-band array antenna using the substrate-integrated waveguide (SIW) technology, indicating a relatively high efficiency and wide bandwidth of greater than 70% and 9%, respectively [7]. In 2010, Reza Azadegan described a microstrip array antenna of good performance, having 26.5 dBi gain at 12 GHz [8]. Kim and Jaggard were the first to use the fractal in the design of antenna array system in 1985, discussing a low-sidelobe random array [9]. The radiation characteristics of deterministic fractal array configurations, such as Cantor linear arrays and Sierpinski carpet planar arrays, have been considered in Refs. [10, 11], developing rapid algorithms for use in efficient radiation pattern computations and adaptive beam forming. A newly developed thinning algorithm is presented which may be employed to reduce considerably both the overall physical size and the total number of elements in a synthesized multiband array [12]. Ref. [13] introduces a type of nature-based design process that applies a specially formulated genetic algorithm (GA) technique to evolve optimal polyfractal array layouts.

Since the large size, inconvenient installation of current receiving antennas widely used in satellite TV systems, the paper proposed a Cantor fractal microstrip array antenna for Ku-band application, focusing on the roles of fractal array in miniaturization and mutual coupling reduction.

2. ANTENNA DESIGN

Based on a FR4 board with the relative permittivity of 2.2, we designed a 4×4 uniform planar array with Cantor elements, as shown in Fig. 1(a). The size of FR4 substrate is designated as $W = 60$ mm, $L = 58$ mm and $h = 0.8$ mm. For the consideration of gain, beam width and sidelobe level, the distance d between elements along two directions is determined as 0.57λ , that is 13.7 mm. The front and side view of a Cantor element is given as in Fig. 1(b), with $W = 14$ mm, $L = 12$ mm, $h = 0.8$ mm, $a = 6.528$ mm, $b = 5.382$ mm, $a_1 = 1.632$ mm, $b_1 = 1.346$ mm. According to power division principle, we also designed an 8×8 uniform planar array with Cantor elements, as shown

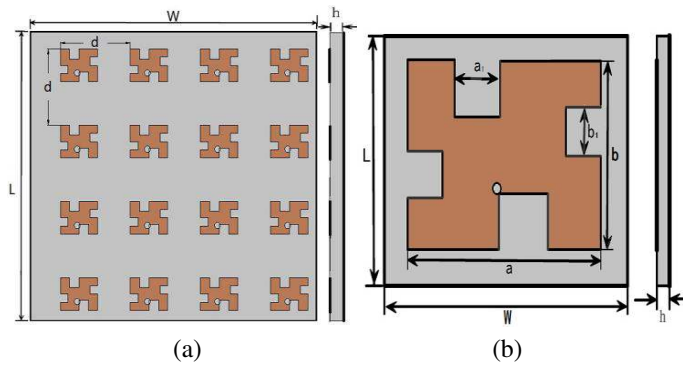


Figure 1: Layout of designed (a) 4×4 Cantor array antenna and (b) Cantor elements array antenna.

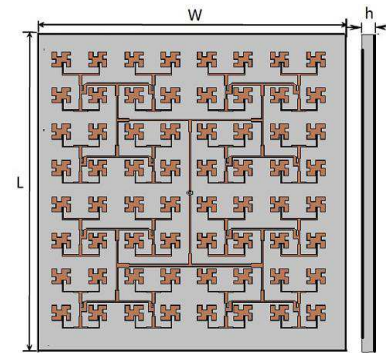


Figure 2: Layout of designed 8×8 Cantor array antenna.

in Fig. 2. The size of FR4 substrate is designated as $W = 120$ mm, $L = 124$ mm and $h = 0.8$ mm, with other dimensions for Cantor elements all same as that of 4×4 array.

Further, the slot coupling technology has been applied to the fractal microstrip array for wider bandwidth and unlimited feeding design. As shown in Figs. 3(a) and (b), a 4×4 H-type coupling slot array and a feeding network are loaded on the Cantor microstrip array as the intermediate layer and the background layer apartly, in which $h_1 = 0.8$ mm, $h_2 = 0.2$ mm.

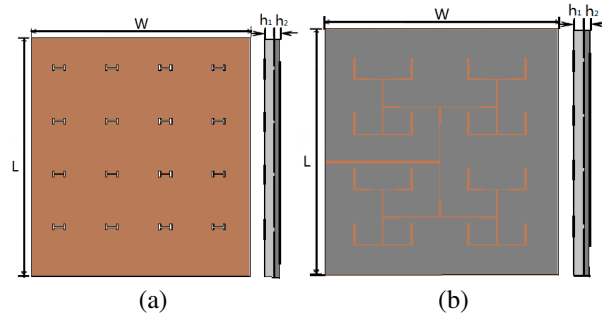


Figure 3: Layout of designed 4×4 array antenna. (a) Coupling slots and (b) feeding network.

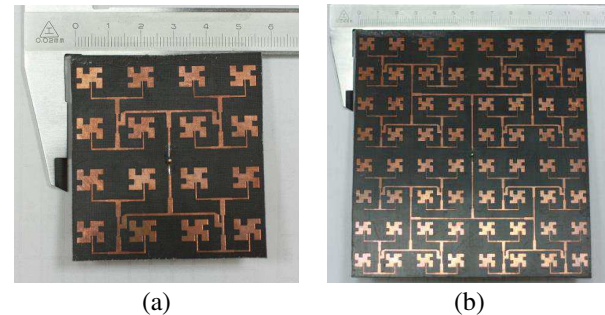


Figure 4: The photos of manufactured samples with (a) 4×4 and (b) 8×8 planar array.

3. SIMULATION AND EXPERIMENT

The photographs of the manufactured Cantor array antennas, $4 \times 4/8 \times 8$ planar array and slot coupling array are shown in Figs. 4(a) ~ (c), respectively. All the manufactured antennas were measured in a microwave semi-anechoic chamber, in which only signals of $0 \sim 180^\circ$ with 15° for the side-feeding array were tested while the omnidirectional normalized pattern of $0 \sim 360^\circ$ for the slot coupling feeding array were obtained. As shown in Fig. 5(a), the measured center frequency for 4×4 planar array is 12.02 GHz, a little drift to lower frequency relative to the simulation, with passband of 11.93 ~ 12.11 GHz and relative bandwidth of 3.08%. As shown in Fig. 5(b), the measured center frequency for 8×8 planar array is 12.06 GHz, an obvious drift to lower frequency relative to the simulation, with passband of 11.78 ~ 12.24 GHz and relative bandwidth of 3.81%. It indicates that the more complicated feeding network for 8×8 gives much greater influence on the array performance. From the measured field patterns of 8×8 array antenna with Cantor elements depicted as in Fig. 6, there is a little angular drift in the maximum radiation direction for E plane due to the feeding microstrip lines.

According to the simulated return loss values as shown in Fig. 7, the designed 4×4 H-type coupling slot array presents wider bandwidth of 11.60 ~ 12.24 GHz, with relative bandwidth of 5.38% which increases nearly doubled by comparing to the designed 4×4 Cantor array with power-division feeding network. While the measured center frequency for 4×4 slot coupling array is 12.06 GHz, an obvious drift to higher frequency relative to the simulation, with passband of 11.78 ~ 12.24 GHz and relative bandwidth of 3.81%. The total passband width is up to 0.66 GHz, which fully proves the broaden role of slot coupling.

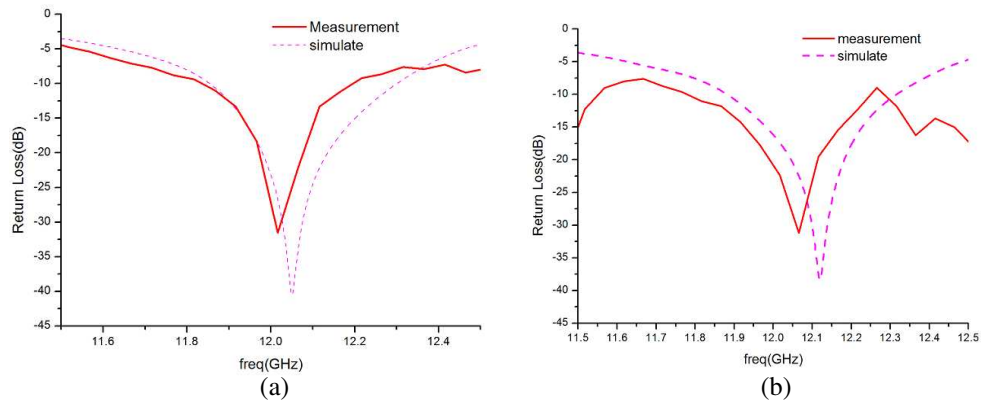


Figure 5: The return loss values of (a) 4×4 and (b) 8×8 planar array antenna with Cantor elements.

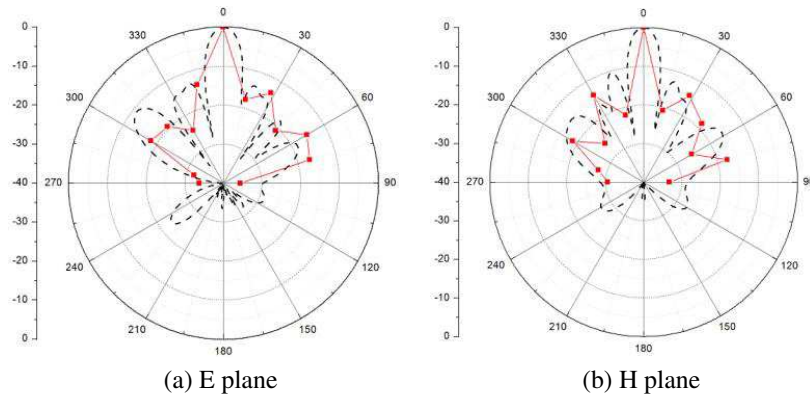


Figure 6: The field patterns of 8×8 array antenna with Cantor elements.

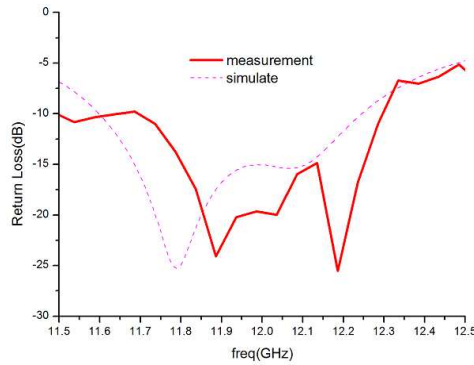


Figure 7: The return loss values of 4×4 slot coupling array antenna.

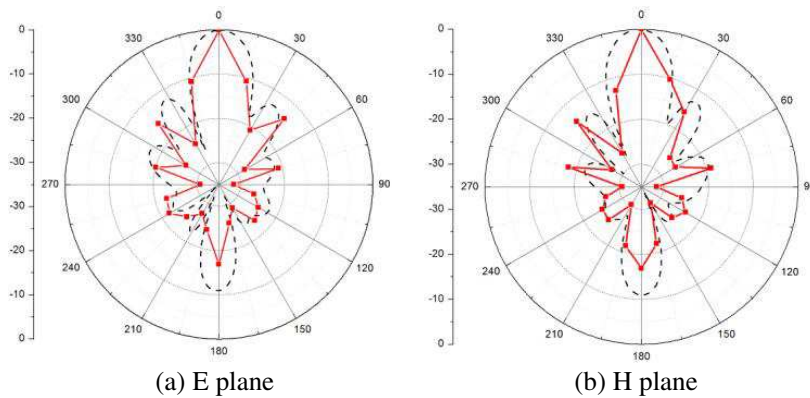


Figure 8: The field patterns of 4×4 slot coupling array antenna.

The simulated field patterns, as depicted in Figs. 8(a) and (b), seem in more accordance with the multi-ports feeding, without coupling and element placement interference from feeding microstrip lines. Since the patch is going to reflect a little in the slot coupling, there would be increased backward sidelobes which made the front to back ratio grow up. On the other hand, the efficiency of its feeding network gets higher, expressed by the peak gain up to 18.05 dBi, because of the independent modulation of the radiation and feeding. From the measured field patterns as in the figure, the result of the mainlobe agrees with the simulation very well, while that of the sidelobes with similar trend and decreased magnitude seems better than the simulation.

4. CONCLUSION

A Cantor fractal array with slot matching feed network has been successfully designed for DTV receiving systems in Ku band. By feeding network supplied with slot coupling technology, the radiation energy of proposed antenna system could be controlled well, which has been proved with 4×4 and 8×8 feeding system. In our design, fractal antenna elements can reduce the outlet of antenna with little radiation loss and the coupling energy be modulated by slots with separated feeding network, providing more reliable multi-ports. It is also found that with the combination of all these techniques, the gain, bandwidths and phase center could be improved further significantly.

ACKNOWLEDGMENT

This work is supported by Fujian Provincial Key Sci-Tech Project (2010HZ0004-1) of China.

REFERENCES

1. Sabban, A., "Applications of MM wave microstrip antenna arrays," *Signals, Systems and Electronics*, Vol. 07, 119–122, 2007.
2. Busuioc, D., A. Borji, M. Shahabadim, et al., "High efficiency antenna array with optimized corporate hybrid feed," *Proc. Int. Symp. Antennas and Propagation Society*, Vol. 7, 1503–1506, 2006.
3. Gottwald, G. and W. Wiesbeck, "Influence of mutual coupling and feed network on directivity and gain [Array Patch Antennas]," *Proc. Eight. Int. Conf. on Antennas and Propagation*, 761–763, 1993.
4. Horng, T.-S. and N. G. Alexopoulos, "Corporate feed design for microstrip arrays," *IEEE Trans. Antennas*, Vol. 41, No. 13, 1615–1624, 1993.
5. Lu, K. H. and T.-N. Chan, "Circularly polarized array antenna with corporate-feed network and series feed elements," *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 10, 3288–3292, 2005.
6. Mohd Shah, M. S. R. and M. K. Suaidi, "Design of 1×2 , 1×4 , and 2×2 dual polarization microstrip array antenna," *Proceedings of IEEE 6th National Conference on Telecommunication Technologies*, 113–116, 2008.
7. Awida, M. H., "Substrate-integrated waveguide Ku-band cavity backed 2×2 microstrip patch array antenna," *Antennas and Wireless Propagation Letters*, Vol. 8, 1054–1056, 2009.
8. Azadegan, R., "A Ku-band planar antenna array for mobile satellite TV reception with linear polarization," *IEEE Transactions on Antennas and Propagation*, Vol. 58, No. 6, 2097–2101, 2010.
9. Kim, Y. and D. L. Jaggard, "The fractal random array," *Proc. IEEE*, Vol. 74, No. 9, 1278–1280, 1986.
10. Werner, D. H., R. L. Haupt, and P. L. Werner, "Fractal antenna engineering: The theory and design of fractal antenna arrays," *IEEE Antennas and Propagation Magazine*, Vol. 41, No. 5, 37–59, 1999.
11. Werner, D. H. and S. Ganguly, "An overview of fractal antenna engineering research," *IEEE Antennas Propagation Magazine*, Vol. 45, No. 1, 3857, 2003.
12. Werner, D. H., M. A. Gingrich, and P. L. Werner, "A self-similar fractal radiation pattern synthesis technique for reconfigurable multiband arrays," *IEEE Transactions on Antennas and Propagation*, Vol. 51, No. 7, 1486–1498, 2003.
13. Petko, J. S. and D. H. Werner, "An autoployploidy-based genetic algorithm for enhanced evolution of linear polyfractal arrays," *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 3, 583–593, 2007.

Wavelength-dependent Switching of a Nonlinear Optical Loop Mirror

O. Pottiez¹, B. Ibarra-Escamilla², and E. A. Kuzin²

¹Centro de Investigaciones en Óptica, Loma del Bosque 115
Col. Lomas del Campestre, Leon, Gto. 37150, Mexico

²Departamento de Optica, Instituto Nacional de Astrofísica Óptica y Electrónica
L. E. Erro 1, Puebla, Pue. 72000, Mexico

Abstract— We propose and study a Nonlinear Optical Loop Mirror (NOLM) whose switching characteristic is wavelength-dependent. The device operation relies on a polarization imbalance between counter-propagating beams in the loop, which is created by the insertion of a high-birefringence fiber (HiBiF). Wavelength-dependent switching arises from the fact that the polarization of one of the beams is altered by the HiBiF in a wavelength-dependent manner. The setup is proposed for wavelength-confined switching and for simultaneous regeneration of two wavelength signals presenting uneven power levels.

1. INTRODUCTION

The Nonlinear Optical Loop Mirror (NOLM) [1] is a versatile and low-cost device whose potential has long been recognized for applications like ultrafast optical switching and signal processing (regeneration, wavelength conversion, demultiplexing, etc.) [2, 3]. Switching is obtained through the optical Kerr effect, when the symmetry in the Sagnac interferometer structure of the device is broken in some way. Although conventional schemes are power-asymmetric (by the use of an asymmetric coupler, or by inserting asymmetrically a gain or loss element in the loop), a birefringence asymmetry (e.g., a wave retarder (WR) inserted asymmetrically in the loop) is also effective, taking advantage of the polarization dependence of the Kerr nonlinearity. Such a polarization-imbalanced NOLM, including a 50/50 coupler and a quarter-WR, was proposed and studied previously [4], and the enhanced flexibility of its switching characteristic, which can be controlled through the WR angle and input polarization, makes it very promising for applications like high-quality data regeneration [5] or passive mode locking of fiber lasers [6].

With most NOLM schemes, the nonlinear response does not depend on the signal wavelength. For some applications however, in particular in Wavelength Division Multiplexing (WDM) systems, wavelength-dependent operation is highly desirable. Unfortunately, very few wavelength-sensitive NOLM designs can be found in the literature. In [7], wavelength dependence was introduced in a NOLM through the insertion of a chirped grating. Wavelength-confined switching (switching at the grating wavelength only) was demonstrated experimentally. In the frame of signal regeneration, efforts are being made to concentrate functions such as amplitude equalization and regeneration of different channels into a single device, for obvious cost-saving reasons. Most of the fiber-based schemes that have been proposed so far for these tasks are based on nonlinear spectral broadening followed by offset filtering [8, 9]. These schemes usually conserve a certain degree of parallelism, as at some point the channels are demultiplexed and processed separately, so that the number of some components (fiber sections, filters, attenuators, delay lines etc.) remains proportional to the number of channels. Therefore, in a sense simultaneous regeneration is not performed through a truly single device.

2. DEVICE DESCRIPTION AND MODELING

As mentioned before, the polarization-imbalanced NOLM relies on a birefringent element, a WR, to provide switching, taking advantage of the polarization-dependent nonlinear phase shift. If the WR is replaced by a piece of high birefringence fiber (HiBiF), consisting of a large number of beat lengths, the phase shift between the x and y components of light propagating through it is much larger than 2π and is given by

$$\Delta\phi = \frac{2\pi}{\lambda_0} \Delta n L_{HB} = 2\pi \frac{L_{HB}}{L_B} \quad (1)$$

where Δn is the refractive index difference, L_B the beat length, L_{HB} the HiBiF length and λ_0 the wavelength. At a slightly different wavelength $\lambda_1 = \lambda_0 + \Delta\lambda$ ($\Delta\lambda \ll \lambda_0$), the phase shift $\Delta\phi$

changes by an amount $\delta = -2\pi L_{HB}\Delta\lambda/(L_B\lambda_0)$, which is a substantial fraction of 2π and thus alters significantly polarization at the HiBiF output if L_{HB}/L_B is large. This wavelength shift thus alters the polarization imbalance in the NOLM loop, and therefore its switching characteristic becomes wavelength-dependent.

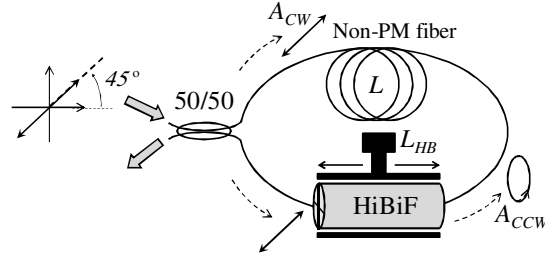


Figure 1: Scheme under study. The case of linear input polarization is illustrated.

The proposed scheme is illustrated in Fig. 1. It includes a 50/50 coupler, a length L of non-polarization-maintaining (non-PM) fiber and a piece of HiBiF inserted asymmetrically in the loop. For proper NOLM operation, polarization imbalance has to be maintained, so that the ellipticity of each beam must be conserved as it propagates down the non-PM fiber, a condition hardly met in practice due to the residual birefringence of the fiber and its typically long length. A nearly isotropic behavior can be easily obtained in practice, however, by applying twist to the fiber [10]. Finally, the HiBiF birefringence (and thus $\Delta\phi$) can be slightly tuned in a range of $\sim \pi$ (e.g., mechanically or thermally), allowing transmission adjustments.

In the continuous-wave approximation, the NOLM behavior can be easily modeled using the Jones matrix formalism. In the circular $[C^+; C^-]$ base, the matrix of the HiBiF (whose axes are set parallel and perpendicular to the plane of the loop) is given by

$$HB = \begin{bmatrix} \cos(\Delta\phi/2) & j \sin(\Delta\phi/2) \\ j \sin(\Delta\phi/2) & \cos(\Delta\phi/2) \end{bmatrix}. \quad (2)$$

In the weak nonlinearity limit [4], the Kerr effect does not alter light ellipticity, so that the Stokes parameter, $A = (|C^+|^2 - |C^-|^2)/P$ (where $P = |C^+|^2 + |C^-|^2$ is the optical power) remains constant. In this case, the nonlinear phase shift in the non-PM loop is conveniently accounted for by a Jones matrix, which writes as

$$F_{CW/CCW} = \begin{bmatrix} \exp[j\gamma(1 - \frac{1}{3}A_{cw/ccw})\frac{P_{in}}{2}L] & 0 \\ 0 & \exp[j\gamma(1 + \frac{1}{3}A_{cw/ccw})\frac{P_{in}}{2}L] \end{bmatrix}, \quad (3)$$

where P_{in} is the NOLM input power, and the subscripts CW and CCW refer to clockwise and counter-clockwise beams, respectively. A_{cw} is equal to the Stokes parameter at the NOLM input (assuming that the coupler does not alter polarization) and A_{ccw} is calculated from the Jones vector of the CCW beam at the HiBiF output. Finally, each crossing of the 50/50 coupler is taken into account by a multiplication by $1/\sqrt{2}$ for the CW field and by $j/\sqrt{2}$ for the CCW field. For a given Jones vector E_{in} at the NOLM input, the output fields of the CW and CCW beams, $E_{out,cw} = 1/2HB.F_{cw}E_{in}$ and $E_{out,ccw} = -1/2F_{ccw}.HB.E_{in}$, respectively, are easily calculated and summed to yield the total output field E_{out} . The NOLM transmission is finally obtained by $T = |E_{out}|^2/|E_{in}|^2$.

3. RESULTS AND DISCUSSION

Figure 2 shows the power-dependent NOLM transmission at four different wavelengths and for two particular input polarization states, namely circular and linear at 45° with respect to the HiBiF axes. A 500-m long high-nonlinearity fiber (nonlinear coefficient $\gamma = 10/\text{W/m}$) is chosen as the non-PM fiber, and the HiBiF consists of 50.25 beat lengths (at $\lambda_0 = 1550 \text{ nm}$). It is clear from the figures that the NOLM transmission is strongly wavelength-dependent. At wavelength λ_3 , the transmission even vanishes ($T = 0$ for any power), which means that the signal at that wavelength is completely reflected back to the input port. This behavior can be useful for demultiplexing. Indeed, assuming for example a dual-wavelength signal at λ_0 and λ_3 , and adjusting the peak power

of the signal at λ_0 to the switching power $P_\pi(\lambda_0) = 3.77$ W, the signals at λ_0 and λ_3 are transmitted and reflected, respectively.

Except when $\Delta\phi = \pi/2$ or 0, the curves of Figs. 2(a) and (b) are different, which means that, as wavelength is varied, the transmission evolves in a way that depends on input polarization. The case of linear input polarization at 45° is particularly interesting: as observed in Fig. 2(b), a wavelength shift only affects switching power, whose value is then given by $P_\pi = 6\pi/(\gamma L|\sin\Delta\phi|)$, whereas minimal and maximal transmission values $T(0) = 0$ and $T(P_\pi) = 1$ are unaltered. Moreover, through slight mechanical/thermal adjustments of the HiBiF, the ratio between switching powers at two given wavelengths can be readily tuned. This configuration is applied to amplitude regeneration of two signals at λ_0 and λ_2 having different power levels. If the average peak powers at each wavelength are in the same ratio than the values of P_π of the corresponding transmission curves, and if total input power is adjusted to set the peak powers slightly above the values of P_π (at ~ 4.4 W and ~ 8.8 W, respectively), then the intensity limiting effect of the NOLM significantly reduces amplitude fluctuations (regeneration of the “1” level). Moreover, the low transmission at low power for any wavelength also allows regeneration of the “0” level in both signals. Such a regeneration is illustrated in Fig. 3 (channels were shifted temporally to avoid nonlinear interactions between them in the loop). This figure was obtained resolving a system of coupled nonlinear partial differential equations to take into account the small dispersion of the high-nonlinearity fiber (0.3 ps/nm/km with 0.02 ps/nm²/km dispersion slope) and the twist-induced group delay difference (assuming a twist of 2 turns/m), however the results do not differ significantly from those of the continuous-wave approach presented here.

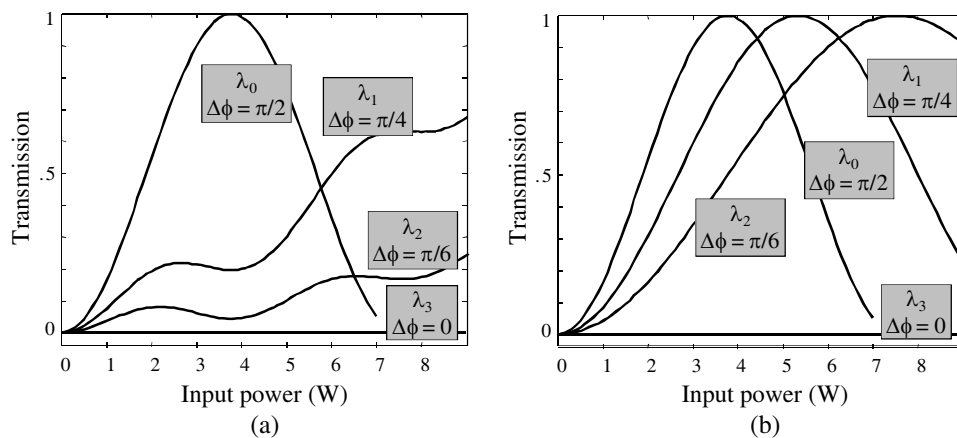


Figure 2: NOLM transmission at four different wavelengths $\lambda_0 = 1550$ nm, $\lambda_1 = 1553.85$ nm, $\lambda_2 = 1555.14$ nm and $\lambda_3 = 1557.71$ nm, in the case of (a) circular input polarization and (b) linear input polarization making a 45° angle with the HiBiF axes.

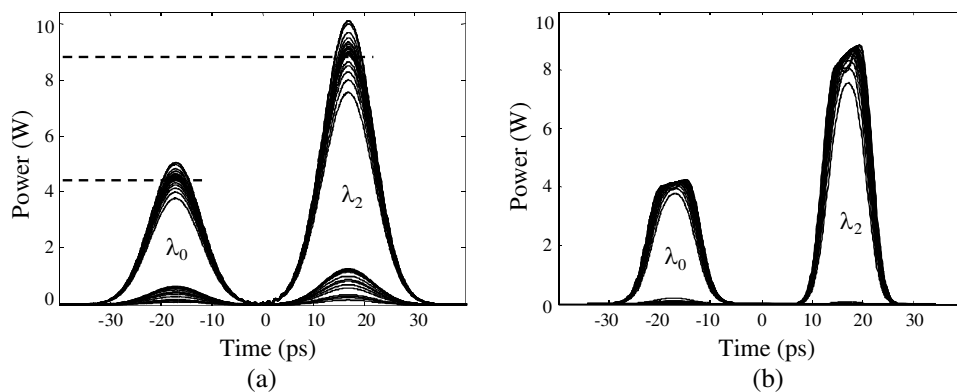


Figure 3: Eye diagrams at the NOLM (a) input and (b) output of a 2-wavelength signal at $\lambda_0 = 1550$ nm and $\lambda_2 = 1555.14$ nm with a power ratio of 1 : 2, assuming that each channel consists of 12-ps pulses with 10% Gaussian amplitude noise as the “1” level and 10% ghost pulses with 50% noise for the “0” level. Transmission curves of Fig. 2(b) (linear input polarization at 45°) were used.

4. CONCLUSION

In this work we propose and study a novel wavelength-sensitive polarization-imbalanced NOLM. It consists of a symmetric coupler, a loop of twisted non-PM fiber and a HiBiF inserted asymmetrically in the loop. Switching is due to the polarization imbalance introduced by the HiBiF and wavelength-dependent operation is obtained as a consequence of the wavelength-dependent phase shift of the HiBiF. The case of linear input polarization at 45° with the HiBiF axes is particularly attractive, as the transmission curve is then a simple sinusoidal function of input power whose switching power only varies with wavelength. The setup is considered for wavelength demultiplexing and for amplitude regeneration of two wavelength channels with uneven power levels. This work contributes to pave the way for the design of a range of novel wavelength-sensitive ultrafast processing devices for WDM systems.

ACKNOWLEDGMENT

This work is funded by CONACYT grant #130681.

REFERENCES

1. Doran, N. J. and D. Wood, "Nonlinear optical loop mirror," *Opt. Lett.*, Vol. 13, 56–58, 1988.
2. Boscolo, S., S. K. Turitsyn, and K. J. Blow, "Nonlinear loop mirror-based all-optical signal processing in fiber-optic communications," *Opt. Fiber Technol.*, Vol. 14, 299–316, 2008.
3. Sotobayashi, H., C. Sawaguchi, Y. Koyamada, and W. Chujo, "Ultrafast walk-off-free nonlinear optical loop mirror by a simplified configuration for 320-Gbit/s time-division multiplexing signal demultiplexing," *Opt. Lett.*, Vol. 27, 1555–1557, 2002.
4. Kuzin, E. A., N. Korneev, J. W. Haus, and B. Ibarra-Escamilla, "Theory of nonlinear loop mirrors with twisted low-birefringence fiber," *J. Opt. Soc. Am. B*, Vol. 18, 919–925, 2001.
5. Pottiez, O., E. A. Kuzin, B. Ibarra-Escamilla, F. Gutierrez-Zainos, U. Ruiz-Corona, and J. T. Camas-Anzueto, *IEEE Photon. Technol. Lett.*, Vol. 17, 154–156, 2005.
6. Pottiez, O., R. Grajales-Coutiño, B. Ibarra Escamilla, E. A. Kuzin, and J. C. Hernandez-Garcia, "Adjustable noise-like pulses from a figure-eight fiber laser," *Appl. Opt.*, Vol. 50, E24–E31, 2011.
7. Pattison, D. A., P. N. Kean, W. Forysiak, I. Bennion, and N. J. Doran, "Bandpass switching in a nonlinear-optical loop mirror," *Opt. Lett.*, Vol. 20, 362–364, 1995.
8. Vasilyev, M. and T. I. Lakoba, "All-optical multichannel 2R regeneration in a fiber-based device," *Opt. Lett.*, Vol. 30, 1458–1460, 2005.
9. Kouloumentas, Ch., P. Vorreau, L. Provost, P. Petropoulos, W. Freude, J. Leuthold, and I. Tomkos, "All-fiberized dispersion-managed multichannel regeneration at 43 Gb/s," *IEEE Photon. Technol. Lett.*, Vol. 20, 1854–1856, 2008.
10. Tanemura, T. and K. Kikuchi, "Circular-birefringence fiber for nonlinear optical signal processing," *J. Lightwave Technol.*, Vol. 24, 4108–4119, 2006.

Complexified Spherical Waves and Their Sources in the Physical Space

A. M. Tagirdzhanov¹ and A. P. Kiselev^{1,2,3}

¹St. Petersburg State University, Russia

²Steklov Mathematical Institute St. Petersburg Department, Russia

³Institute of Problems of Mechanical Engineering, Russia

Abstract— We address spherical waves complexified by a complex shift in a coordinate of the point source. These waves have been studied since the early 1970s in both time-harmonic and non-time-harmonic cases as exact localized solutions of the wave equation. We deal with the fundamental mode described by $u = \frac{f(\theta_*)}{R_*}$, where $R_* = \sqrt{x^2 + y^2 + (z - ia)^2}$, $a > 0$ is a free positive constant, $\theta_* = R_* - ct$ is a complex phase and $f(\theta_*)$ is an arbitrary function describing the waveform. Such a function satisfies the inhomogeneous wave equation $u_{xx} + u_{yy} + u_{zz} - c^{-2}u_{tt} = F$ with a certain source function $F = F(x, y, z, t)$, which is a generalized function supported by a 2D surface in the real 3D physical space. Here, $c > 0$ is the constant wave speed. The function F is dependent on the waveform f as well as on the definition of the branch of the square root in the “complex distance” R_* . Unlike several earlier studies, in which sources in the complex space were discussed, we focus on explicitly finding the source function F in the real physical space for a rather arbitrary waveform f . For different definitions of the branch of the square root, we present an example of the waveform describing a Gaussian-localized wave packet.

1. INTRODUCTION

The classical spherical wave is a solution to the wave equation having the form [1]

$$u = \frac{f(\theta)}{R}, \quad \theta = R - ct, \quad (1)$$

where $R = \sqrt{x^2 + y^2 + z^2}$ and the *waveform* f is an arbitrary function of θ . Function (1) satisfies

$$\Delta u - c^{-2}u_{tt} = -4\pi\delta(x)\delta(y)\delta(z)f(-ct), \quad \Delta u = u_{xx} + u_{yy} + u_{zz}, \quad (2)$$

where $c > 0$ is a constant wave speed.

The complexified spherical wave is obtainable by a complex shift $z \rightarrow z - ia$, $a > 0$,

$$R \rightarrow R_* = \sqrt{\rho^2 + (z - ia)^2}, \quad \rho = \sqrt{x^2 + y^2}. \quad (3)$$

We come up with

$$u = \frac{f(\theta_*)}{R_*}, \quad \theta_* = R_* - ct, \quad (4)$$

where $f(\theta_*)$ is an arbitrary function of a complex argument. The *complexified distance* R_* is a multivalued function and its rigorous definition requires introducing a branch cut. Therefore, u has a jump and satisfies the *inhomogeneous* wave equation

$$\Delta u - c^{-2}u_{tt} = F(x, y, z, t). \quad (5)$$

The *source function* F is a generalized function localized on a certain 2D surface \mathcal{S} in the real physical 3D space \mathbb{R}^3 . This surface depends on the choice of the cut for R_* .

The above complexification procedure was first proposed for the time-harmonic case (i.e., for the waveform $f(\theta) = e^{ik\theta} = e^{ikR - i\omega t}$ where $k = \omega/c$ is the wave number, and $\omega > 0$ is the frequency) by Izmet'ev [2] and Deschamps [3]. They were interested in finding an exact solution showing the Gaussian beam localization, such as described earlier by an approximate approach (see, e.g., a review [4]). For the non-time-harmonic case, several waveforms describing beam-like and packetlike propagation were considered in [5, 6]. These wavefields were not strongly localized with respect to transverse variables x, y .

Here, we aim, first, to calculate F for a general waveform f and, second, to present an example of a waveform exhibiting a Gaussian packetlike behavior. We use the approaches of [7] and [8].

2. SOURCE FUNCTION

For any choice of the branch cut, R_* has a jump on a certain axisymmetric surface $\mathcal{S} \subset \mathbb{R}^3$ with the boundary $\mathcal{C} = \{\rho = a, z = 0\}$. The surface \mathcal{S} is a support [9] of a generalized function F in (5), see [7] for more detail. We focus on two characteristic examples of the cut, which give dissimilar types of wavefield behavior. The *beam choice* [6] of the branch is characterized by

$$\mathcal{S} = \{\rho \geq a, z = 0\}, \quad R_* > 0 \text{ for } \rho > a, z = +0. \quad (6)$$

The *source choice* [6] is characterized by

$$\mathcal{S} = \{\rho \leq a, z = 0\}, \quad R_* > 0 \text{ for } \rho > a, z = 0. \quad (7)$$

In the time-harmonic case, the source function F was presented in [7] for a general cut. Using a similar approach, we find the function F for a general waveform and a general cut. In the case of beam choice (6), the result is as follows

$$F = F_1 \delta'(z) + F_2 \delta(z), \quad (8)$$

where

$$F_1 = \frac{f(\theta_0^+) + f(-\theta_0^-)}{r_0} H(\rho - a), \quad F_2 = -ia \left(\frac{f'(\theta_0^+) - f'(-\theta_0^-)}{r_0^2} - \frac{f(\theta_0^+) + f(-\theta_0^-)}{r_0^3} \right) H(\rho - a) \quad (9)$$

and $\theta_0^\pm = r_0 \mp ct$ with $r_0 = R_*|_{z=+0} = \sqrt{\rho^2 - a^2}$. Here,

$$H(\rho - a) = \begin{cases} 1, & \rho \geq a, \\ 0, & \rho < a \end{cases}$$

is the Heaviside function. At $\rho = a$, F_2 is non-integrable and, to consider it as a generalized function [9], we need a regularization procedure similar to that in [7].

We omit the result for a general cut.

3. GAUSSIAN WAVE PACKET

Let

$$f(\theta_*) = \exp(2ka(1 - \sqrt{2Q})), \quad Q = 1 - \frac{i\theta_*}{2a}, \quad (10)$$

where k is a free parameter and the branch of \sqrt{Q} is specified by the condition $\text{Re}\sqrt{Q} \geq 0$. Under the condition

$$ka \gg 1, \quad (11)$$

this waveform describes a highly localized Gaussian wave packet. The idea of using function (10) was inspired by [8], where the Bateman-type solutions to the homogeneous wave equation were considered.

We present the corresponding asymptotic expressions for u in the paraxial area,

$$\rho \ll |z - ia|, \quad k\rho^4 \ll |z - ia|^3, \quad (12)$$

and in the far-field,

$$R = \sqrt{\rho^2 + z^2} \gg a, \quad R \gg ka^2, \quad (13)$$

for choices of the branch cut, made in accordance with (6) and (7). For the beam choice (6), the wavefield behavior is similar to that found in [8]. The most interesting is the source choice (7), when the wavefield behavior crucially differs from that in [8].

3.1. The Beam Choice

In the paraxial area described by (12), the near-peak behavior of u is given by

$$u \approx \frac{1}{z - ia} \exp \left(ik \left(\alpha + \frac{\beta\rho^2}{4(z^2 + a^2)} \right) - \frac{k\alpha^2}{4a} - \frac{ka\rho^2}{2(z^2 + a^2)} \right), \quad (14)$$

where $\alpha = z - ct$, $\beta = z + ct$. Under the far-field condition (13), we obtain

$$u \approx \frac{1}{R} \exp \left(\frac{ikA}{\sqrt{1 + 2 \sin^2(\chi/2)}} - 2ka \left(\sqrt{1 + 2 \sin^2(\chi/2)} - 1 \right) - \frac{k}{4a} \frac{A^2}{(1 + 2 \sin^2(\chi/2))^{3/2}} \right), \quad (15)$$

for $z > 0$ and

$$u \approx -\frac{1}{R} \exp \left(\frac{-ikB}{\sqrt{1 + 2 \cos^2(\chi/2)}} - 2ka \left(\sqrt{1 + 2 \cos^2(\chi/2)} - 1 \right) - \frac{k}{4a} \frac{B^2}{(1 + 2 \cos^2(\chi/2))^{3/2}} \right), \quad (16)$$

for $z < 0$. Here, $A = R - ct$, $B = R + ct$ and χ defined by $\cos \chi = z/R$ is the inclination angle.

For $z > 0$ and $ka \gg 1$, u is Gaussian-localized at $|\chi| \ll 1$, where (15) reduces to

$$u \approx \frac{1}{R} \exp \left(ikA - \frac{kA^2}{4a} - \frac{ka\chi^2}{2} \right). \quad (17)$$

For $z < 0$ and $ka \gg 1$, u is Gaussian-localized at $|\pi - \chi| \ll 1$, where (16) reduces to

$$u \approx -\frac{1}{R} \exp \left(-ikB - \frac{kB^2}{4a} - \frac{ka(\pi - \chi)^2}{2} \right). \quad (18)$$

The paraxial formulas (14), (17)–(18) are much alike to those presented in [8] for the Bateman-type solution. The wavefields corresponding to cases (14) and (17) are shown in Figure 1.

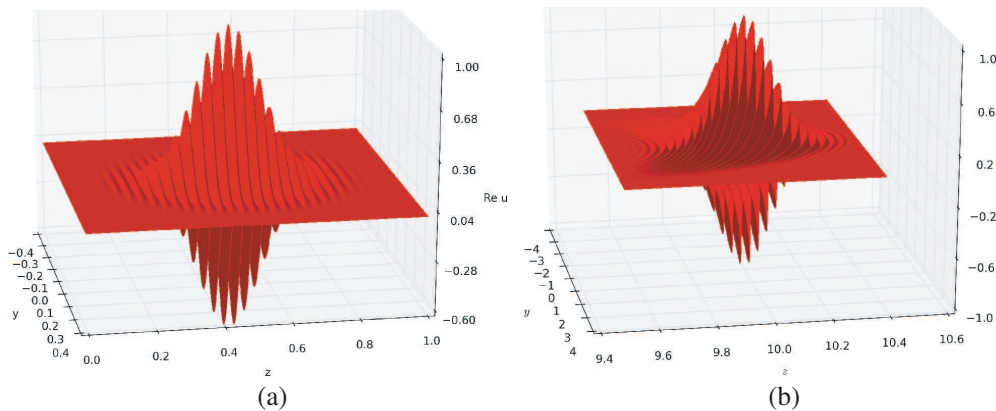


Figure 1: $\text{Re } u$ for $ka = 200$ and $x = 0$, at two instants of time (a) $t = 0.5$, (b) $t = 10$. The wavefields are identical for the cases of beam choice and source choice.

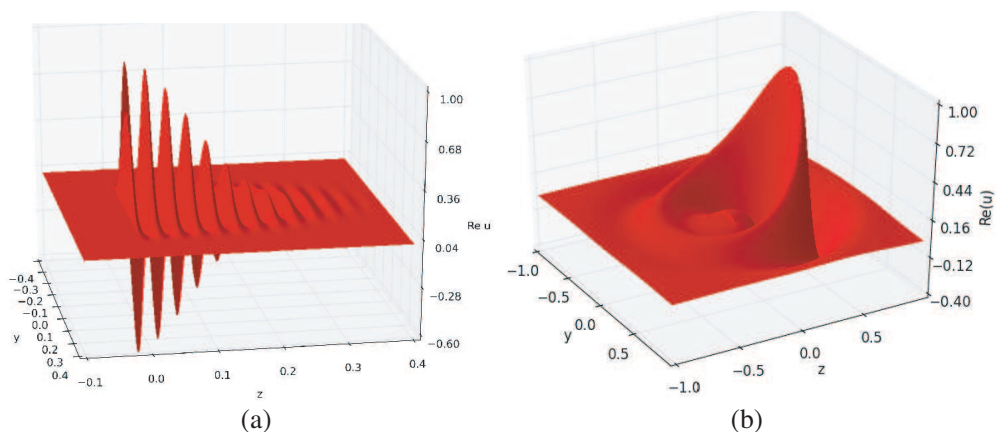


Figure 2: $\text{Re } u$ at $x = 0$ for the source choice of the cut at (a) $ka = 200$, $t = 0$, and (b) $ka = 2$, $t = 0.5$.

3.2. The Source Choice

In the case of source choice (7), the wavefield behavior for $z > 0$ is the same as above. For $z < 0$, the wavefield behavior crucially differs from (14). Instead of (14), we have

$$u \approx \frac{1}{z - ia} \exp\left(-ik \frac{\beta \rho^2}{4(a^2 + z^2)} - ka - \frac{k\beta^2}{4a}\right), \quad z < 0, \quad (19)$$

which has a Gaussian localization only in the longitudinal direction but is exponentially small at $ka \gg 1$.

In the far-field, the wavefield behavior is described by (15) for both positive and negative values of z . Expression (16) demonstrates the decay of the wavefield with increasing χ . Wavefield (19) is shown at Figure 2(b).

4. CONCLUSION

Unlike the sourceless Bateman-type solutions [4, 8], the complexified spherical waves are generated by some field sources that can be associated with wavefield distributions of certain antennas. We are planning to continue the discussion of the similarities and dissimilarities of wavefields (4) with the same waveform, but different choices of the cut, with the Bateman-type solution. Also, we aim to extend the above analysis to higher-order modes.

ACKNOWLEDGMENT

The support from Russian Foundation for Basic Research grant 11-01-00407-a is acknowledged.

REFERENCES

1. Courant, R. and D. Hilbert, *Methods of Mathematical Physics*, Vol. 1, Wiley-Interscience, New York, 1953.
2. Izmet'sev, A. A., "One parameter wave beams in free space," *Radiophys. Quant. Electron.*, Vol. 13, 1062–1068, 1970.
3. Deschamps, G. A., "Gaussian beam as a bundle of complex rays," *Electron. Lett.*, Vol. 7, 684–685, 1971.
4. Kiselev, A. P., "Localized light waves: Paraxial and exact solutions of the wave equation (a review)," *Optics and Spectroscopy*, Vol. 102, No. 4, 603–622, 2007.
5. Heyman, E. and L. B. Felsen, "Complex-source pulsed-beam fields," *J. Opt. Soc. Am. A*, Vol. 6, No. 6, 806–817, 1989.
6. Heyman, E. and L. B. Felsen, "Gaussian beam and pulsed-beam dynamics: Complex-source and complex-spectrum formulations within and beyond paraxial asymptotics," *J. Opt. Soc. Am. A*, Vol. 18, No. 7, 1588–1611, 2001.
7. Tagirdzhanov, A. M., A. S. Blagovestchenskii, and A. P. Kiselev, "'Complex source' wavefields: Sources in real space," *J. Phys. A: Math. Theor.*, Vol. 44, No. 42, 425203, 2011.
8. Kiselev, A. P. and M. V. Perel, "Highly localized solutions of the wave equation," *J. Math. Phys.*, Vol. 41, No. 4, 1934–1955, 2000.
9. Gel'fand, I. M. and G. E. Shilov, *Generalized Functions. Vol. I: Properties and Operations*, Academic Press, New York, 1964.

Spatiotemporal Properties of Broadband Axicon Fields

R. Dutta, K. Saastamoinen, J. Turunen, and A. T. Friberg

Department of Physics and Mathematics, University of Eastern Finland, Joensuu FI-80101, Finland

Abstract— Propagation-invariant fields generated by achromatic refractive and diffractive axicons using isodiffracting Gaussian beams are considered. In the case of full spectral coherence, the temporal shapes of the resulting approximately propagation-invariant pulses are examined and compared to the incoming field. Additionally, corresponding spectrally incoherent (stationary) fields are investigated and the temporal coherence properties of the generated fields are considered.

1. INTRODUCTION

Generation of propagation-invariant fields [1] can be achieved in various manners, such as optical resonators [2], computer holography [3] or spatial light modulators (SLM) [4]. However, it often relies on the usage of axicons [5, 6] which in the most traditional form are refractive elements with rotationally symmetric conical surfaces. Modern fabrication techniques have allowed diffractive elements, or diffractive axicons [7], to be used instead. Specifically, axicons transform incident Gaussian beams into Bessel-like beams [8]. Specifically, (achromatic) refractive axicons typically produce propagation-invariant Bessel-X waves or pulses [9] and diffractive axicons generate pulsed Bessel beams [10]. They represent non-diffracting or propagation-invariant wave fields as a mathematical solution of the free-space scalar wave equation [11], though in reality the generated fields are only approximately propagation-invariant [12].

While the initial research was concentrated on the generation of propagation-invariant beams using (quasi)monochromatic sources, production of broad-band Bessel beams [8, 13] has attracted much attention recently due to the availability of wideband sources such as super-luminescent diodes or supercontinuum sources. Inclusion of partial optical coherence to propagation-invariance [14] has led to further investigation in the topic of diffraction-free wideband light, such as spectrally partially coherent polychromatic fields [15] and the acknowledgement that ideal polychromatic propagation-invariant fields are spatially partially coherent in space–time domain [16].

In this work we analyze the properties of approximately propagation-invariant fields generated by refractive and diffractive axicons in space-time domain when the axicons are illuminated with collimated isodiffracting Gaussian beams with different spectral properties. In particular, we concentrate on examining the temporal properties of spectrally fully coherent pulses and coherence properties of spectrally incoherent radiation.

2. THEORY

Let us consider an isodiffracting beam with a rotationally symmetric spectral field distribution in the form

$$U(\rho'; \omega) = \left(\frac{2}{\pi w^2} \right)^{1/2} \left(\frac{\omega}{\bar{\omega}} \right)^{1/2} \sqrt{S(\omega)} \exp \left(-\frac{\omega}{\bar{\omega}} \frac{\rho'^2}{w^2} \right), \quad (1)$$

where

$$S(\omega) = \frac{S_0}{\Gamma(2n)\bar{\omega}} \left(2n \frac{\omega}{\bar{\omega}} \right)^{2n} \exp \left(-2n \frac{\omega}{\bar{\omega}} \right) \quad (2)$$

is the spectral density of the field. Here, S_0 is a constant, ρ' is the transverse radial coordinate, n is a real number, $\bar{\omega}$ is the peak angular frequency of $S(\omega)$, and w is the $1/e$ width of the field at $\omega = \bar{\omega}$. Additionally, Γ denotes the Gamma function. Next, we consider a rotationally symmetric achromatic refractive axicon (at the plane $z = 0$) which can be described by a complex-amplitude transmission function

$$t(\rho'; \omega) = \sqrt{\eta(\omega)} \exp \left[-i(\omega/c) \sin \theta(\omega) \rho' \right], \quad (3)$$

where c is the speed of light in vacuum and θ is the cone angle of the axicon. For achromatic axicons diffraction efficiency $\eta(\omega)$ is unity and $\sin \theta(\omega) = \sin \theta$, while for ideal diffractive axicons $\eta(\omega)$ is a constant and $\sin \theta = 2\pi c/d\omega$, where d is the radial period of the circular grating.

Now, we utilize the Fresnel diffraction formula along with Eqs. (1) and (3), and further apply the method of stationary phase [17] for single integrals to find the space-frequency domain field behind the axicon. At a distance z , the propagation-invariant field takes on the form

$$U(\rho, z; \omega) = C\omega\sqrt{S(\omega)} \exp\left(-\frac{\omega}{\bar{\omega}} \frac{z^2}{w^2} \sin^2 \theta\right) J_0\left(\frac{\omega}{c} \sin \theta \rho\right) \exp\left[i\omega\left(\frac{z}{v} + \frac{\rho^2}{2cz}\right)\right], \quad (4)$$

where C is a constant, $v = c/(1 - \sin^2 \theta/2)$ is the phase velocity, and J_0 is the Bessel function of the first kind. For a diffractive axicon, we can see that the Bessel beam is independent of frequency, i.e., diffractive axicons will generate spatially coherent Bessel beams irrespective of the spectral coherence properties of the incident field.

We proceed to examine the time-domain properties of the spectrally fully coherent fields before and after the axicons. For this purpose, we utilize the Fourier-transform for Eqs. (1) and (4) to obtain the temporal fields and further the intensity distributions. For the field before the axicon, the intensity can be expressed as

$$I(\rho'; t) = I_0 \frac{n^3}{\left[(n + \rho'^2/w^2)^2 + (\bar{\omega}t)^2\right]^{n+3/2}}, \quad (5)$$

where $I_0 = I(0; 0)/n^3$. The intensity profile has half-width at half-maximum (HWHM) at $t = T$ as

$$\bar{\omega}T = \sqrt{2^{1/(n+3/2)} - 1} \left(n + \rho'^2/w^2\right). \quad (6)$$

The temporal field after the achromatic axicon is found to be

$$U(\rho, z; t) = -i \left[\frac{(2n)^{2n} S_0 \sin^2 \theta z \bar{\omega}}{\pi \Gamma(2n) z_R c}\right]^{1/2} \int_0^\infty x^{n+1} J_0\left(\frac{\bar{\omega}}{c} \rho \sin \theta x\right) \exp(-ax) dx, \quad (7)$$

where $z_R = \bar{\omega}w^2/2c$, $x = \omega/\bar{\omega}$, and

$$a = n + \left(\frac{z}{L}\right)^2 - i \left(\frac{\bar{\omega}\rho^2}{2zc} - \bar{\omega}t_r\right), \quad (8)$$

with $t_r = t - z/v$ as the retarded time and $L = w/\sin \theta$ as the effective propagation-invariant range of the axicon field. An analytical solution of this integral is not possible for an arbitrary value of ρ , but it can be solved numerically. However, here we restrain ourselves to on-axis ($\rho = 0$) intensity distribution of the field after the achromatic axicon. This yields an analytical expression

$$I(0, z; t_r) = I_{0r} \frac{[n + (z/L)^2]^{2(n+2)}}{\{[n + (z/L)^2]^2 + (\bar{\omega}t_r)^2\}^{n+2}}, \quad (9)$$

where $I_{0r} = I(0, z; 0)$. This has the HWHM value at $t_r = T_r$ as

$$\bar{\omega}T_r = \sqrt{2^{1/(n+2)} - 1} [n + (z/L)^2]. \quad (10)$$

It can readily be seen from Eq. (10) that the pulse width after the axicon increases with axial distance. This effect originates from the fact that the spectral amplitude gets narrower and the axial spectral phase is linear as z increases [see Eq. (4)]. For small values of n (wide spectra) the change of axial pulse width is significant, and further it does not depend on cone angle.

As opposed to the refractive axicon, the diffractive axicon allows an analytical solution for the temporal field as

$$U(\rho, z; t_r) = -\frac{4\pi i}{d} (2n)^n \left[\frac{S_0 z}{\pi \bar{\omega} \Gamma(2n) z_R}\right]^{1/2} J_0\left(\frac{2\pi \rho}{d}\right) \left(\frac{b}{a_n}\right)^{(n+1)/2} K_{-1-n}(2\sqrt{a_n b}), \quad (11)$$

where K denotes the modified Bessel function of the second kind,

$$a_n(\rho, z, t_r) = n + i \left[\bar{\omega}t_r - \frac{\pi}{w/d} \frac{(\rho/d)^2}{z/L}\right], \quad (12)$$

and

$$b(z) = \left(\frac{z}{L}\right)^2 + i\pi\frac{z}{L}\frac{w}{d}. \quad (13)$$

From Eq. (11) we may obtain the intensity profile as $I(\rho, z; t_r) = |U(\rho, z; t_r)|^2$, however it does not allow an analytical expression for HWHM even if $\rho = 0$.

Finally, we proceed to consider spectrally incoherent (stationary) fields with the same spectrum defined in Eq. (2) and focus on the temporal coherence properties before and after the two axicons. Utilizing the Wiener-Khintchine theorem, we may evaluate the mutual coherence function from the spectral density. For the incident field, this leads to the complex degree of temporal coherence which can be expressed as

$$\gamma(\rho', \tau) = \left(\frac{n + \rho'^2/w^2}{n + \rho'^2/w^2 + i\bar{\omega}\tau/2} \right)^{2(n+1)}, \quad (14)$$

where τ is the time difference. The HWHM then is

$$\bar{\omega}\Theta_i = 2\sqrt{2^{1/(n+1)} - 1} \left(n + \frac{\rho'^2}{w^2} \right). \quad (15)$$

On axis, we can express the complex degree of coherence of the field after the achromatic axicon as

$$\gamma(0, z; \tau) = \left[\frac{n + (z/L)^2}{n + (z/L)^2 + i\bar{\omega}\tau/2} \right]^{2n+3}, \quad (16)$$

with HWHM as

$$\bar{\omega}\Theta_r = 2\sqrt{2^{1/(2n+3)} - 1} [n + (z/L)^2]. \quad (17)$$

Similarly, in the case of diffractive axicon the expression of the complex degree of coherence can be written as

$$\gamma(\rho, z; \tau) = \left(\frac{2n}{2n + i\bar{\omega}\tau} \right)^{n+1/2} \frac{K_{-1-2n}(2\sqrt{a_nb})}{K_{-1-2n}(2\sqrt{2nb})}, \quad (18)$$

where $a_n(\tau) = 2n + i\bar{\omega}\tau$ and

$$b(z) = \frac{8\pi^2 c^2 z^2}{d^2 w^2 \bar{\omega}^2}. \quad (19)$$

As expected, HWHM cannot be obtained analytically.

3. RESULTS AND OBSERVATIONS

Next we proceed to present numerical evaluations of the temporal behavior of the propagation-invariant fields. Using Eqs. (5) and (9), we have first calculated the on-axis temporal intensity profile of the incident pulse and after the refractive axicon in Fig. 1 with $n = 1$ at positions $z = L/2$ and $z = L$. For simplicity, we have plotted the intensity as a function of dimensionless variables $\bar{\omega}t$ for the incoming field and $\bar{\omega}t_r$ for the generated field. First, we immediately see that the all fields are sharply peaked and symmetrical. A second obvious feature is the pulse broadening along with z , corresponding to the broadening of the initial pulse along radial direction as Eq. (6) suggests. In non-axial case, as Eq. (7) indicates, the Bessel field exhibits modulation of the intensity profile.

The intensity distribution of the field after the diffractive axicon is plotted in a similar fashion in Fig. 2, with dimensionless quantity $w/d = 10$, $n = 1$, $d = 1.5 \mu\text{m}$, and angular frequency $\bar{\omega}$ corresponding to the wavelength 550 nm. In the figure, intensities have been scaled for ease of comparison. Once again, it is seen that the pulse broadens along axial distance z , but a distinctive feature compared to the refractive axicon is the non-symmetry of the intensity profile with a visible tailing effect and time delay. The time delay becomes more apparent as we increase the transverse (off-axis) coordinates while the pulse shape remains more or less the same.

Finally, we consider the temporal coherence of stationary fields according to Eqs. (14), (16), and (18). Figs. 3(a) and 3(b) show the degree of temporal coherence as a function of $\bar{\omega}\tau$ after both axicons and for the incoming field at different axial distances z . It is seen that the temporal coherence profile for the field after the refractive axicon broadens significantly along with z , while the opposite is the case for diffractive axicon.

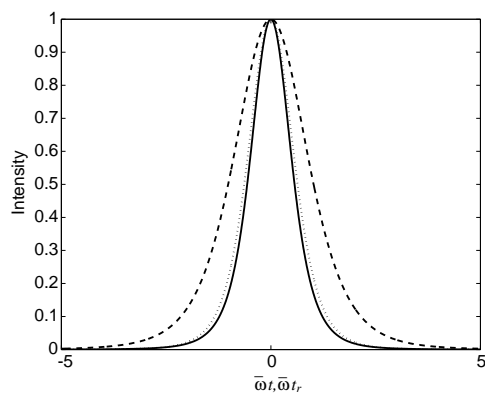


Figure 1: Intensity of the field on axis ($\rho = 0$) before the refractive axicon (solid line) and after it at distances $Z = L/2$ (dotted line), and $z = L$ (dashed line).

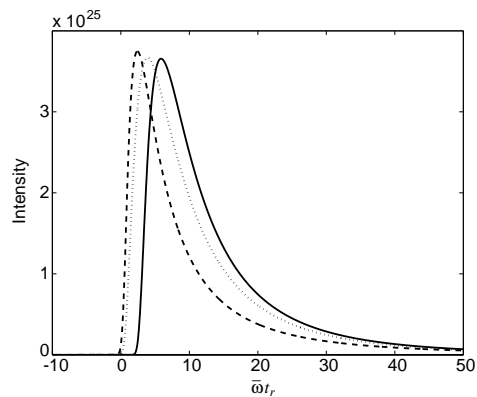


Figure 2: Intensity profiles for the field after the diffractive axicon on the optical axis at $z = L/2$ (dashed line), $z = L$ (dotted line), and off axis at point $\rho/d = 2.5$ (solid line).

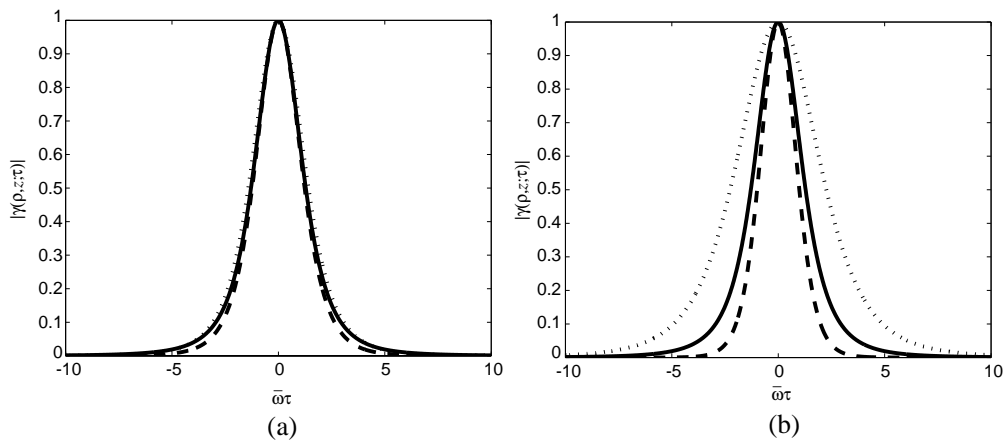


Figure 3: Temporal coherence profiles for the incident field and the fields after the axicons at (a) $z = L/2$ and at (b) $z = L$. In both figures solid line: incident field, dashed line: diffractive axicon, dotted line: refractive axicon.

4. CONCLUSIONS

We have examined approximate propagation-invariant fields generated from an isodiffracting field with achromatic refractive and diffractive axicon. Different spectral coherence properties were considered as a starting point, restricted to examining either fully coherent pulsed fields and incoherent (stationary) fields. The temporal intensities of the pulsed fields after the axicons were calculated, which showed the broadening of the pulses compared to the incoming field. However, differences arose between the fields generated by different axicons. While the refractive axicon provided a field with symmetric, sharply peaked intensity, the diffractive axicon yielded a broader, non-symmetric and time-delayed profile with a clear tailing effect. Additionally, the temporal coherence profiles of the axicon fields exhibited faster decrease in coherence with axial distance for the diffractive axicon rather than the refractive one.

ACKNOWLEDGMENT

This work was supported by the Academy of Finland through the Photonics Program (Grant No. 135027).

REFERENCES

1. Turunen, J. and A. T. Friberg, "Propagation-invariant optical fields," *Progr. Opt.*, Vol. 54, 1–88, 2009.
2. Turunen, J. and A. T. Friberg, "Spatially partially coherent Fabry-Perot modes," *J. Opt. Soc. Am. A*, Vol. 11, 227–235, 1994.

3. Turunen, J., A. Vasara, and A. T. Friberg, "Holographic generation of diffraction-free beams," *Appl. Opt.*, Vol. 27, 3959–3962, 1988.
4. Hernández-Hernández, R. J., R. A. Terborg, I. Ricardez-Vargas, and K. Volke-Sepúlveda, "Experimental generation of Mathieu-Gauss beams with a phase-only spatial light modulator," *Appl. Opt.*, Vol. 49, 6903–6909, 2010.
5. McLeod, J. H., "The axicon: A new type of optical element," *J. Opt. Soc. Am.*, Vol. 44, 592–597, 1954.
6. Jaroszewicz, Z., A. Burvall, and A. T. Friberg, "Axicon—the most important optical element," *Opt. Phot. News*, Vol. 16, 34–39, 2005.
7. Perez, M. V., C. Gomez-Reino, and J. M. Cuadrado, "Diffraction patterns and zone plates produced by thin layer axicons," *Opt. Acta*, Vol. 33, 1161–1176, 1986.
8. Cheng, Z., F. Wu, D. Fan, and X. Fang, "A precise method for analyzing Bessel-like beams generated by broadband waves," *Opt. Laser Technol.*, Vol. 52, 87–90, 2013.
9. Saari, P. and K. Reivelt, "Evidence of X-shaped propagation-invariant localized light waves," *Phys. Rev. Lett.*, Vol. 79, 4135–4138, 1997.
10. Porras, M. A., "Diffraction effects in few-cycle optical pulses," *Phys. Rev. E*, Vol. 65, 026606, 2002.
11. Durnin, J., "Exact solutions for nondiffracting beams. I. The scalar theory," *J. Opt. Soc. Am. A*, Vol. 4, 651–654, 1987.
12. Durnin, J., J. J. Miceli, and J. H. Eberly, "Diffraction free beams," *Phys. Rev. Lett.*, Vol. 58, 1499–1501, 1987.
13. Fischer, P., C. T. A. Brown, J. E. Morris, C. Lopez-Mariscal, E. M. Wright, W. Sibbett, and K. Dholakia, "White light propagation invariant beams," *Opt. Express*, Vol. 13, 6657–6666, 2005.
14. Turunen, J., A. Vasara, and A. T. Friberg, "Propagation-invariance and self-imaging in variable-coherence optics," *J. Opt. Soc. Am. A*, Vol. 8, 282–289, 1991.
15. Saastamoinen, K., J. Turunen, P. Vahimaa, and A. T. Friberg, "Spectrally partially coherent propagation-invariant fields," *Phys. Rev. A*, Vol. 80, 053804, 2009.
16. Turunen, J., "Space-time coherence of polychromatic propagation-invariant fields," *Opt. Express*, Vol. 16, 20283–20294, 2008.
17. Friberg, A. T., "Stationary-phase analysis of generalized axicons," *J. Opt. Soc. Am. A*, Vol. 13, 743–750, 1996.

Semi-automatic Polarimetric SAR Image Classification by MD PSO Based Dynamic Clustering

Turker Ince¹, Serkan Kiranyaz², and Moncef Gabbouj²

¹Department of Electrical and Electronics Engineering, Izmir University of Economics, Izmir, Turkey

²Department of Signal Processing, Tampere University of Technology, Finland

Abstract— In this study, a new systematic approach for semi-automatic classification of polarimetric synthetic aperture radar (PolSAR) image is proposed. The feature extraction block utilizes traditionally used SAR features including the complete coherency (or covariance) matrix information, features derived from various target decomposition theorems, the backscattering power and the selected texture features from gray-level cooccurrence matrix (GLCM). Classification of the information in multi-dimensional PolSAR data space by dynamic clustering is addressed as an optimization problem and recently proposed multi-dimensional particle swarm optimization (MD PSO) technique is applied to find optimal clusters in a given input data space, distance metric and a proper validity index function. An experimental study is performed using the fully polarimetric San Francisco Bay AIRSAR dataset to analyze and compare the results of classification with the state of the art techniques.

1. INTRODUCTION

Image and data classification techniques play an important role in the automatic analysis and interpretation of remote sensing data. Particularly polarimetric synthetic aperture radar (SAR) data poses a challenging problem in this field due to complexity of measured information from its multiple polarimetric channels. Recently, the number of applications which use data provided by the SAR systems having fully polarimetric capability have been increasing. Over the past decade, there has been extensive research in the area of the segmentation and classification of polarimetric SAR data. In the literature, the classification algorithms for polarimetric SAR can be divided into three main classes: 1) classification based on physical scattering mechanisms inherent in data [1, 2]; 2) classification based on statistical characteristics of data [3, 4] and 3) classification based on image processing techniques [5, 6]. Additionally, there has been several works using some combinations of the above classification approaches [1, 3]. Although the supervised approaches to PolSAR classification problem have generally higher accuracy for smaller datasets, their performance depend highly on the availability of limited labeled training samples and their generalization performance is usually lower for larger datasets. On the other hand, the unsupervised clustering based approaches have the advantage of efficient utilization of unlabeled test samples together with labeled training samples for increasing generalization performance.

As one of the pioneer works, Lee et al. [3] proposed an unsupervised classification method based on combination of polarimetric target decomposition (specifically the Cloude and Pottier decomposition [7]) and the maximum likelihood classifier using the complex Wishart distribution. The unsupervised Wishart classifier has an iterative procedure based on the well-known *K-means* algorithm, and has become a preferred benchmark algorithm due to its computational efficiency and generally good performance. However, this classifier still has some significant drawbacks since it entirely relies on *K-means* for actual clustering, such as it may converge to local optima, the number of clusters should be fixed *a priori*, its performance is sensitive to the initialization and its convergence depends on several parameters.

In this study, a new semi-automatic polarimetric SAR (PolSAR) image classification method based on multi-dimensional particle swarm optimization (MD PSO) based dynamic clustering of multi-dimensional information retrieved from fully polarimetric SAR data is proposed. The main advantages of the proposed approach are 1) its ability to perform a global search to find optimum (with respect to given validity index function and distance metric) clusters (classes) in multi-dimensional data space of a polarimetric SAR image, 2) it automatically determines appropriate number of classes for a given input data space, and 3) its robustness against variations of the parameters and initialization. The performance of the proposed classifier is evaluated using the fully polarimetric San Francisco Bay dataset acquired by the NASA/Jet Propulsion Laboratory Airborne SAR (AIRSAR) at L-band. The classification results are compared with state of the art unsupervised $H/A/\alpha$ Wishart classifier.

2. POLARIMETRIC SAR DATA PROCESSING

Polarimetric SAR (PolSAR) features can generally be divided into three categories: 1) the features extracted directly from the polarimetric SAR data and its different transforms such as the scattering matrix, and from which the Stokes matrix, the covariance matrix, and the coherency matrix can be derived; 2) the features based on the polarimetric target decomposition theorems, which are used for information extraction in PolSAR. The coherent decomposition theorems such as the Pauli decomposition [7], the Krogager decomposition [8], the Cameron decomposition [9], and SDH (Sphere, Diplane, Helix) decomposition [10] aim to express the measured scattering matrix by the radar as the combination of scattering responses of coherent scatterers. In the other category, the incoherent decompositions such as the Freeman decomposition [11], the Huynen decomposition [12], and the Cloude-Pottier (eigenvector-eigenvalue or $H/\alpha/A$) decomposition [7] employ the second order polarimetric representations of PolSAR data (such as covariance matrix or coherency matrix) to characterize distributed scatterers; 3) additionally, image processing features such as texture and color are commonly used to improve the classification accuracy. Each feature has its own strength and weaknesses for discriminating different SAR class types and employing multiple features and different combinations can significantly improve the SAR image classification [13]. The following briefly describes the set of features employed in the proposed classification framework.

PolSAR systems often measure the complex scattering matrix, $[S]$, produced by a target under study with the objective to infer its physical properties. Assuming linear horizontal and vertical polarizations for transmitting and receiving, $[S]$ can be expressed as,

$$[S] = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix}. \quad (1)$$

Reciprocity theorem applies in a monostatic system configuration, $S_{hv} = S_{vh}$. For coherent scatterers only, the decompositions of the measured scattering matrix $[S]$ can be employed to characterize the scattering mechanisms of such targets. One way to analyze coherent targets is the Pauli decomposition [3], which expresses $[S]$ in the so-called Pauli basis as,

$$S = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} = \alpha [S]_a + \beta [S]_b + \gamma [S]_c \quad \text{where} \quad (2)$$

$$\left\{ [S]_a = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, [S]_b = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, [S]_c = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$$

and where $\alpha = (S_{hh} + S_{vv})/\sqrt{2}$, $\beta = (S_{hh} - S_{vv})/\sqrt{2}$, $\gamma = \sqrt{2}S_{hv}$.

Hence by means of the Pauli decomposition, all polarimetric information in $[S]$ could be represented by combining the intensities $|\alpha|^2$, $|\beta|^2$ and $|\gamma|^2$, which determine the power scattered by different types of scatterers such as single- or odd-bounce scattering, double- or even-bounce scattering, and orthogonal polarization by volume scattering.

Alternatively, the second order polarimetric descriptors of the average polarimetric covariance $\langle [C] \rangle$ and coherency $\langle [T] \rangle$ matrices can be derived from $[S]$ and employed to extract physical information from the observed scattering process. Due to presence of speckle noise and random vector scattering from surface or volume, PolSAR data are often multi-look processed by averaging n neighboring pixels. By using the Pauli based scattering matrix for a pixel i , $k_i = [S_{hh} + S_{vv}, S_{hh} - S_{vv}, 2S_{hv}]^T / \sqrt{2}$, the multi-look coherency matrix $\langle [T] \rangle$ can be written as

$$\langle [T] \rangle = \frac{1}{n} \sum_{i=1}^n k_i k_i^{*T} \quad (3)$$

Both coherency $\langle [T] \rangle$ and covariance $\langle [C] \rangle$ are 3×3 Hermitian positive semi definite matrices, and since they can be converted into one another by a linear transform, both are equivalent representations of the target polarimetric information.

The Cloude-Pottier decomposition [7] is based on eigenanalysis of the polarimetric coherency matrix, $\langle [T] \rangle$:

$$\langle [T] \rangle = \lambda_1 e_1 e_1^{*T} + \lambda_2 e_2 e_2^{*T} + \lambda_3 e_3 e_3^{*T} \quad (4)$$

where $\lambda_1 > \lambda_2 > \lambda_3 \geq 0$ are real eigenvalues and the corresponding orthonormal eigenvectors e_i (representing three scattering mechanisms) are

$$e_i = e^{i\phi_i} \left[\cos \alpha_i, \sin \alpha_i \cos \beta_i e^{i\delta_i}, \sin \alpha_i \sin \beta_i e^{i\gamma_i} \right]^T \quad (5)$$

Cloude and Pottier defined entropy H , average of set of four angles $\bar{\alpha}$, $\bar{\beta}$, $\bar{\delta}$, and $\bar{\gamma}$, and anisotropy A for analysis of the physical information related to the scattering characteristics of a medium:

$$H = - \sum_{i=1}^3 p_i \log_3 p_i \quad \text{where} \quad p_i = \frac{\lambda_i}{\sum_{i=1}^3 \lambda_i} \quad (6)$$

$$\bar{\alpha} = \sum_{i=1}^3 p_i \alpha_i, \quad \bar{\beta} = \sum_{i=1}^3 p_i \beta_i, \quad (7)$$

$$\bar{\delta} = \sum_{i=1}^3 p_i \delta_i, \quad \bar{\gamma} = \sum_{i=1}^3 p_i \gamma_i \quad A = \frac{p_2 - p_3}{p_2 + p_3}. \quad (8)$$

Due to basis invariance of the target decomposition, the above parameters are roll invariant hence they do not depend on orientation of target about the radar line of sight. Also, information about target's total backscattered power can be determined by the *Span* as,

$$Span = \sum_{i=1}^3 \lambda_i. \quad (9)$$

Additionally, three texture measures, *contrast*, *entropy* and *dissimilarity*, are extracted from normalized gray level co-occurrence matrices (GLCMs) which are calculated using interpixel distance of 2 and averaging over four possible orientation settings ($\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$).

3. DYNAMIC CLUSTERING BY MD PSO

Data clustering is the process of identifying natural groupings in a multidimensional data based on some distance metric (e.g., Euclidean). *K-means* [14] is a well known and widely used clustering method, which first assigns each data point to one of the K cluster centroids and then updates them to the mean of their associated points. As a hard clustering method, *K-means* suffers from the following drawbacks:

- The number of clusters K , needs to be set in advance.
- The performance of the method depends on the initial (random) centroid positions as the method converges to the closest local optima.
- The method is also dependent on the data distribution.

A hard clustering technique based on the PSO has been shown to outperform *K-means*, FCM, KHM and some other state-of-the-art clustering methods in any (evaluation) criteria [15]. PSO-based clustering outperforms many well-known clustering methods, it still suffers from two major drawbacks: the number of clusters, K , (being the solution space dimension as well) should still be specified in advance and similar to other PSO applications, the method tends to trap in local optima particularly when the complexity of the clustering scheme increases. This also involves the dimension of the solution space, i.e., convergence to optimum number of *true* clusters can only be guaranteed for low dimensions.

The multi-dimensional PSO (MD PSO) technique, along with the fractional global best scheme (FGBF), was earlier proposed to determine the solution space dimension in a multidimensional search space and avoid the premature convergence problem [16]. The clustering validity indices (CVI) are used as the fitness functions for clustering and the same validity index used by Omran et al. in their work [15], which is formed as a regularization function fusing both compactness and separation criteria, is employed. Due to space limitations, further details and pseudo codes for MD PSO and FGBF and their mutual application over multidimensional search spaces can be obtained from [16].

4. PROPOSED POLSAR IMAGE CLASSIFICATION FRAMEWORK

In this study, we propose a semi-supervised classification approach based on dynamic clustering of multi-dimensional information retrieved from fully polarimetric SAR data. Data clustering is addressed as an optimization problem and two techniques, MD PSO and FGBF, are employed which extend particle swarm optimization (PSO) in a proper way to find optimal (number of) clusters in a multi-dimensional space.

From the acquired polarimetric SAR data at input of the proposed classification system, the feature extraction unit forms the multi-dimensional feature vector as explained in Section 2. Next as a preprocessing step, normalization and *K-means* clustering is applied to effectively limit the dimension of the solution space to a reasonable number (e.g., $N \leq 512$). The aim of the preprocessing is to increase the efficiency and reduce the computational complexity of the proposed dynamic clustering process without significantly impacting the clustering performance.

As discussed in Section 3, the clustering problem requires the determination of the solution space dimension (i.e., number of clusters, K) and an effective mechanism to avoid local optima traps, (i.e., both dimensionally and spatially) particularly in complex clustering schemes in high dimensions (e.g., $K > 10$). These requirements justify the use of the dynamic clustering technique based on MD PSO with FGBF.

Finally, the class labeling of each cluster prototype (cluster centroid) can be accomplished either manually by the expert or automatically by using the available training data for each class. In the latter case, the class of all points within a cluster is determined by the majority of labeled training points in the cluster. Then, the overall classification of entire PolSAR image can be accomplished by back propagating the labels of cluster prototypes in such a way that points closest to a particular prototype (using the same distance metric in the feature space) are assigned its label. The overview of the proposed system is shown in Fig. 1.

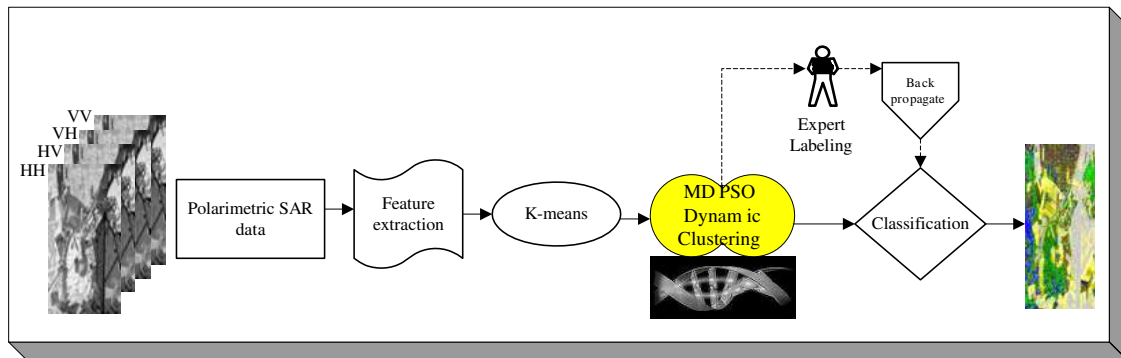


Figure 1: Proposed MD PSO clustering based PolSAR image classification system.

5. EXPERIMENTAL RESULTS

The NASA/Jet Propulsion Laboratory Airborne SAR (AIRSAR) L-band data of the San Francisco Bay is used for performance evaluation of the proposed classifier. The original four-look fully polarimetric SAR data of the San Francisco Bay, having a dimension of 900×1024 pixels, provides good coverage of both natural (sea, mountains, forests, etc.) and man-made targets (buildings, streets, parks, golf course, etc.) with a more complex inner structure. In this study, the speckle filter suggested by Lee et al. [17] is employed with 5×5 window to preserve the texture information as recommended. For clustering algorithm, the typical internal PSO parameters (c_1 , c_2 and w) are used as in [16], and the two critical PSO parameters, swarm size (S) and number of iterations, ($IterNo$) are set as 200 and 2000, respectively.

The convergence plots for one run of MD PSO based clustering fitness score (CVI) and true solution space dimension (number of clusters) are plotted in Fig. 2. For this run MD PSO achieved fitness score of 83.2. To validate MD PSO clustering performance, we also run exhaustive number of *K-means* clustering runs (i.e., 100 times) where K is set to the optimal cluster number found out by that MD PSO (clustering) operation to make a fair comparison. For a San Francisco Bay dataset, the average and the best (minimum) CVI scores achieved by *K-means* were 103.7 and 95.4. Hence, MD PSO clustering CVI was about 13% lower than *K-means*.

The classification results of the proposed PolSAR image classification system and state of the art unsupervised Wishart $H/A/\alpha$ classifier [3] (obtained using PolSARPro 2.0 software [18]) for complete San Francisco Bay image are shown in Fig. 3. Note that the competing algorithm requires the number of data classes (clusters) to be set *a priori* from the user, which was set to 16 ($N = 16$). From the figure, even though both algorithms were able to differentiate reasonably well the large uniform regions corresponding to main classes of scattering such as the ocean, vegetation, and building areas, the proposed system achieves more homogeneous classification results.

Additionally, to test pixel classification accuracy, we selected manually (mostly from homogeneous parts of the image) areas for three main classes, the sea (14268 pixels), urban areas (9455 pixels), and the vegetated zones (4593 pixels) as labeled data. After processing the complete (labeled and unlabeled) data by the proposed classifier, with the number of classes fixed to 3, overall 95% classification accuracy was achieved. The confusion matrix of the proposed method for the labeled data is given in Table 1. From the results, the proposed classifier had most difficulty in separation of vegetated zones from urban areas.

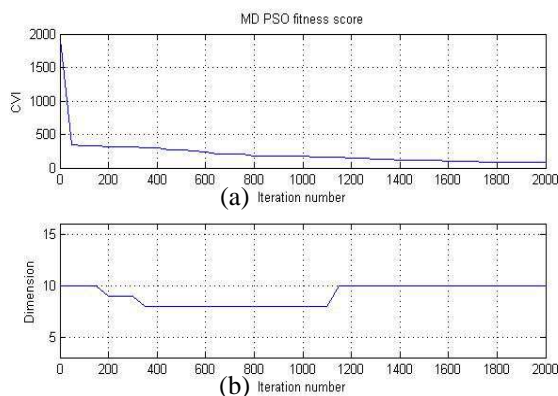


Figure 2: (a) Fitness score and (b) dimension plots vs. iteration number for an MD PSO based clustering run.

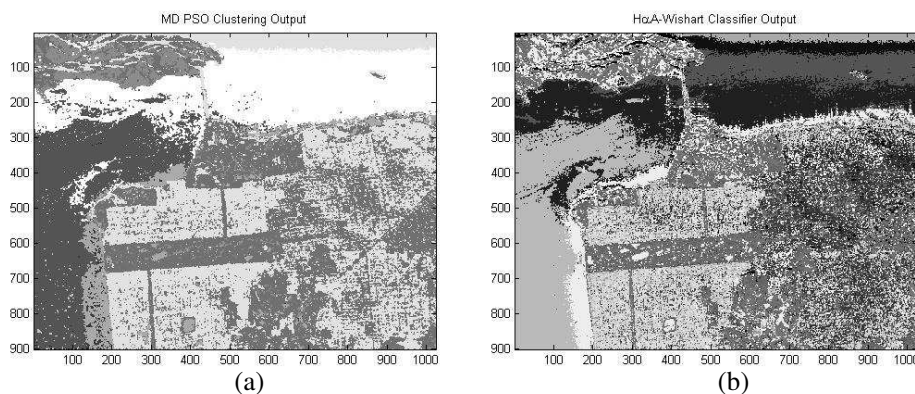


Figure 3: The classification results for San Francisco Bay dataset of the proposed system ((a) $N = 10$) and state of the art $H/\alpha/A$ Wishart classifier ((b) $N = 16$).

Table 1: Confusion matrix for the labeled data.

Ground Truth	Classification Result		
	Sea	Urb	Veg
Sea	14256	2	10
Urb	0	8177	1278
Veg	0	118	4475

6. CONCLUSIONS

In this paper, a semi-automatic PolSAR image classifier system based on MD PSO dynamic clustering has been proposed. The promising results were obtained for the San Francisco Bay dataset and compared to state of the art Wishart $H/A/\alpha$ classifier. Future work will focus on application of the proposed dynamic clustering approach to semi-supervised classifiers (such as RBF or SVM). Also, more experiments using larger real data sets will be done for further validation.

REFERENCES

1. Pottier, E. and J. S. Lee, "Unsupervised classification scheme of POLSAR images based on the complex Wishart distribution and the $H/A/\alpha$ -polarimetric decomposition theorem," *Proc. of the 3rd EUSAR 2000 Conf.*, May 2000.
2. Van Zyl, J. J., "Unsupervised classification of scattering mechanisms using radar polarimetry data," *IEEE Trans. Geosci. Remote Sensing*, Vol. 27, 36–45, Jan. 1989.
3. Lee, J. S., M. R. Grunes, T. Ainsworth, L.-J. Du, D. Schuler, and S. R. Cloude, "Unsupervised classification using polarimetric decomposition and the complex Wishart classifier," *IEEE Trans. Geosci. Remote Sensing*, Vol. 37, No. 5, 2249–2257, Sep. 1999.
4. Wu, Y., K. Ji, W. Yu, and Y. Su, "Region-based classification of polarimetric SAR images using Wishart MRF," *IEEE Geosci. Rem. Sens. Lett.*, Vol. 5, No. 4, 668–672, 2008.
5. Tan, C. P., K. S. Lim, and H. T. Ewe, "Image processing in polarimetric SAR images using a hybrid entropy decomposition and maximum likelihood (EDML)," *Proc. Int. Symposium on Image and Signal Processing and Analysis (ISPA)*, 418–422, Sep. 2007.
6. Ye, Z. and C.-C. Lu, "Wavelet-based unsupervised SAR image segmentation using hidden markov tree models," *Proc. of the 16th International Conference on Pattern Recognition (ICPR'02)*, Vol. 2, 20729, 2002.
7. Cloude, S. R. and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sensing*, Vol. 35, 68–78, Jan. 1997.
8. Krogager, E., "A new decomposition of the radar target scattering matrix," *Electron. Letters*, Vol. 26, No. 18, 1525–1526, 1990.
9. Cameron, W. L., N. N. Youssef, and L. K. Leung, "Feature motivated polarization scattering matrix decomposition," *Proceedings of the IEEE International Radar Conference*, 549–557, Arlington, VA, USA, May 1990.
10. Krogager, E. and Z. H. Czyz, "Properties of the sphere, diplane, helix (target scattering matrix decomposition)," *Proc. JIPR-3*, J. Saillard, et al., Eds., 106–114, Nantes, France, Mar. 1995.
11. Freeman, A. and S. L. Durden, "A three-component scattering model for polarimetric SAR data," *IEEE Trans. Geosci. Remote Sensing*, Vol. 36, No. 3, 963–973, 1998.
12. Huynen, J. R., "Phenomenological theory of radar targets," Ph.D. Dissertation, Tech. University of Delft, Delft, The Netherlands, 1970.
13. She, X. L., J. Yang, and W. J. Zhang, "The boosting algorithm with application to polarimetric SAR image classification," *Proceedings of the 1st Asian and Pacific Conference on Synthetic Aperture Radar (APSAR'07)*, 779–783, Huangshan, China, Nov. 2007.
14. Hammerly, G. and C. Elkan, "Alternatives to the k-means algorithm that find better clusters," *Proc. of the 11th ACM CIKM*, 600–607, 2002.
15. Omran, M. G., A. Salman, and A. P. Engelbrecht, "Dynamic clustering using particle swarm optimization with application in image segmentation," *Pattern Analysis and Applications*, Vol. 8, 332–344, 2006.
16. Kiranyaz, S., T. Ince, A. Yildirim, and M. Gabbouj, "Fractional particle swarm optimization in multi-dimensional search space," *IEEE Transactions on Systems, Man, and Cybernetics — Part B*, Vol. 40, No. 2, 298–319, 2010.
17. Lee, J. S., M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implications for classification," *IEEE Trans. Geosci. Remote Sensing*, Vol. 37, No. 5, 2363–2373, 1999.
18. The Polarimetric SAR Data Processing and Educational Tool (PolSARPro), [Online] Available: <http://earth.esa.int/polsarpro/datasets.html>.

New Propagation Regime for Nonlinear Guided Waves: Coupled Electromagnetic TE-TM Wave Propagation

D. V. Valovik

Penza State University, 40 Krasnaya Street, Penza 440026, Russia

Abstract— Coupled electromagnetic TE and TM wave propagation in a nonlinear plane layer is considered. Nonlinearity is described by Kerr law. Physical problem is reduced to a nonlinear two-parameter eigenvalue problem for Maxwell's equations. Numerical results are given.

1. INTRODUCTION

In this paper, nonlinear coupled surface electromagnetic TE-TM wave propagation in a layer with Kerr nonlinearity is considered. It is known [1, 2] that in this case new propagation regime exists. In this regime we have sum of two independent waves (TE and TM) that create new (nonlinear) polarization (so called coupled TE-TM wave). TE and TM waves that form this nonlinear TE-TM wave we call pseudopolarizations (pseudo-TE and pseudo-TM, respectively). It is proved in [2] that each pseudopolarization propagates with its own propagation constant (γ_E and γ_M for pseudo-TE and pseudo-TM waves, respectively) and at its own frequency (ω_E and ω_M , respectively). The case when $\omega_E = \omega_M$ was investigated in detail in [1], where short background of nonlinear guided waves is also given. This regime is seemed to be interesting because this model can be applied to study nonlinear interaction of two types of waves (pseudo-TE and pseudo-TM) at different frequencies.

2. STATEMENT OF THE PROBLEM

Let us consider electromagnetic waves propagating through a nonlinear homogeneous isotropic nonmagnetic dielectric layer. The permittivity inside the layer is described by Kerr law. The layer is located between two half-spaces: $x < 0$ and $x > h$ in Cartesian coordinate system $Oxyz$. The half-spaces are filled with isotropic nonmagnetic media without any sources and characterized by constant permittivities $\varepsilon_1 \geq \varepsilon_0$ and $\varepsilon_3 \geq \varepsilon_0$, respectively, where ε_0 is the permittivity of free space. Everywhere below $\mu = \mu_0$ is the permeability of free space. Maxwell's equations have the form [3]

$$\text{rot}\tilde{\mathbf{H}} = \partial_t\tilde{\mathbf{D}}, \quad \text{rot}\tilde{\mathbf{E}} = -\partial_t\tilde{\mathbf{B}}, \quad (1)$$

where $\tilde{\mathbf{D}} = \varepsilon\tilde{\mathbf{E}}$, $\tilde{\mathbf{B}} = \mu\tilde{\mathbf{H}}$; $\partial_t = \partial/\partial t$; $(\tilde{\mathbf{E}}, \tilde{\mathbf{H}})$ represents the total field. From system (1), we obtain

$$\text{rot}\tilde{\mathbf{E}} = -\partial_t(\mu\tilde{\mathbf{H}}), \quad \text{rot}\tilde{\mathbf{H}} = \partial_t(\varepsilon\tilde{\mathbf{E}}). \quad (2)$$

We assume that the fields $\tilde{\mathbf{E}}$, $\tilde{\mathbf{H}}$ exist at two frequencies ω_E and ω_M in space

$$\begin{aligned} \tilde{\mathbf{E}}(x, y, z, t) &= \mathbf{E}_E^+(x, y, z) \cos \omega_E t + \mathbf{E}_E^-(x, y, z) \sin \omega_E t + \mathbf{E}_M^+(x, y, z) \cos \omega_M t + \mathbf{E}_M^-(x, y, z) \sin \omega_M t; \\ \tilde{\mathbf{H}}(x, y, z, t) &= \mathbf{H}_E^+(x, y, z) \cos \omega_E t + \mathbf{H}_E^-(x, y, z) \sin \omega_E t + \mathbf{H}_M^+(x, y, z) \cos \omega_M t + \mathbf{H}_M^-(x, y, z) \sin \omega_M t, \end{aligned}$$

where indices E or M correspond, respectively, to TE or TM waves.

Let us form complex amplitudes \mathbf{E} , \mathbf{H}

$$\mathbf{E} = \mathbf{E}_E + \mathbf{E}_M, \quad \mathbf{H} = \mathbf{H}_E + \mathbf{H}_M, \quad (3)$$

where $\mathbf{E}_E = \mathbf{E}_E^+ + i\mathbf{E}_E^-$, $\mathbf{H}_E = \mathbf{H}_E^+ + i\mathbf{H}_E^-$ and $\mathbf{E}_M = \mathbf{E}_M^+ + i\mathbf{E}_M^-$, $\mathbf{H}_M = \mathbf{H}_M^+ + i\mathbf{H}_M^-$ and $\mathbf{E}_E = (0, E_y, 0)^T$, $\mathbf{H}_E = (H_x, 0, H_z)^T$ and $\mathbf{E}_M = (E_x, 0, E_z)^T$, $\mathbf{H}_M = (0, H_y, 0)^T$; $(\cdot)^T$ is the transposition operation. Denote by

$$\mathbf{E}_{\omega_E \omega_M} := \mathbf{E}_E e^{-i\omega_E t} + \mathbf{E}_M e^{-i\omega_M t}, \quad \mathbf{H}_{\omega_E \omega_M} := \mathbf{H}_E e^{-i\omega_E t} + \mathbf{H}_M e^{-i\omega_M t}. \quad (4)$$

As it is known (see, for example, [4, 5]), Kerr law has the form $\varepsilon = \varepsilon_2 + \alpha|\hat{\mathbf{E}}|^2$, where ε_2 is a constant part of the permittivity, α is the coefficient of nonlinearity, and $\hat{\mathbf{E}}$ is a complex amplitude of a monochromatic wave $\hat{\mathbf{E}}e^{-i\omega t}$. The medium is supposed to be isotropic. In the case

under consideration we obtain $|\mathbf{E}_{\omega_E \omega_M}| = |\mathbf{E}_E e^{-i\omega_E t} + \mathbf{E}_M e^{-i\omega_M t}| = |\mathbf{E}_E + \mathbf{E}_M| = |\mathbf{E}|$, where \mathbf{E} is defined by formula (3). Thus, the Kerr law is fulfilled.

In the case under consideration we obtain that Maxwell's equations depend on t in the same way as if ε was constant. This allows us to write down Equation (2) for field (4)

$$\begin{aligned} \operatorname{rot}(\mathbf{E}_E e^{-i\omega_E t} + \mathbf{E}_M e^{-i\omega_M t}) &= i\mu\omega_E \mathbf{H}_E e^{-i\omega_E t} + i\mu\omega_M \mathbf{H}_M e^{-i\omega_M t}, \\ \operatorname{rot}(\mathbf{H}_E e^{-i\omega_E t} + \mathbf{H}_M e^{-i\omega_M t}) &= -i\varepsilon\omega_E \mathbf{E}_E e^{-i\omega_E t} - i\varepsilon\omega_M \mathbf{E}_M e^{-i\omega_M t}. \end{aligned} \quad (5)$$

Thus, we gave proof of possibility to investigate Equation (5) for the complex amplitudes \mathbf{E} and \mathbf{H} instead of Equation (1) for the fields $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{H}}$.

The complex amplitudes \mathbf{E} , \mathbf{H} (see (3)) must satisfy Equation (5); the continuity condition for tangential field components on the boundaries $x = 0$, $x = h$; the radiation condition at infinity, where electromagnetic field exponentially decays as $|x| \rightarrow \infty$ in the domains $x < 0$ and $x > h$.

The solutions to Equation (5) are sought for in the entire space.

It is well known (see, for example, [3]) that in the case of constant permittivity surface wave can be represented as a superposition of TE and TM waves. This means that in this case general solution to Maxwell's Equation (2) is a linear combination of TE and TM waves. At this rate if we substitute TE or TM fields in Equation (2) we shall be convinced that the components H_z , H_x (for TE wave) and components E_z , E_x (for TM wave) do not depend on y . Further, it turns out that (see, for example, [5]) for Kerr nonlinearity and for each polarization separately (TE and TM) exists surface waves that do not depend on y . We shall assume that in the nonlinear case under consideration the components E_z , E_x , H_z , H_x do not depend on y [5].

Waves propagating along the boundary z depend harmonically on z . In other words, dependence on z for the components of the fields has the form $e^{i\gamma z}$, where γ is unknown spectral parameter (propagation constant). It is clear that $|\hat{\mathbf{E}} e^{i\gamma z}| = |\hat{\mathbf{E}}|$ does not depend on z if $\operatorname{Im} \gamma = 0$. That is to say, system (5) depend linearly on $e^{i\gamma z}$. On the other hand, for each pseudopolarizations (TE and TM) different spectral parameters can be chosen. We can chose γ_E for TE waves and γ_M for TM waves [1] (see Fig. 1). Thus, the components E_x , E_y , E_z , H_x , H_y , H_z have the form

$$\begin{aligned} E_x &\equiv E_x(x) e^{i\gamma_M z}, & E_y &\equiv E_y(x) e^{i\gamma_E z}, & E_z &\equiv E_z(x) e^{i\gamma_M z}, \\ H_x &\equiv H_x(x) e^{i\gamma_E z}, & H_y &\equiv H_y(x) e^{i\gamma_M z}, & H_z &\equiv H_z(x) e^{i\gamma_E z}, \end{aligned}$$

and the fields $\tilde{\mathbf{E}}$, $\tilde{\mathbf{H}}$ have the form

$$\begin{aligned} \tilde{\mathbf{E}} &= \operatorname{Re} \left\{ (E_x(x) e^{i(\gamma_M z - \omega_M t)}, E_y(x) e^{i(\gamma_E z - \omega_E t)}, E_z(x) e^{i(\gamma_M z - \omega_M t)})^T \right\}, \\ \tilde{\mathbf{H}} &= \operatorname{Re} \left\{ (H_x(x) e^{i(\gamma_E z - \omega_E t)}, H_y(x) e^{i(\gamma_M z - \omega_M t)}, H_z(x) e^{i(\gamma_E z - \omega_E t)})^T \right\}. \end{aligned}$$

After simple transformations from (5) (normalizing accordingly with the formulae $\tilde{x} = kx$, $\frac{d}{dx} = k \frac{d}{d\tilde{x}}$, $\tilde{\gamma}_E = \frac{\gamma_E}{k}$, $\tilde{\gamma}_M = \frac{\gamma_M}{k}$, $\tilde{\varepsilon}_j = \frac{\varepsilon_j}{\varepsilon_0}$, ($j = 1, 2, 3$), $\tilde{\alpha} = \frac{\alpha}{\varepsilon_0}$, where $k^2 = \omega_M^2 \varepsilon_0 \mu_0$, introducing notation $X := iE_x$, $Y := E_y$, $Z := E_z$, $\tau := \omega_E^2 \omega_M^{-2}$, and omitting the tilde symbol) we obtain

$$\begin{cases} \gamma_M^2 X - \gamma_M Z' = \varepsilon X, \\ \gamma_E^2 Y - Y'' = \tau \varepsilon Y, \\ \gamma_M X' - Z'' = \varepsilon Z, \end{cases} \quad \text{where } \varepsilon = \begin{cases} \varepsilon_1, & x < 0 \\ \varepsilon_2 + \alpha(X^2 + Y^2 + Z^2), & 0 < x < h \\ \varepsilon_3, & x > h. \end{cases} \quad (6)$$

Introduce the notation $k_{E1}^2 = \gamma_E^2 - \tau \varepsilon_1$, $k_{E3}^2 = \gamma_E^2 - \tau \varepsilon_3$, $k_{M1}^2 = \gamma_M^2 - \varepsilon_1$, $k_{M3}^2 = \gamma_M^2 - \varepsilon_3$.

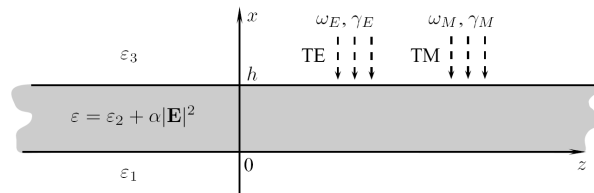


Figure 1: Geometry of the problem.

System (6) is linear in the half-spaces $x < -h$ and $x > h$. Its solutions (in accordance with the radiation condition) have the forms, respectively

$$\left\{ \begin{array}{l} X(x) = C_M^{(0)} e^{xk_{M1}}, \\ Y(x) = C_E^{(0)} e^{xk_{E1}}, \\ Z(x) = \gamma_M^{-1} k_{M1} C_M^{(0)} e^{xk_{M1}} \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} X(x) = C_M^{(h)} e^{-(x-h)k_{M3}}, \\ Y(x) = C_E^{(h)} e^{-(x-h)k_{E3}}, \\ Z(x) = -\gamma_M^{-1} k_{M3} C_M^{(h)} e^{-(x-h)k_{M3}}. \end{array} \right. \quad (7)$$

The constants $C_M^{(0)}$ and $C_E^{(0)}$ ($C_M^{(h)}$ and $C_E^{(h)}$) are supposed to be known (initial conditions).

Tangential components of electromagnetic field are known to be continuous at the interfaces. In this case tangential components are E_y , E_z , H_y , H_z . It follows from the continuity of H_y that $Z' - \gamma_M X$ is continuous at the interfaces. It follows from above that the transmission conditions for the functions X , Y , Y' , Z have the form

$$\begin{aligned} [Z' - \gamma_M X]|_{x=0} &= [Z' - \gamma_M X]|_{x=h} = 0, \\ [Y]|_{x=0} &= [Y]|_{x=h} = 0, \\ [Y']|_{x=0} &= [Y']|_{x=h} = 0, \\ [Z]|_{x=0} &= [Z]|_{x=h} = 0, \end{aligned} \quad (8)$$

where $[f]|_{x=x_0} = \lim_{x \rightarrow x_0-0} f(x) - \lim_{x \rightarrow x_0+0} f(x)$.

Definition 1. The pair (γ_E, γ_M) is called coupled eigenvalues of the problem if for prescribed values $C_E^{(0)}$ and $C_M^{(0)}$ there are nontrivial functions X , Y , Z such that they satisfy system (6) for $0 < x < h$; for $h < 0$, $x > h$ they defined by (7), respectively; and they satisfy transmission conditions (8). The functions X , Y , Z are called eigenfunctions of the problem.

The main problem (problem P) is to determine coupled eigenvalues.

3. DISPERSION EQUATIONS AND NUMERICAL RESULTS

It can be proved [1, 2] that the system of dispersion equations have the form

$$\left\{ \begin{array}{l} C_E^{(h)} g_E(h, \gamma_E) = \alpha \frac{Q_E(h, \gamma_E, \gamma_M)}{\sin 2k_E h}, \\ C_M^{(h)} k_M g_M(h, \gamma_M) = \alpha \frac{Q_M(h, \gamma_E, \gamma_M)}{\sin 2k_M h}, \end{array} \right. \quad (9)$$

where

$$\begin{aligned} g_E(h, \gamma_E) &= (k_E^2 - k_{E1} k_{E3}) \sin 2k_E h - k_E (k_{E1} + k_{E3}) \cos 2k_E h, \\ g_M(h, \gamma_M) &= (\varepsilon_1 \varepsilon_3 k_M^2 - \varepsilon_2^2 k_{M1} k_{M3}) \sin 2k_M h - \varepsilon_2 k_M (\varepsilon_1 k_{M3} + \varepsilon_3 k_{M1}) \cos 2k_M h, \end{aligned}$$

and $Q_E(h, \gamma_E, \gamma_M)$, $Q_M(h, \gamma_E, \gamma_M)$ are given in [2].

Coupled eigenvalues (propagation constants) can be determined from system (9).

If we put $\alpha = 0$ in (9) we obtain the equations $g_E(\gamma_E) = 0$ and $g_E(\gamma_M) = 0$, which are dispersion equations for a constant permittivity in the layer.

The following theorem is the main result of this paper (proof see in [1, 2]).

Theorem 1. Let $\varepsilon_2 > \max(\varepsilon_1, \varepsilon_3) > 0$ and $\tilde{\gamma}_E$, $\tilde{\gamma}_M$ be solutions to the equations $g_E(\gamma_E) = 0$, $g_E(\gamma_M) = 0$. Then it is possible to choose α such that the pair $(\hat{\gamma}_E, \hat{\gamma}_M)$ exists in the vicinity of $(\tilde{\gamma}_E, \tilde{\gamma}_M)$. The pair $(\hat{\gamma}_E, \hat{\gamma}_M)$ is a solution to problem P

Consider the layer of thickness $h = 8$ with $\varepsilon_1 = 1$, $\varepsilon_2 = 4$, $\varepsilon_3 = 1$, and $\alpha = 0.001$, $\omega_E = \omega_M$.

In the case of $\alpha = 0$ for the linear cases we have, for TE waves and TM waves, respectively

$$\begin{aligned} \gamma_E^{\text{lin}} &= 1.132, & \gamma_E^{\text{lin}} &= 1.478, & \gamma_E^{\text{lin}} &= 1.720, & \gamma_E^{\text{lin}} &= 1.879, & \gamma_E^{\text{lin}} &= 1.970; \\ \gamma_M^{\text{lin}} &= 1.036, & \gamma_M^{\text{lin}} &= 1.351, & \gamma_M^{\text{lin}} &= 1.653, & \gamma_M^{\text{lin}} &= 1.852, & \gamma_M^{\text{lin}} &= 1.964. \end{aligned}$$

The gray vertical and horizontal segments in the left lower corner in Figs. 5, 7, 8 correspond to the propagation constants γ_E^{lin} and γ_M^{lin} of the linear problems.

In Figs. 2, 7, 8 the points of intersections of black and blue curves are nonlinear coupled propagation constants.

In Figs. 3–6 below solid curve (red) corresponds to eigenfunction X ; dashdot curve (blue) corresponds to Y ; dash curve (green) corresponds to Z .

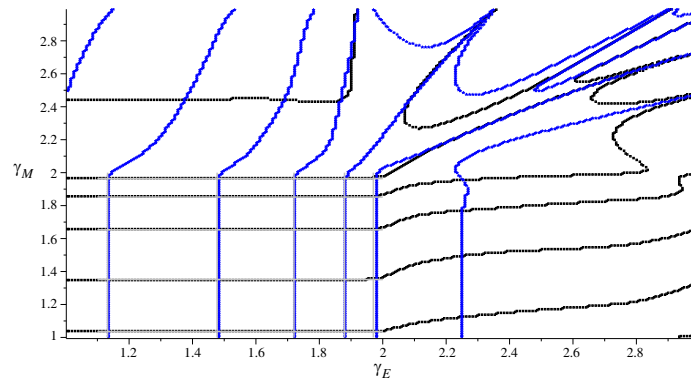


Figure 2: Dispersion curves. Parameters: $C_E^{(0)} = 1$ and $C_M^{(0)} = 1$, $\alpha = 0.001$.

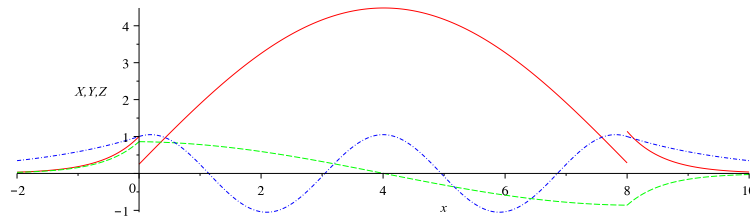


Figure 3: Eigenfunctions. Parameters: $C_E^{(0)} = 1$, $C_M^{(0)} = 1$, $\alpha = 0$, $\gamma_E^{\text{lin}} = 1.132$, $\gamma_M^{\text{lin}} = 1.964$.

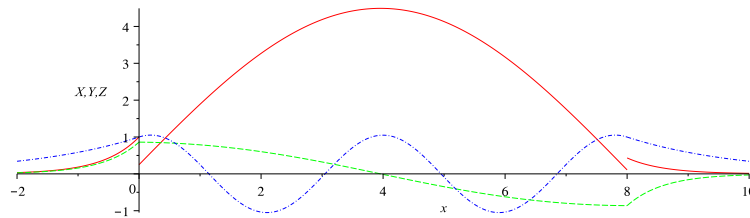


Figure 4: Eigenfunctions. Parameters: $C_E^{(0)} = 1$, $C_M^{(0)} = 1$, $\alpha = 0.001$, $(\gamma_E, \gamma_M) = (1.135, 1.967)$.

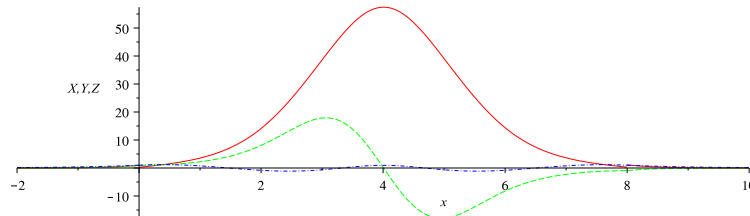


Figure 5: Eigenfunctions. Parameters: $C_E^{(0)} = 1$, $C_M^{(0)} = 1$, $\alpha = 0.001$, $(\gamma_E, \gamma_M) = (1.373, 2.44)$.

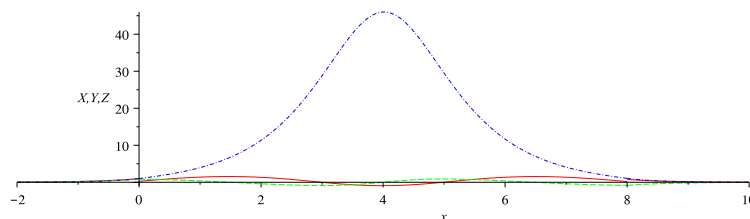


Figure 6: Eigenfunctions. Parameters: $C_E^{(0)} = 1$, $C_M^{(0)} = 1$, $\alpha = 0.001$, $(\gamma_E, \gamma_M) = (2.249, 1.767)$.

In Fig. 3 eigenfunctions of the linear problem are depicted.

In Figs. 4–6 eigenfunctions of the nonlinear problem are depicted.

Coupled eigenvalues chosen for depicted eigenfunctions can be easily found in Fig. 2.

As it is expected the eigenfunction X has a finite jump at the interfaces; the eigenfunction Y is continuous and smooth at the interfaces; the eigenfunction Z is continuous at the interface.

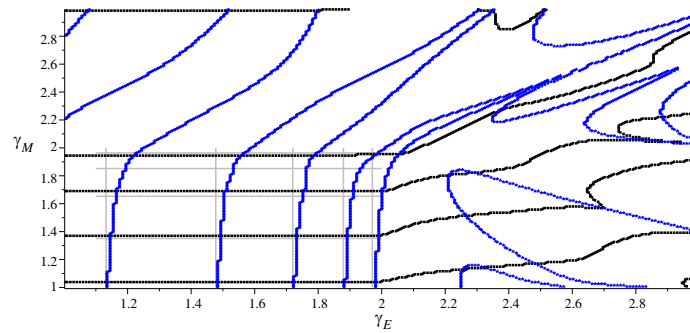


Figure 7: Dispersion curves. Parameters: $C_E^{(0)} = 1$ and $C_M^{(0)} = 10$, $\alpha = 0.001$.

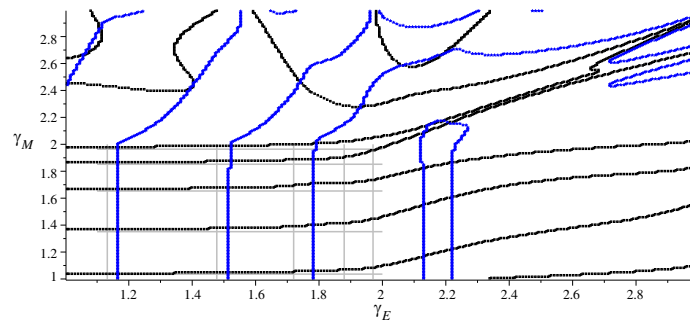


Figure 8: Dispersion curves. Parameters: $C_E^{(0)} = 10$ and $C_M^{(0)} = 1$, $\alpha = 0.001$.

What happens in the nonlinear case if we change $C_E^{(0)}$ or $C_M^{(0)}$? To find out see Figs. 7, 8.

As it is expected propagation constants of the linear problems (vertical and horizontal segments of gray color in Figs. 7, 8) do not depend on $C_E^{(0)}$ and $C_M^{(0)}$.

4. CONCLUSION

Some peculiarities of the nonlinear problem should be marked: (a) new nonlinear waves exist in the regions where linear waves do not exist 4; (b) Nonlinear coupled propagation constants depend on the values $C_E^{(0)}$ and $C_M^{(0)}$ (see system (9) and Figs. 7, 8); (c) In coupled nonlinear wave it is probably possible to control one wave by changing the parameters of the other.

Natural question is to check experimentally existence of coupled surface TE-TM waves.

ACKNOWLEDGMENT

This study is supported by the RFBR (Grants Nos. 11-07-00330-A, 12-07-97010-R A), The Ministry of Education and Science of the Russian Federation (Grant No. 14.B37.21.1950).

REFERENCES

1. Smirnov, Y. G. and D. V. Valovik, "Coupled electromagnetic te-tm wave propagation in a layer with kerr nonlinearity," *Journal of Mathematical Physics*, Vol. 53, No. 12, 123530-1-24, 2012.
2. Valovik, D. V., "On the problem of nonlinear coupled electromagnetic TE-TM wave propagation," *Journal of Mathematical Physics*, Vol. 54, No. 4, 042902-1-14, 2013.
3. Stretton, J. A., *Electromagnetic Theory*, McGraw Hill, New York, 1941.
4. Landau, L. D., E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media, Vol. 8 (Course of Theoretical Physics)*, Butterworth-Heinemann, Oxford, 1993.
5. Boardman, A. D., P. Egan, F. Lederer, U. Langbein, and D. Mihalache, *Third-order Nonlinear Electromagnetic TE and TM Guided Waves*, Elsevier Science Publication, 1991; Reprinted from *Nonlinear Surface Electromagnetic Phenomena*, Eds. H.-E. Ponath and G. I. Stegeman.

Reconstruction of Permittivity and Permeability Tensors of Anisotropic Materials in a Rectangular Waveguide from the Reflection and Transmission Coefficients at Different Frequencies

Yu. G. Smirnov¹, Yu. V. Shestopalov², and E. D. Derevyanchuk¹

¹Penza State University, Russia

²Karlstad University, Sweden

Abstract— This paper is devoted to the study of inverse problem of the permittivity and permeability tensor reconstruction of anisotropic materials in the form of diaphragms (sections) in a single-mode waveguide of rectangular cross section from the transmission and reflection coefficients measured. Results of numerical modeling are presented for different types of anisotropic materials. The developed solution techniques for the inverse problem under study can be applied in optics, nanotechnology, and design of microwave devices.

1. INTRODUCTION

Determination of electromagnetic parameters of anisotropic dielectrics bodies of complicated structure is an urgent problem. However, as a rule, these parameters cannot be directly measured (because of composite character of the material and small size of samples), so that numerical solution to the corresponding forward and inverse electromagnetic problems must be applied [1]. In this paper we consider the inverse problem of the permittivity determination of a multi-sectional diaphragm aimed at reconstructing permittivity and permeability tensors of anisotropic materials in the form of diaphragms (sections) in a single-mode waveguide of rectangular cross section from the transmission and reflection coefficients measured. Numerical results confirm the efficiency of the method.

2. INVERSE PROBLEM

Assume that a waveguide $P = \{x: 0 < x_1 < a, 0 < x_2 < b, -\infty < x_3 < \infty\}$ with the perfectly conducting boundary surface ∂P is given in the cartesian coordinate system. A three-dimensional body Q ($Q \subset P$) is placed in the waveguide; the body has the form of a diaphragm adjacent to the waveguide walls (Fig. 1).

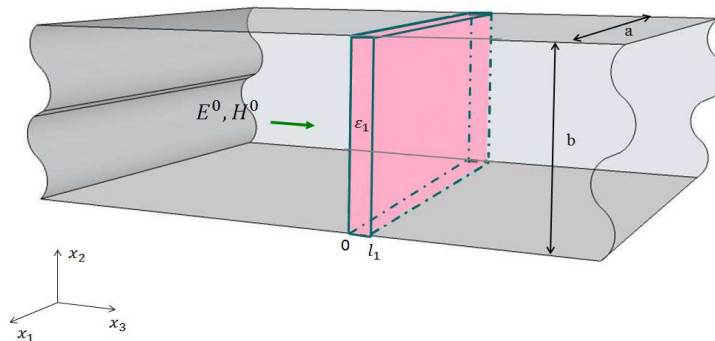


Figure 1. Diaphragm in a waveguide.

Domain $P \setminus \bar{Q}$ is filled with an isotropic and homogeneous layered medium having constant permeability ($\mu_0 > 0$) and constant permittivity ($\varepsilon_0 > 0$), the sections of the diaphragm

$$Q_0 = \{x: 0 < x_1 < a, 0 < x_2 < b, -\infty < x_3 < 0\}$$

$$Q_1 = \{x: 0 < x_1 < a, 0 < x_2 < b, 0 < x_3 < l_1\},$$

$$Q_2 = \{x: 0 < x_1 < a, 0 < x_2 < b, l_2 < x_3 < +\infty\}$$

each with an anisotropic medium having diagonal permittivity and permeability tensors:

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon_{11} & 0 & 0 \\ 0 & \varepsilon_{22} & 0 \\ 0 & 0 & \varepsilon_{33} \end{pmatrix}, \quad \hat{\mu} = \begin{pmatrix} \mu_{11} & 0 & 0 \\ 0 & \mu_{22} & 0 \\ 0 & 0 & \mu_{33} \end{pmatrix}$$

Length l_1 of diaphragm is known. The electromagnetic field inside and outside the object in the waveguide is governed by Maxwell's equations

$$\begin{aligned} \operatorname{rot} \mathbf{H} &= -i\omega \hat{\varepsilon} \mathbf{E} \\ \operatorname{rot} \mathbf{E} &= i\omega \hat{\mu} \mathbf{H}, \end{aligned} \quad (1)$$

where \mathbf{E} and \mathbf{H} are the vectors of the electric and magnetic field intensity and ω is the circular frequency.

Assume that $\pi/a < k_0 < \pi/b$, where k_0 is the wavenumber, $k_0^2 = \omega^2 \varepsilon_0 \mu_0$. In this case, only one wave H_{10} propagates in the waveguide without attenuation (we have a single-mode waveguide [2]).

The incident electrical field is

$$\mathbf{E}^0 = \mathbf{e}_2 A \sin\left(\frac{\pi x_1}{a}\right) e^{-i\gamma_0 x_3} \quad (2)$$

with a known A and $\gamma_0 = \sqrt{k_0^2 - \pi^2/a^2}$. Solving the forward problem for Maxwell's equations with the aid of (2), we obtain explicit expressions for the field inside every section of diaphragm Q and outside the diaphragm:

$$\begin{aligned} E^{(0)} &= \sin\left(\frac{\pi x_1}{a}\right) (Ae^{-i\gamma_0 x_3} + Be^{i\gamma_0 x_3}), \quad x \in Q_0, \\ E^{(1)} &= \sin\left(\frac{\pi x_1}{a}\right) (C_1 e^{-i\gamma_1 x_3} + D_1 e^{i\gamma_1 x_3}), \\ E^{(2)} &= \sin\left(\frac{\pi x_1}{a}\right) F e^{-i\gamma_2 x_3}, \quad \gamma_2 = \gamma_0 \end{aligned} \quad (3)$$

Substituting (3) into Maxwell's equations we obtain an explicit formula for γ_1

$$\gamma_1 = \sqrt{\left(\omega^2 \mu_{11} \mu_{33} \varepsilon_{22} - \frac{\pi^2}{a^2} \mu_{11}\right) / \mu_{33}} \quad (4)$$

From the conditions on the boundary surfaces $L := \{x_3 = 0, x_3 = l_1\}$ of the diaphragm sections

$$[E_\tau]_L = 0, \quad [H_\tau]_L = 0, \quad (5)$$

where E_τ , H_τ are tangential components of E and H and square brackets $[\cdot]$ denote the function jump over the boundary surfaces, applied to (3), (4) we obtain using conditions (5) a system of equations for the unknown coefficients

$$\begin{cases} A + B = C_1 + D_1 \\ \frac{\gamma_0}{\mu_0} (B - A) = \frac{\gamma_1}{\mu_{11}} (D_1 - C_1) \\ C_1 e^{-i\gamma_1 l_1} + D_1 e^{i\gamma_1 l_1} = F e^{-i\gamma_0 l_1} \\ \frac{\gamma_1}{\mu_{11}} (D_1 e^{i\gamma_1 l_1} - C_1 e^{-i\gamma_1 l_1}) = \frac{\gamma_0}{\mu_{11}} (-F e^{-i\gamma_0 l_1}), \end{cases} \quad (6)$$

In system (6) coefficients A , B , C_1 , D_1 , F , $\hat{\varepsilon}$, $\hat{\mu}$ are supposed to be complex.

Inverse problem P1(2): find (complex) permittivity $\hat{\varepsilon}$ (permeability) ($\hat{\mu}$) of section from the known transmission coefficient F .

Inverse problem Q1(2): find (complex) permittivity $\hat{\varepsilon}$ (permeability) ($\hat{\mu}$) of section from the known reflection coefficient B .

2.1. Permittivity Tensor Reconstruction

Consider inverse problem $P1$. Expressing C_1, D_1 via F we obtain a recurrent formula that couples amplitudes A and F :

$$\frac{Ae^{i\gamma_0 l_1}}{F^{(1)}} = \left(\cos(\gamma_1^{(1)} l_1) + i \sin(\gamma_1^{(1)} l_1) \left(\frac{\gamma_0 \mu_{11}}{2\gamma_1^{(1)} \mu_0} + \frac{\gamma_1^{(1)} \mu_0}{2\gamma_0 \mu_{11}} \right) \right), \quad (7)$$

where $\gamma_1^{(1)} := \gamma_1$, γ_1 is expressed by formula (4). As a result we obtain a nonlinear equation with one unknown ε_{22} . Solving this system, we can find only the unknown ε_{22} . However, we cannot find other components of permittivity tensors: ε_{11} and ε_{33} . To do this we have determined the tensor transformation that changes the positions of its diagonal components (in our case, ε_{11} and ε_{22} are interchanged as well as ε_{33} and ε_{22}). First, to interchange components ε_{11} and ε_{22} we can carry out these two transformations sequentially.

$$A_1 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

Then

$$\varepsilon^{\hat{(2)}} = A_2^{-1} A_1^{-1} \hat{\varepsilon} A_1 A_2, \quad \varepsilon^{\hat{(2)}} = \begin{pmatrix} \varepsilon_{33} & 0 & 0 \\ 0 & \varepsilon_{11} & 0 \\ 0 & 0 & \varepsilon_{22} \end{pmatrix}$$

A_1 and A_2 are the diaphragm rotation to the angle $\varphi = \frac{\pi}{2}$ w.r.t. Ox_2 and Ox_1 axes, respectively. Both transformations A_1, A_2 can be realized in practise. The permeability tensor $\hat{\mu}$ is transformed in the same manner. Then for new tensors $\varepsilon^{\hat{(2)}}, \mu^{\hat{(2)}}$, the inverse problem is reduced to numerical solution of the following nonlinear equation:

$$\frac{Ae^{i\gamma_0 l_1}}{F^{(2)}} = \left(\cos(\gamma_1^{(2)} l_1) + i \sin(\gamma_1^{(2)} l_1) \left(\frac{\gamma_0 \mu_{33}}{2\gamma_1^{(2)} \mu_0} + \frac{\gamma_1^{(2)} \mu_0}{2\gamma_0 \mu_{33}} \right) \right), \quad \gamma_1^{(2)} = \sqrt{\frac{(\omega^2 \mu_{33} \mu_{22} \varepsilon_{11} - \frac{\pi^2}{a^2} \mu_{33})}{\mu_{22}}} \quad (8)$$

Solving this equation, we can find only unknown ε_{11} . Next we consider the last transformation interchanging components ε_{33} and ε_{22} .

$$A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

Then

$$\varepsilon^{\hat{(3)}} = A_3^{-1} \hat{\varepsilon} A_3, \quad \varepsilon^{\hat{(3)}} = \begin{pmatrix} \varepsilon_{11} & 0 & 0 \\ 0 & \varepsilon_{33} & 0 \\ 0 & 0 & \varepsilon_{22} \end{pmatrix}$$

A_3 is the diaphragm rotation to the angle $\varphi = \frac{\pi}{2}$ w.r.t Ox_1 . Transformation A_3 can be also realized in practise. Permeability tensor $\hat{\mu}$ is transformed in the same manner and the inverse problem then for new tensors $\varepsilon^{\hat{(3)}}, \mu^{\hat{(3)}}$ is reduced to numerical solution of the nonlinear equation:

$$\frac{Ae^{i\gamma_0 l_1}}{F^{(3)}} = \left(\cos(\gamma_1^{(3)} l_1) + i \sin(\gamma_1^{(3)} l_1) \left(\frac{\gamma_0 \mu_{33}}{2\gamma_1^{(3)} \mu_0} + \frac{\gamma_1^{(3)} \mu_0}{2\gamma_0 \mu_{33}} \right) \right), \quad \gamma_1^{(3)} = \sqrt{\frac{(\omega^2 \mu_{11} \mu_{22} \varepsilon_{33} - \frac{\pi^2}{a^2} \mu_{11})}{\mu_{22}}} \quad (9)$$

Solving this system, we can find only the unknown ε_{33} .

Combining systems (7), (8), (9) we obtain the resulting nonlinear system of equations. Solving this system we can find all permittivity tensor components: $\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}$.

3. PERMEABILITY TENSOR RECONSTRUCTION

Consider inverse problem $P2$. Solving Equation (7) we can find only unknown μ_{11} . To determine other diagonal tensor components we should find such tensor transformations that diagonal tensor

components would change their positions on the main diagonal, (in our case, μ_{11} and μ_{22} are interchanged as well as μ_{11} and μ_{33}) and each transformations could be implemented in practice. Such transformation was found.

$$A_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Then

$$\hat{\mu}^{(2)} = A_2^{-1} \hat{\mu} A_2, \quad \mu^{(2)} = \begin{pmatrix} \mu_{33} & 0 & 0 \\ 0 & \mu_{22} & 0 \\ 0 & 0 & \mu_{11} \end{pmatrix}$$

A_2 is the diaphragm rotation to the angle $\varphi = \frac{\pi}{2}$ w.r.t Ox_2 . Permittivity tensor $\hat{\varepsilon}$ is transformed in the same manner. Then we obtain

$$\frac{Ae^{i\gamma_0 l_1}}{F^{(2)}} = \left(\cos(\gamma_1^{(2)} l_1) + i \sin(\gamma_1^{(2)} l_1) \left(\frac{\gamma_0 \mu_{33}}{2\gamma_1^{(2)} \mu_0} + \frac{\gamma_1^{(2)} \mu_0}{2\gamma_0 \mu_{33}} \right) \right), \quad \gamma_1^{(2)} = \sqrt{\frac{(\omega^2 \mu_{33} \mu_{11} \varepsilon_{22} - \frac{\pi^2}{a^2} \mu_{33})}{\mu_{11}}} \quad (10)$$

Solving this equation, we can find only μ_{33} . Next,

$$A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \hat{\mu}^{(3)} = A_3^{-1} \hat{\mu} A_3, \quad \mu^{(3)} = \begin{pmatrix} \mu_{22} & 0 & 0 \\ 0 & \mu_{11} & 0 \\ 0 & 0 & \mu_{33} \end{pmatrix}$$

A_3 is the diaphragm rotation to the angle $\varphi = \frac{\pi}{2}$ w.r.t Ox_1 and this transformation can be also realized in practise. Permittivity tensor $\hat{\varepsilon}$ is changed in the same manner. Then we obtain

$$\frac{Ae^{i\gamma_0 l_1}}{F^{(3)}} = \left(\cos(\gamma_1^{(3)} l_1) + i \sin(\gamma_1^{(3)} l_1) \left(\frac{\gamma_0 \mu_{11}}{2\gamma_1^{(3)} \mu_0} + \frac{\gamma_1^{(3)} \mu_0}{2\gamma_0 \mu_{11}} \right) \right), \quad \gamma_1^{(3)} = \sqrt{\frac{(\omega^2 \mu_{11} \mu_{22} \varepsilon_{11} - \frac{\pi^2}{a^2} \mu_{11})}{\mu_{22}}} \quad (11)$$

From this equation, we can find μ_{22} . Combining systems (7), (10), (11) we obtain the resulting nonlinear system of equations; solving this system we can find all permeability tensor components: μ_{11} , μ_{22} , μ_{33} .

Now we consider inverse problem $Q1$. Expressing C_1 , D_1 via F we obtain a formula that couples amplitudes A and B :

$$A = B \left(\frac{\gamma_0^2 \mu_{11}^2 + (\gamma_1^{(1)})^2}{\gamma_0^2 \mu_{11}^2 - (\gamma_1^{(1)})^2} - 2i \cot(\gamma_1^{(1)} l_1) \frac{\gamma_0 \gamma_1^{(1)} \mu_{11}}{\gamma_0^2 \mu_{11}^2 - (\gamma_1^{(1)})^2} \right) \quad (12)$$

where $\gamma_1^{(1)} := \gamma_1$, γ_1 is expressed by formula (4). Solution to inverse problem $Q1(Q2)$ repeats that of inverse problem $P1(P2)$ except general formula (7) which is replaced by (12); the tensor transformations are the same. Solving the corresponding nonlinear system of equations, we can find all permittivity (permeability) tensor components: $\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}$ ($\mu_{11}, \mu_{22}, \mu_{33}$).

3.1. Numerical results

In Table 1 numerical results for inverse problem $P1$ are presented. Parameters of the one-sectional diaphragm are $a = 2$ cm, $b = 1$ cm, $c = 2$ cm, and $l_1 = 2$ cm; the excitation frequency $f = 11.93$ GHz.

$$\hat{\mu} = \begin{pmatrix} 1.4 & 0 & 0 \\ 0 & 1.7 & 0 \\ 0 & 0 & 1.5 \end{pmatrix}$$

The first, second, and third columns of the table show, respectively, the values of transmission coefficient $\frac{F}{A}$ and the calculated and true values of the components of the permeability tensor.

In Table 2 numerical results for inverse problem $P2$ are presented.

Parameters of the one-sectional diaphragm are $a = 2$ cm, $b = 1$ cm, $c = 2$ cm, and $l_1 = 2$ cm; the excitation frequency $f = 11.93$ GHz.

$$\hat{\varepsilon} = \begin{pmatrix} 1.1 & 0 & 0 \\ 0 & 1.1 & 0 \\ 0 & 0 & 1.3 \end{pmatrix}$$

Table 1.

$\frac{F^{(i)}}{A}$	Calculated $\hat{\varepsilon}$	True $\hat{\varepsilon}$
$-0.262 - i \cdot 0.965$	$\begin{pmatrix} 1.099908 & 0 & 0 \\ 0 & 1.200046 & 0 \\ 0 & 0 & 1.299904 \end{pmatrix}$	$\begin{pmatrix} 1.1 & 0 & 0 \\ 0 & 1.2 & 0 \\ 0 & 0 & 1.3 \end{pmatrix}$
$-0.241 - i \cdot 0.97$		
$-0.606 - i \cdot 0.795$		
$-0.000789 - i \cdot 0.00128$	$\begin{pmatrix} 1.103124i & 0 & 0 \\ 0 & -1.199928 & 0 \\ 0 & 0 & 1.299904 \end{pmatrix}$	$\begin{pmatrix} 1.1i & 0 & 0 \\ 0 & -1.2 & 0 \\ 0 & 0 & 1.3 \end{pmatrix}$
$-0.012 + i \cdot 0.032$		
$-0.606 - i \cdot 0.795$		

Table 2.

$\frac{F^{(i)}}{A}$	Calculated $\hat{\mu}$	True $\hat{\mu}$
$0.315 - i \cdot 0.945$	$\begin{pmatrix} 1.2003 & 0 & 0 \\ 0 & 1.4126 & 0 \\ 0 & 0 & 1.79996 \end{pmatrix}$	$\begin{pmatrix} 1.2 & 0 & 0 \\ 0 & 1.4 & 0 \\ 0 & 0 & 1.8 \end{pmatrix}$
$-0.411 - i \cdot 0.909$		
$-0.066 - i \cdot 0.998$		
$-0.692 + i \cdot 0.714$	$\begin{pmatrix} 1.80003 & 0 & 0 \\ 0 & 3.26789 & 0 \\ 0 & 0 & -1.499961 \end{pmatrix}$	$\begin{pmatrix} 1.8 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1.5 \end{pmatrix}$
$-0.0046 - i \cdot 0.00442$		
$-0.893 + i \cdot 0.247$		

The first, second, and third columns of the table show, respectively, the values of transmission coefficient $\frac{F}{A}$ and the calculated and true values of the components of the permeability tensor.

In Table 3 numerical results for inverse problem $Q1$ are presented. Parameters of the one-sectional diaphragm are $a = 2$ cm, $b = 1$ cm, $c = 2$ cm, and $l_1 = 2$ cm; the excitation frequency $f = 11.93$ GHz.

$$\hat{\mu} = \begin{pmatrix} 1.4 & 0 & 0 \\ 0 & 1.7 & 0 \\ 0 & 0 & 1.5 \end{pmatrix}$$

The first, second, and third columns of the table show, respectively, the values of reflection coefficient B and the calculated and true values of the components of the permeability tensor.

Table 3.

$B^{(i)}$	Calculated $\hat{\varepsilon}$	True $\hat{\varepsilon}$
$-0.014 + i \cdot 0.023$	$\begin{pmatrix} 1.10067 & 0 & 0 \\ 0 & 1.2033 - i0.036 & 0 \\ 0 & 0 & 1.3011 \end{pmatrix}$	$\begin{pmatrix} 1.1 & 0 & 0 \\ 0 & 1.2 & 0 \\ 0 & 0 & 1.3 \end{pmatrix}$
$0.0067 - i \cdot 0.01$		
$-0.00334 + i \cdot 0.019$		

4. CONCLUSION

In this work, we have developed a method of reconstructing the permeability and permittivity tensors in a rectangular waveguide and a mathematical model of the wave propagation through a one-layered dielectric in a waveguide. For the case of a one-sectional diaphragm, determination of the layer tensor permittivity (permeability) is reduced to the solution of a nonlinear equation system obtained by using a general formula that couples the transmission (reflection) coefficient and the incident field amplitude. We have also determined the tensor transformations that can be implemented in practise. Using the developed method we perform numerical solution to the inverse problem of the permittivity and permeability tensors reconstruction in a broad range of parameter variation. Numerical results confirm the efficiency of the method.

ACKNOWLEDGMENT

This work is partially supported by the Russian Foundation for Basic Research, projects 11-07-00330-a and 12-07-97010-p-a, by a grant of the Ministry of Education and Science of the Russian Federation 14.B37.21.1950, and the Visby Program of the Swedish Institute.

REFERENCES

1. Shestopalov, Y. and Y. Smirnov, "Determination of permittivity of an inhomogeneous dielectric body in a waveguide," *Inverse Problems*, Vol. 27, 095010–095022, 2011.
2. Jackson, J. D., *Classical Electromagnetics*, John Wiley & Sons, New York, 1967.

Parameter Optimization of Waveguide Filters Employing Analysis of Closed-form Solution

P. Tomasek¹ and Y. V. Shestopalov²

¹Tomas Bata University in Zlin, The Czech Republic

²Karlstad University, Sweden

Abstract— This work aims at computation and optimization of transmission coefficients of waveguide filters formed by one- or multi-sectional diaphragms in a waveguide of rectangular cross-section. An approach for designing band-stop and band-pass filters is proposed employing analysis of closed-form solutions and numerical multi-parameter optimization.

1. INTRODUCTION

In this work, we consider parameter optimization of filters created on the basis of multi-layered parallel-plane dielectric diaphragms in waveguides of rectangular cross section. This topic is interesting as far as alternative possibility of designing specific frequency selective filters is concerned. These filters can for example suppress a desired band of frequencies.

The calculation of transmission coefficient is based on the fundamental mathematical background [1,2] and an analytical approach developed in [2–5]. The parameter optimization method and local optimization techniques are described in [6,7] and applied in [11]. The filters under study are frequency selective filters which are multi-layer structures in a waveguide [6,9]. A specific behaviour of the transmission coefficient characteristic for these structures is governed by frequency, parameters of layers, and geometry of a given filter [12].

The first experiments with optimization of filters using evolutionary approach (a genetic algorithm) were performed in [8], further experiments were performed in [14]; the considered filters consisted of a frequency selective surface (FSS) placed on the transverse plane in a rectangular waveguide causing its filtering behaviour. Due to the boundary conditions imposed on the metal waveguide walls, the FSS results to be infinite in extent, allowing the use of fast and efficient method of moment solvers to analyse its properties. A genetic algorithm has been adopted to optimize the filter performance. The techniques [6–8,11,14] that make use of complicated inclusions in waveguides to design FSSs and filters employ approximate numerical schemes for calculating the transmission characteristics that produce non-controllable errors in the filter design.

An approach presented in this article is virtually free from this drawback because it is based on explicit formulas for the transmission coefficient obtained and applied in [7–10] and careful investigation of its properties as a function of several complex variables (see [6]) and therefore provides more robust determination of the filter characteristics and goals. The method allows simultaneous analysis and design of not only narrow band-pass but also wider band-pass and band-stop filters — rectangular waveguides loaded with layered dielectric diaphragms.

2. STATEMENT OF THE PROBLEM

Assume that a single-mode waveguide [13] of rectangular cross section with a perfectly conducting boundary surface contains a body which has the form of a diaphragm, an insert separated into several sections adjacent to the waveguide walls. The waveguide is filled with an isotropic and homogeneous layered medium of constant permittivity. The sections of the diaphragm are filled each with a medium of a constant permittivity (Fig. 1 illustrates an example where a and b corresponds to the width and height of the waveguide respectively, l_i relates to the thickness of a layer).

Solving the forward problem for Maxwell's equations we obtain explicit expressions for the field inside every section of the diaphragm and outside the diaphragm. An explicit expression for the transmission coefficient of a multi-sectional diaphragm

$$F = 2 \prod_{j=0}^n \gamma_j \frac{A e^{i\gamma_0 l_n}}{\gamma_n p_{n+1} + \gamma_0 q_{n+1}} \quad (1)$$

is obtained from a system of equations related to transmission conditions on the boundary surfaces on the diaphragm sections [7–10]; here F is the transmission coefficient, A is the amplitude of the

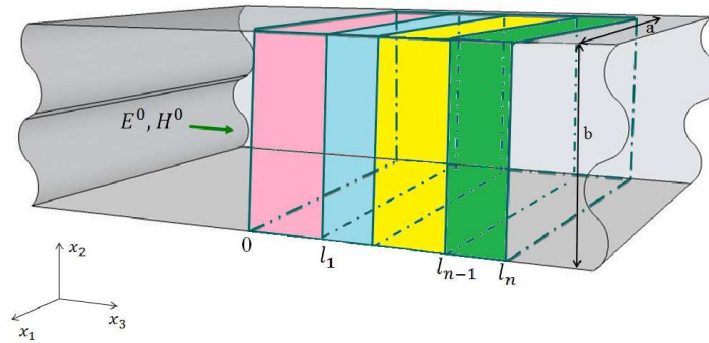


Figure 1: The propagation scheme of multilayered diaphragms in a waveguide [4].

incident wave, n is the number of sections in the diaphragm, and other quantities and symbols γ , p , q , l are explained in [7, 8]. We see that the transmission coefficient is a function of several complex variables; therefore its direct analysis meets substantial difficulties. Although given explicitly, formula (1) can hardly allow one to apply analytical methods for parameter optimization when the number of sections n exceeds two and becomes extremely complicated when n is several tens or more. In view of this fact, numerical optimization procedures should be applied.

3. OPTIMIZATION

A frequency range, an initial geometry with design variables (e.g., width and height of the waveguide cross section), and optimization goals must be set (e.g., for a specific band-stop filter) before performing the optimization of a waveguide filter.

The transmission coefficient depends on frequency and other parameters forming the parameter vector of the filter which specifies the geometry (defined by design variables) and sections positions. A n optimization method searches for the set of parameters which satisfies the given objectives, at least approximately, being thus in a certain sense optimal. In our analysis, we use an explicit representation for the transmission coefficient of a multi-sectional diaphragm (1).

An optimization goal is defined by the frequency ranges where the transmission coefficient must be lower or greater than a threshold value set by a user. In our work, optimization was performed numerically using a local optimizer *fmincon* [15] in Matlab (a possible alternative is *fminsearchbnd* [10]) applied to the result of analysis of the diaphragm transmission coefficient given explicitly by (1).

4. EXPERIMENTS

For test purposes of this study, specific optimization goals were defined: to create a band-stop filter (that suppresses a specific frequency band) applied in the frequency range from 10 to 30 GHz and attenuating frequencies from 19 to 21 GHz while in the lower and higher bands (from 10 to 17 GHz and from 23 to 30 GHz) the frequencies should be passed.

We have made several numerical experiments; the most suitable one which has demonstrated the best results and met the prescribed band-stop filter conditions is presented below. A diaphragm made of twenty sections is used; the materials of the sections are polyimide (relative permittivity 3.4) and air (relative permittivity 1.0006). These materials were repeated ten times to form ten pairs (Fig. 1). Two parameters (design variables) were optimized in this experiment: the thicknesses of each polyimide and air sections. The filter parameter values are as follows: the amplitude of the incident wave is 1.0, the waveguide width and height are, respectively, 20 and 5 mm, and the thicknesses of the sections varies from 0.5 to 50 mm. The optimized transmission coefficients are presented in Fig. 2.

The optimization procedure results in an excellent filter which suppresses the desired band of frequencies and passes the others. The optimized thicknesses are 0.602808 mm for polyimide sections and 6.88111 mm for air sections.

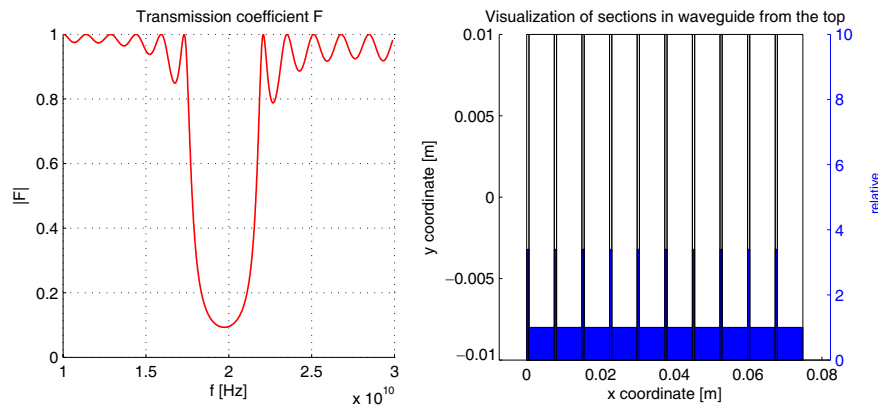


Figure 2: Result of optimization of a 20-sectional diaphragm with respect to two design variables: thicknesses of two periodically changing materials having definite permittivities.

5. CONCLUSION

A method of optimizing multi-layered parallel-plane dielectric diaphragms in waveguides of rectangular cross section is proposed and tested.

Results of numerical experiments confirm the possibility of using rectangular guides loaded with parallel-plane dielectric medium (sections) in the waveguide filter design and construction as a band-stop filter to suppress a desired band of frequencies. Other types of filters, such as band-pass, low-pass, or high-pass filters can be also created and optimized using the proposed technique.

Further work in this direction should prove these theoretical studies by the results of actual measurements.

ACKNOWLEDGMENT

This work was supported by the internal grant of Tomas Bata University in Zlin: IGA/FAI/2013/007 and by the Visby project of the Swedish Institute.

REFERENCES

1. Shestopalov, Y. and Y. Smirnov, "Inverse scattering in guides," *Journal of Physics: Conference Series*, Vol. 346, 012019, 2012.
2. Shestopalov, Y. and Y. Smirnov, "Numerical-analytical methods for the analysis of forward and inverse scattering by dielectric bodies in waveguides," *Proc. 19th Int. Conf. on Microwaves, Radar and Wireless Communications (MIKON-2012)*, 127–132, Warsaw, Poland, 2012.
3. Smirnov, Y., Y. Shestopalov, and E. Derevyanchuk, "Reconstructing permittivity of layered dielectrics in a rectangular waveguide," *Proc. Algebra, Geometry, Mathematical Physics (AGMP) Int. Conf.*, 13 p., Mulhouse, France, 2011, Springer Proceedings Series, 2013, in press.
4. Smirnov, Y., Y. Shestopalov, and E. Derevyanchuk, "Solution to the inverse problem of reconstructing permittivity of an n-sectional diaphragm in a rectangular waveguide," *Proc. Algebra, Geometry, Mathematical Physics (AGMP) Int. Conf.*, 13 p., Mulhouse, France, 2011, Springer Proceedings Series, 2013, in press.
5. Smirnov, Y., Y. Shestopalov, and E. Derevyanchuk, "Permittivity reconstruction of layered dielectrics in a rectangular waveguide from the transmission coefficients at different frequencies," *Proc. Int. Workshop on Inverse Problems and Large-scale Modeling*, 12 p., Sunne, Sweden, May 1–6, 2012, Springer Proceedings Series, 2013, in press.
6. Chan, R., R. Mittra, and T. Cwik, "Techniques for analyzing frequency selective surfaces — A review," *IEEE*, Vol. 76, No. 12, 1593–1615, 1988.
7. Antoniou, A. and W. S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, Springer (India) Pvt. Limited, 2009.
8. Yin, L., Z. Qian, W. Hong, X. W. Zhu, and Y. Chen, "The application of genetic algorithm in E-plane waveguide filter design," *Int. J. Infrared Millimeter Waves*, Vol. 21, 303–308, 2000.
9. Wu, T. K., *Frequency Selective Surfaces and Grid Arrays*, Wiley, 1995.
10. D'Errico, J., "fminsearchbnd, fminsearchcon — File exchange — matlab central [online]," <http://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd>, Feb. 6, 2012, [cit. 2012-12-20].

11. Elhawil, A., G. Koers, L. Zhang, J. Stiens, and R. Vounckx, "Comparison between two optimisation algorithms to compute the complex permittivity of dielectric multilayer structures using a free-space quasi-optical method in W-band," *IET Science, Measurement & Technology*, Vol. 3, No. 1, 13–21, Institution of Engineering and Technology, Stevenage, ROYAUME-UNI, 2009.
12. Goldsmith, P. S., "The quasi optical techniques," *IEEE*, Vol. 80, 1729–1747, 1992.
13. Jackson, J. D., *Classical Electromagnetics*, Wiley, 1967.
14. Monorchioand, A., G. Manara, U. Serra, G. Marola, and E. Pagana, "Design of waveguide filters by using 21 genetically optimized frequency selective surfaces," *IEEE Microwave and Wireless Components Letters*, Vol. 15, 407–409, 2005.
15. The MathWorks Inc. Mathworks Nordic, "Find minimum of constrained nonlinear multivariable function — matlab [online]," <http://www.mathworks.se/help/optim/ug/fmincon.html>, [cit. 2012-12-20].

Computing Ground-Wave Electric Field at MF Band via FDTD Method

F. T. Pachón-García, J. M. Paniagua-Sánchez, M. Rufo-Pérez, and A. Jiménez-Barco
Polytechnic School, University of Extremadura, Av. de la Universidad s/n, Cáceres 10003, Spain

Abstract— This paper presents a study about calculations of electric field at medium frequency (MF: 300 kHz–3 MHz), based on 2D finite-difference time-domain (FDTD) method. Variations of electric field with land conductivity and with typified mountains are evaluated, and the results are validated with Millington expressions and the finite element method-parabolic equation (FEM-PE), agreeing well. The interest of introducing this technique for calculating electric field levels is that it includes reflection, diffraction, surface wave, and takes into account backscattering. Therefore, it provides a more robust solution in non-ideal environments with difficult terrain, being this a first step to analyse how suitable its use is, in terms of time consuming and accuracy of results for this frequency band.

1. INTRODUCTION

Ground wave propagation has been widely discussed since its inception, both from a theoretical and practical [1, 2] point of view. However new developments around applicability of computational techniques, highly costly in resources, are making that more robust and accurate methods can be applied for calculating electric field levels in real environments with all its peculiarities, analysing conductivity changes as well as terrain elevations. Concerning this issue, new techniques are being applied to the calculation of ground wave [3]. Note the practical implementation of the methods of finite-difference time-domain (FDTD) introduced by A. Taflové [4] and used by many authors [5–9] to analyse radio propagation of signal, and to design antennas or to implement circuits in different frequency bands. The disadvantage of FDTD is the high required processing level, but lately the tendency is changing with the use of supercomputers and GPUs (Graphics Processing Units) [10].

The present study implements a 2D FDTD to be used for the calculation of electric field levels along certain profiles, quantifying the variations of levels. The model will be validated on the one hand, and for flat land, with a proprietary application that incorporates Millington expressions using the UexWAVE software, based on GRWAVE from ITU [11], and on the other hand, for typified Gaussian mountains, with the FEMIX software [12, 13] made by IEEE members.

It should be emphasized that among the virtues of the FDTD technique is to consider all propagation phenomena such as reflection, diffraction, surface wave, backscattering and forward propagation, thus providing a complete solution [14], although with high utilization of computational resources.

2. MODEL AND FORMULATION

It has been implemented a computer program named ‘MF-FDTD’ (Medium Frequency-FDTD), which is based on FDTD technique, considering the cases of mountainous terrain with different sizes, and with specific radio-electric constants.

The application MF-FDTD is designed for the calculation of electric field levels in two dimensions, and it assumes that the structure 2D is extended infinitely in the direction or axis that is not considered; in our case, all physical and electromagnetic parameters are only function of the x and y coordinates [4]. Maxwell’s equations can be decomposed into two sets of equations, the TE (transverse electric) and TM modes (transverse magnetic) [15]. Then, for our purpose, the electric field is taken perpendicular to the z axis, with the working components H_z , E_x and E_y . Actually, considering a 2D space is a simplification of a 3D real environment, where reflections occurring in other objects, that are not in the plane of study, may affect to our plane, nevertheless it serves to obtain a first approximation of the level fluctuations, with manageable calculations without using excessive computational resources.

Calculation equations are shown in expressions (1), (2) and (3), where ε (F/m), μ (H/m), σ (S/m) and σ^* (Ω /m) are the permittivity, permeability, conductivity and equivalent magnetic losses (the latter considered 0 in this paper). Meanwhile the index i , j are the number of cell in the x and y direction respectively and n refers to timing instants. The magnetic component is calculated

at discrete timing steps Δt , $2\Delta t$, $3\Delta t$, \dots , $n\Delta t$, and electrical components are calculated at $t/2$, $3\Delta t/2$, \dots , $(n + 1/2)\Delta t$.

$$E_x|_{i,j+1/2}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i,j+1/2}\Delta t}{2\varepsilon_{i,j+1/2}}}{1 + \frac{\sigma_{i,j+1/2}\Delta t}{2\varepsilon_{i,j+1/2}}} \right) E_x|_{i,j+1/2}^{n-1/2} + \left(\frac{\frac{\Delta t}{\varepsilon_{i,j+1/2}}}{1 + \frac{\sigma_{i,j+1/2}\Delta t}{2\varepsilon_{i,j+1/2}}} \right) \left(\frac{H_z|_{i,j+1}^n - H_z|_{i,j}^n}{\Delta y} \right) \quad (1)$$

$$E_y|_{i-1/2,j+1}^{n+1/2} = \left(\frac{1 - \frac{\sigma_{i-1/2,j+1}\Delta t}{2\varepsilon_{i-1/2,j+1}}}{1 + \frac{\sigma_{i-1/2,j+1}\Delta t}{2\varepsilon_{i-1/2,j+1}}} \right) E_y|_{i-1/2,j+1}^{n-1/2} + \left(\frac{\frac{\Delta t}{\varepsilon_{i-1/2,j+1}}}{1 + \frac{\sigma_{i-1/2,j+1}\Delta t}{2\varepsilon_{i-1/2,j+1}}} \right) \left(-\frac{H_z|_{i,j+1}^n - H_z|_{i-1,j+1}^n}{\Delta x} \right) \quad (2)$$

$$H_z|_{i,j+1}^{n+1} = \left(\frac{1 - \frac{\sigma_{i,j+1}^*\Delta t}{2\mu_{i,j+1}}}{1 + \frac{\sigma_{i,j+1}^*\Delta t}{2\mu_{i,j+1}}} \right) H_z|_{i,j+1}^n + \left(\frac{\frac{\Delta t}{\mu_{i,j+1}}}{1 + \frac{\sigma_{i,j+1}^*\Delta t}{2\mu_{i,j+1}}} \right) \left(\frac{E_x|_{i,j+3/2}^{n+1/2} - E_x|_{i,j+1/2}^{n+1/2}}{\Delta y} - \frac{E_y|_{i+1/2,j+1}^{n+1/2} - E_y|_{i-1/2,j+1}^{n+1/2}}{\Delta x} \right) \quad (3)$$

In relation to the source, a Gaussian pulse is used, and for the termination of the simulation space, PML (Perfectly Matched Layer) absorbing boundary condition is conformed. In terms of selecting temporal-spatial resolution, Courant condition has been followed [4].

3. SIMULATION RESULTS

In simulations, the source and receptor are located one cell above the ground. Spatial resolution is the same as x and y direction being 12, 10 and 6 m the resolution of the cell for the frequencies of 0.5, 1 and 2 MHz, respectively. The power of the transmitter is fixed to 1 kW for simplicity when comparing results with validation software. The atmosphere has been simulated with values of standard refractivity index [16].

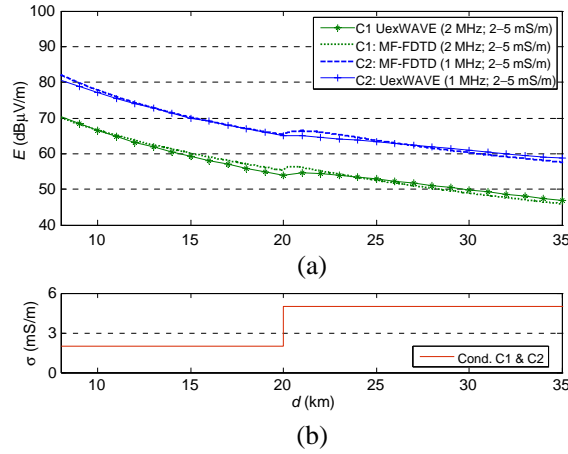


Figure 1: (a) E field variation with distance, (b) considering changes of conductivity for 1 and 2 MHz. Dashed line FDTD, and solid one Millington application.

Firstly, profiles with changes in conductivities and validation with UexWAVE software is presented. Mixed paths are exposed, with changes in terrain conductivity, having chosen two frequencies, 1 and 2 MHz. A transition of terrain values has been included $\{\sigma, \varepsilon_r\}$, from $\{2 \text{ mS/m}, 20.5\}$ to $\{5 \text{ mS/m}, 26.5\}$ quantifying the variation in levels. Comparing the results with the curves from UexWAVE application, we can observe in Figure 1 an almost perfect fit, although some differences are detected in the transitions. These variations are in part justified by the numerical approximations of Millington method, as indicated in ITU 368 [11], which can produce errors up to 3 dB. These differences are greater just in the transition, throwing the FDTD method more oscillating values than Millington numerical approximation (differences up to 2 dB), and they are justified globally because FDTD is a full propagation method.

Secondly, mountainous terrain and validation with FEMIX software is exposed. To calibrate and validate these results, FEMIX program [12] was used; it is based on finite element method-parabolic equation (FEM-PE) and although it only includes forward propagation and not backscattering, it can serve to set the behavior of the signal and to use the output as a reference. The selected type of mountain is the standard Gaussian shape, with 1500 m high, intended so to get noticeable effects.

Concerning the frequencies to compute along the frequency band under study, 0.5, 1 and 2 MHz have been chosen, being in this way the results easily contrastable. Regarding the time consuming, the worst case was the associated with 2 MHz, with more than 7 hours in the simulation.

Overall, Figure 2 shows how in front of the mountain the level increases just when the altitude of the mountain begins to climb (and this is consistent both in FDTD and FEM-PE curves) and once the maximum height of the mountain is passed, a sharp decline in the levels is observed in the area that can be named ‘shadow zone’, suffering a gradual recovery of signal as we move away from the mountain, tending the FDTD line toward the GRWAVE curve which excludes elevated terrains.

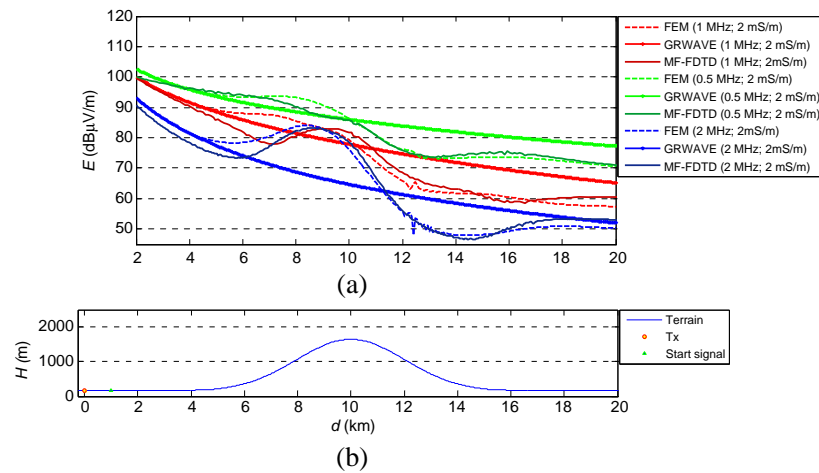


Figure 2: (a) E field variation with distance at 0.5, 1 and 2 MHz, (b) considering a terrain with a Gaussian mountain with 1500 m high, centered at 10 km from the transmitter. Solid line FDTD method, and dashed line FEMIX software. Terrain constants $\{\sigma: 2 \text{ mS/m}, \epsilon_r: 10\}$.

Note that the longer the wavelength, the better the obstacles are dodged, that is to say, less oscillation detected by including the mountain in the propagation path. At higher frequencies, larger oscillations of levels and more affectation of the obstacle is noticed, as it was expected. It should be stressed that because of FDTD includes backscattering and that the shown FDTD line is the result of a smoothing process of the original signal (to avoid small fluctuations due to the grid used) there are some notable differences in the front of the mountain, for example in front of the case of 1 MHz.

Thirdly, a general case of influence of irregularities in the terrain is studied. In this case 1 MHz is considered in a scenario that includes three mountains with different heights and widths as shown in the bottom of Figure 3, being the first hill 250 m high; the second mountain, located far away

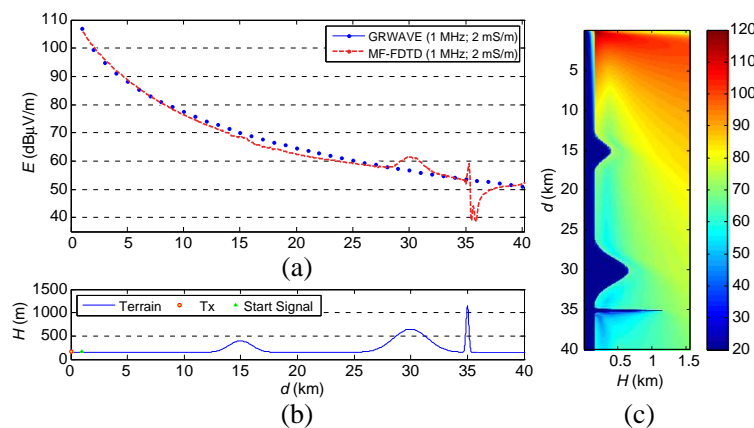


Figure 3: (a) E field variation at 1 MHz with distance when (b) considering three mountains in the path with terrain constants $\{\sigma: 2 \text{ mS/m}, \epsilon_r: 10\}$. Dashed line: FDTD; solid line GRWAVE. (c) Level distribution (dB μ V/m) in the environment.

from the first, with 500 m high, and placed at 30 km from the transmitter, is very close to the third one with 1000 m high and a shape denominated ‘Knife Edge’. Evaluating the results, a soft oscillation with little relevance in levels is obtained at the first hill, a level increase in front of the second one and a posterior decay behind it (where the effect of backscattering coming from the third one must be added) is noticed, and a large decrease in level behind the third mountain is detected, that gradually is being recovered to levels close to those that GRWAVE curve shows. This example illustrates very well the signal behavior in a changing environment. On the right side of the figure, there is an annexed graph of the simulated scenario which is very illustrative of the level distribution.

4. CONCLUSION

It has been created an application based on 2D FDTD to calculate the electric field levels with the distance at Medium Wave, taking into account the affectation of the inclusion of mountains along the path and the changes in soil conductivity. On the one hand, different tests in a flat ground have been performed, with changes of conductivities, comparing the results with those coming from a program which is based on GRWAVE application from ITU, but incorporating the Millington method. On the other hand, it has been tested in cases where there are mountains or elevations of terrain, being the results contrasted with those coming from FEMIX software, which is based on finite element method-parabolic equation, agreeing the results very well.

ACKNOWLEDGMENT

This work has been partially supported by the ‘Gobierno de Extremadura, Consejería de Empleo Empresa e Innovación’ and the ‘European Social Fund’. The authors thank to the Professors G. Apaydin, and L. Sevgi the facilitation of the encrypted program FEMIX to perform some validations, and also to CénitS (Centro Extremeño de iNvestigación, Innovación Tecnológica y Supercomputación) its support for carrying out some simulations in Lusitania supercomputer.

REFERENCES

1. Guerra, D., U. Gil, D. Vega, et al., “Medium wave digital radio mondiale (DRM) field strength time variation in different reception environments,” *IEEE Transactions on Broadcasting*, Vol. 52, No. 4, 483–491, 2006.
2. DeMinco, N., “Propagation prediction techniques and antenna modeling (150 to 1705 kHz) for intelligent transportation systems (ITS) broadcast applications,” *IEEE Antennas and Propagation Magazine*, Vol. 42, No. 4, 9–34, 2000.
3. Sevgi, L., “Groundwave modeling and simulation strategies and path loss prediction virtual tools,” *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 6, 1591–1598, 2007.
4. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, Artech House, Inc., Norwood, MA, 2000.
5. Zhou, L., X. Xi, J. Liu, et al., “LF ground-wave propagation over irregular terrain,” *IEEE Transactions on Antennas and Propagation*, Vol. 59, No. 4, 1254–1260, 2011.
6. Faghihi, F. and H. Heydari, “Time domain physical optics for the higher-order FDTD modeling in electromagnetic scattering from 3-D complex and combined multiple materials objects,” *Progress In Electromagnetics Research*, Vol. 95, 87–102, 2009.
7. Simpson, J. J. and A. Taflove, “A review of progress in FDTD Maxwell’s equations modeling of impulsive subionospheric propagation below 300 kHz,” *IEEE Transactions on Antennas and Propagation*, Vol. 55, No. 6, 1582–1590, 2007.
8. Gao, S., L. W. Li, and A. Sambell, “FDTD analysis of a dual-frequency microstrip patch antenna,” *Progress In Electromagnetics Research*, Vol. 54, 155–178, 2005.
9. Akleman, F. and L. Sevgi, “A novel finite-difference time-domain wave propagator,” *IEEE Transactions on Antennas and Propagation*, Vol. 48, No. 3, 839–842, 2000.
10. Liu, Y., Z. Liang, and Z. Q. Yang, “A novel FDTD approach featuring two-level parallelization on PC cluster,” *Progress In Electromagnetics Research*, Vol. 80, 393–398, 2008.
11. International Telecommunications Union (ITU), “Ground-wave propagation curves for frequencies between 10 kHz and 30 MHz,” ITU-R P.368-9, Geneva, 2007.
12. Apaydin, G. and L. Sevgi, “Numerical investigations of and path loss predictions for surface wave propagation over sea paths including hilly island transitions,” *IEEE Transactions on Antennas and Propagation*, Vol. 58, No. 4, 1302–1314, 2010.

13. Apaydin, G. and L. Sevgi, “FEM-based surface wave multimixed-path propagator and path loss predictions,” *IEEE Antennas and Wireless Propagation Letters*, Vol. 8, 1010–1013, 2009.
14. Sullivan, D. M., *Electromagnetic Simulation Using the FDTD Method*, Institute of Electrical and Electronics Engineers, Inc., New York, 2000.
15. Sevgi, L., F. Akleman, and L. B. Felsen, “Groundwave propagation modeling: Problem-matched analytical formulations and direct numerical techniques,” *IEEE Antennas and Propagation Magazine*, Vol. 44, No. 1, 55–75, 2002.
16. International Telecommunications Union (ITU), “The radio refractive index: Its formula and refractivity data,” ITU-R P.453-10, Geneva, 2012.

Longitudinal Voltages, Induced by Parallel Overhead Transmission Lines Magnetic Field

N. B. Rubtsova², M. Sh. Misrikhanov¹, S. G. Murzin¹, and A. Yu. Tokarskij³

¹JSC Federal Network Company, Moscow, Russian Federation

²Federal State Budgetary Institution ‘Research Institute of Occupational Health’ under the Russian Academy of Medical Sciences, Moscow, Russian Federation

³JSC Federal Network Company Branch “Main Power Networks of the Center”
Moscow, Russian Federation

Abstract— Longitudinal voltage values, induced by magnetic fields currents in parallel overhead transmission line are determined usually by derivative of Carson’ integral equations. These equations application have broad “dead space” conditionally the spacing of transmission lines and ground conductivity that lead to great calculative error. The method of longitudinal voltage values induced by magnetic field in parallel overhead transmission lines calculation for any distance between lines is suggested with full coincidence with Carson’ equation calculation (in work space) without “dead space” (with high calculation error).

Under parallel overhead transmission line (TL) repair work linemen safety requires to know the voltage that equal voltage (electromotive force) induced in repaired TL by working parallel TL currents magnetic field (MF), since linemen may fall this longitudinal voltage under conductor disconnection possibility.

Two parallel single-wire TL treat: line 1 and line 2 disposed at “ a_{12} ” distance apart (see Fig. 1). Line 1 is live at operative mode, and current \dot{I}_1 flows in its wire, line 2 is disconnected, and its section l length is grounded on ends. Accept equation $\mu_A = \mu_E = \mu_0$ for air and earth magnetic conductivity. Reverse line, similar to line 2 wire, will arrange under line 2 wire at a depth h_w . This line 2 wire will metal connect with grounding conductor’ slopes closed circuit “wire 2 — grounding conductor—reverse wire” forming. Magnetic field created by \dot{I}_1 direct current induces \dot{E}_2 voltage in derived circuit that calculates excluding earth conductivity (earth is ideal dielectric) by equation:

$$\dot{E}_2 = -j \frac{\omega \mu_0 l \dot{I}_1}{4\pi} \ln \frac{a_{12}^2 + (h_1 + h_w)^2}{a_{12}^2 + (h_1 - h_2)^2} = -j \frac{\omega \mu_0 l \dot{I}_1}{2\pi} \ln \frac{r_{1R.w.}}{r_{12}}, \quad (1)$$

where: $\omega = 2\pi f$ — angular frequency, f — AC current \dot{I}_1 frequency, μ_0 — permeability of vacuum.

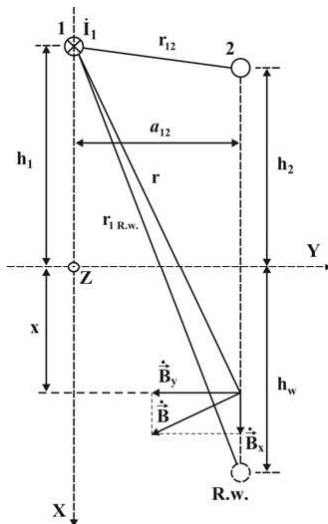


Figure 1: Two parallel single-wire TL.

In case of length l reverse wire section of line 2 is absent, and earth has finite resistivity ρ_E , voltage \dot{E}_2 may be determined by equation $F_{12} = -2jJ(r, \theta)$, where $J(r, \theta) = P + jQ$ — is Carson's integral [1, 2], r and θ — integral characteristics:

$$r = r'_{12} \sqrt{\omega\mu_0/\rho_E}, \quad \theta = \arctg a_{12}/(h_1 + h_2), \quad r'_{12} = r_{1R.w.} \quad \text{at } h_w = h_2. \quad (2)$$

\dot{E}_2 voltage is determined by equation [1, 2]:

$$\dot{E}_2 = -j \frac{\omega\mu_0 l \dot{I}_1}{2\pi} \left(\ln \frac{r'_{12}}{r_{12}} + F_{12} \right), \quad F_{12} = 2 \int_0^\infty \frac{1}{\nu + \sqrt{\nu^2 - (k_E^2 - k_0^2)}} e^{-\nu(h_1+h_2)} \cos \nu a_{12} d\nu, \quad (3)$$

where k_E — earth characteristic parameter, k_0 — air characteristic parameter, $k_E^2 - k_0^2 = -j\omega\mu_0/\rho_E$. F_{12} is series expansion [1, 2], and for $r \leq 0.25$ parameter is calculate by expansion:

$$F_{12} = -0.0772 - \ln \frac{r}{2} + \frac{2 \cos \psi}{3} r \cos \theta - j \left(\psi - \frac{2 \sin \psi}{3} r \cos \theta \right), \quad (4)$$

and for $r \geq 5$ — by expansion:

$$F_{12} = \frac{2}{r} \left(\cos \theta e^{-j\psi} - \frac{\cos 2\theta}{r} e^{-j2\psi} + \frac{\cos 3\theta}{r^2} e^{-j3\psi} - \frac{3 \cos 5\theta}{r^4} e^{-j5\psi} \right). \quad (5)$$

Taking into account that $\psi = \pi/4$, $\sin \psi = \cos \psi = \sqrt{2}/2$, from the first Equation (2) $r'_{12} = r\sqrt{\rho_E/(\omega\mu_0)}$, intake [2] $\delta_E = \sqrt{2\rho_E/(\omega\mu_0)}$ — as the earth penetration depth (the depth of electromagnetic wave attenuation in $e = 2.72$ times), will result $r'_{12} = \frac{r\delta_E}{\sqrt{2}}$, $\cos \theta = \frac{h_1+h_2}{r'_{12}} = \sqrt{2} \frac{h_1+h_2}{r\delta_E}$, from first Equation (2) consideration (3) will receive for $r \leq 0.25$ parameter:

$$\dot{E}_2^{(6)} = -j \frac{\omega\mu_0 l \dot{I}_1}{2\pi} \left[\ln \frac{\sqrt{2}\delta_E}{r_{12}} - j \frac{\pi}{4} + \frac{2}{3} \frac{h_1 + h_2}{\delta_E} (1 + j) - 0.0772 \right]. \quad (6)$$

Similarly for $r \geq 5$ parameter:

$$\dot{E}_2^{(7)} = -j \frac{\omega\mu_0 l \dot{I}_1}{2\pi} \left[\ln \frac{r'_{12}}{r_{12}} + \frac{2}{r} \left(\cos \theta e^{-j\psi} - \frac{\cos 2\theta}{r} e^{-j2\psi} + \frac{\cos 3\theta}{r^2} e^{-j3\psi} - \frac{3 \cos 5\theta}{r^4} e^{-j5\psi} \right) \right], \quad (7)$$

where $r'_{12} = \sqrt{a_{12}^2 + (h_1 + h_2)^2}$ and $r_{12} = \sqrt{a_{12}^2 + (h_1 - h_2)^2}$.

There is break space or “dead space” on r parameter, from 0.25 to 5, in which Equations (6) and (7) have very high error, and as result — miscount.

From the Equation (2) receive:

$$r'_{12} = \sqrt{a_{12}^2 + (h_1 + h_2)^2} = r \left| \sqrt{j\omega\mu_0/\rho_E} \right| = r \cdot \left| \sqrt{\rho_E/(j\omega\mu_0)} \right|.$$

If consider that $a_{11'} \gg h_1 + h_{1'}$, then $h_1 + h_{1'} \approx 0$ may be accepted, in that case:

$$a_{12} = r \left| \sqrt{\rho_E/(j\omega\mu_0)} \right| = r \sqrt{\rho_E/(\omega\mu_0)} = r\delta_E / \sqrt{2}. \quad (8)$$

“Dead space” borders for earth resistivity ρ_3 different values are shown in Table 1: $a_{12}^{\delta_E \max}$ — maximal distances between lines under which expression (6) use is possible, and minimal $a_{12}^{\delta_E \min}$ distance, from which expression (7) use is admissible.

Consequently, voltages induced by parallel lines magnetic field calculation by equations derived by Carson's integral with δ_E electromagnetic wave earth penetration depth use, in the r parameter disruption space gives incorrect result.

Table 1: $a_{12 \max}^{\delta_E}$ and $a_{12 \min}^{\delta_E}$ values for ρ_E earth resistivity.

ρ_E	Oh · m	1	5	10	50	100	500	1000
$a_{12 \max}^{\delta_E}$	m	13	28	40	89	126	281	398
$a_{12 \min}^{\delta_E}$	m	252	563	796	1800	2516	5627	7958

For this defect elimination, we introduce into voltage induced by line 2 in “ l length wire — grounding conductor — earth” circuit calculation reverse current [2] in this circuit and earth equivalent depth:

$$h_{EQ} = \frac{2}{\gamma'} \sqrt{\frac{e\rho_E}{\omega\mu_0}} = \frac{2}{\gamma'} \sqrt{\frac{e}{2\pi\mu_0}} \sqrt{\frac{\rho_E}{f}} = \frac{2}{1,781} \sqrt{\frac{2,718}{2\pi \cdot 4\pi \cdot 10^{-7}}} \sqrt{\frac{\rho_E}{f}} = 658,898 \sqrt{\frac{\rho_E}{f}} \approx 660 \sqrt{\frac{\rho_E}{f}},$$

where: e — natural logarithmic base, $\gamma' = 1,781$ from Euler' constant $\ln \gamma' = 0.5772$, f — line current frequency.

We assume that magnetic field created by \dot{I}_1 current, penetrate into the earth not more than h_{EQ} m and reverse wire of concerned line 2 circuit is placed at h_{EQ} depth. In that case substitution h_{EQ} instead of h_w into Equation (1), we derive the expression for calculation of voltage induced by created \dot{I}_1 current in line 1 wire MF in line 2 grounded circuit:

$$\dot{E}_{2w1} = -j \frac{\omega\mu_0 l \dot{I}_1}{4\pi} \ln \frac{a_{12}^2 + (h_1 + h_{EQ})^2}{a_{12}^2 + (h_1 - h_2)^2}. \quad (9)$$

Current \dot{I}_1 magnetic field induces in the earth electric field (EF), its strength \dot{E}_1 is determined by expression:

$$\dot{E}_1(x, y) = -j \frac{\omega\mu_0 \dot{I}_1}{4\pi} \ln \frac{y^2 + (h_1 + h_{EQ})^2}{y^2 + (h_1 + x)^2}.$$

Considering the earth as isotropic medium we can determine current density $\dot{\eta}_1(x, y)$ in it created by EF strength $\dot{E}_1(x, y)$ induced by \dot{I}_1 current MF by equation:

$$\dot{\eta}_1(x, y) = \frac{\dot{E}_1(x, y)}{\rho_E} = -j \frac{\omega\mu_0 \dot{I}_1}{4\pi\rho_E} \ln \frac{y^2 + (h_1 + h_{EQ})^2}{y^2 + (h_1 + x)^2}.$$

The currents induced in the earth by magnetic field are vortex.

Generated by $\dot{\eta}_1(x, y)$ current density MF induction $\dot{B}_{\eta y}(X, Y)$ component, is determined by expression:

$$\dot{B}_{\eta y}(X, Y) = \frac{\mu_0}{2\pi} \int_{y^-}^{y^+} \int_0^{2h_{EQ}} \frac{\dot{\eta}_1(x, y) (X - x)}{(X - x)^2 + (Y - y)^2} dx dy.$$

$\dot{E}_{2\eta}$ voltage, generated in grounding part of line 2 wire circuit by magnetic flux of $\dot{B}_{\eta y}(X, Y)$ induction, is determined as:

$$\dot{E}_{2\eta} = -\frac{\omega^2 \mu_0^2 \dot{I}_1 l}{8\pi^2 \rho_E} \int_{-h_2}^{h_{EQ}} \int_{y^-}^{y^+} \int_0^{2h_{EQ}} \ln \frac{y^2 + (h_1 + h_{EQ})^2}{y^2 + (h_1 + x)^2} \frac{(X - x)}{(X - x)^2 + (a_{12} - y)^2} dx dy dX. \quad (10)$$

Since line 1 is single-wire, the \dot{I}_1 current of \dot{E}_1 source (Fig. 2) passes through the wire 1 to Z_L load' resistance the another end of its is grounded, and passing through the earth in the form of reverse current \dot{I}_{RC1} returns to the grounding end of \dot{E}_1 source.

Considering all earth volume from the surface to $2h_{EQ}$ depth as isotropic medium, elementary reverse current $d\dot{I}_{RC1}$, passing through elementary channel with length

$$l_{RC} + 2r = l_{RC} + 2\sqrt{x^2 + y^2}$$

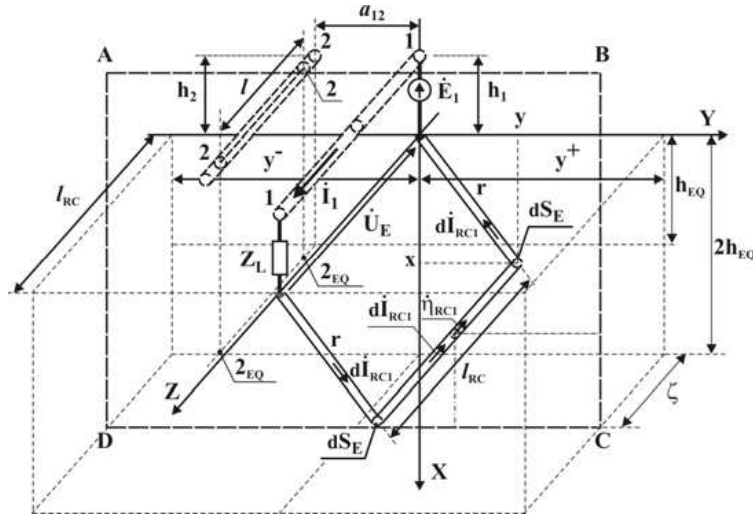


Figure 2: To the definition of $\dot{\eta}_{OT}$ line 1 reverse current to the earth density

and cross-sectional area dS_E , may be evaluated as:

$$d\dot{I}_{RC1} = \frac{\dot{U}_E}{\rho_E (l_{RC} + 2\sqrt{x^2 + y^2})} dS_E,$$

where \dot{U}_E — is voltage between source and load groundings.

In $ABCD$ plane, transversely to line phase wires, disposed at ζ distance from \dot{E}_1 source, reverse current density $\dot{\eta}_{RC}$ is determined by equation:

$$\dot{\eta}_{RC}(x, y) = \dot{\eta}_{RC} = \frac{\dot{U}_E}{\rho_E (l_{RC} + 2\sqrt{x^2 + y^2})}.$$

Since reverse current is line 1 current \dot{I}_1 but passes in reverse direction, the voltage \dot{U}_E value is determined by expression:

$$\dot{U}_E = - \frac{\dot{I}_1}{\int_{y^-}^{y^+} \int_0^{2h_{EQ}} \frac{1}{\rho_E (l_{RC} + 2\sqrt{x^2 + y^2})} dx dy}. \quad (11)$$

MF induction $\dot{B}_{\eta_{RC}y}(X, Y)$ component OY -axis, created by reverse current \dot{I}_{RC1} , is determined by expression:

$$\dot{B}_{\eta_{RC}y}(X, Y) = \frac{\mu_0}{2\pi} \int_{y^-}^{y^+} \int_0^{2h_{EQ}} [\dot{\eta}_{RC}(x, y) (X - x)] / [(X - x)^2 + (Y - y)^2] dx dy.$$

Voltage \dot{E}_{2RC} , created in line 2 wire grounded part circuit by magnetic flux of $\dot{B}_{\eta_{RC}y}(X, Y)$ induction, can be determined by equation:

$$\dot{E}_{2RC} = -j \frac{\omega \mu_0 l}{2\pi} \int_{-h_1}^{h_{EQ}} \int_{y^-}^{y^+} \int_0^{2h_{EQ}} \frac{\dot{U}_E (X - x)}{\rho_E (l_{RC} + 2\sqrt{x^2 + y^2}) [(X - x)^2 + (a_{12} - y)^2]} dx dy dX.$$

Voltage full value is calculated as:

$$\dot{E}_2^{(12)} = \dot{E}_{2w1} + \dot{E}_{2\eta} + \dot{E}_{2RC}. \quad (12)$$

We can examine \dot{E}_2 value changes, calculated by expressions (6) $\dot{E}_2^{(6)}$ and (7) $\dot{E}_2^{(7)}$ that were received by Carson' integral, and by Equation (12) $\dot{E}_2^{(12)}$ under distance between lines a_{12} increase from 10 to 50000 m, take the earth resistivity value $\rho_E = 50 \text{ Ohm} \cdot \text{m}$. Line 1 current $\dot{I}_1 = 4000 \text{ A}$, its frequency $f = 50 \text{ Hz}$. Then reverse current equivalent depth $h_{EQ} = 660 \text{ m}$. Taking $y^+ = -y^- = 100000 \text{ m}$, by expression (11) we can calculate voltage value $\dot{U}_E = -50.13 \text{ V}$.

There are shown in Fig. 3 the changes of voltage modules \dot{E}_{2w1} , $\dot{E}_{2\eta}$, \dot{E}_{2RC} , $\dot{E}_2^{(12)}$ and $\arg(\dot{E}_2^{(12)})$ under a_{12} distance increase from 100 to 2000 m.

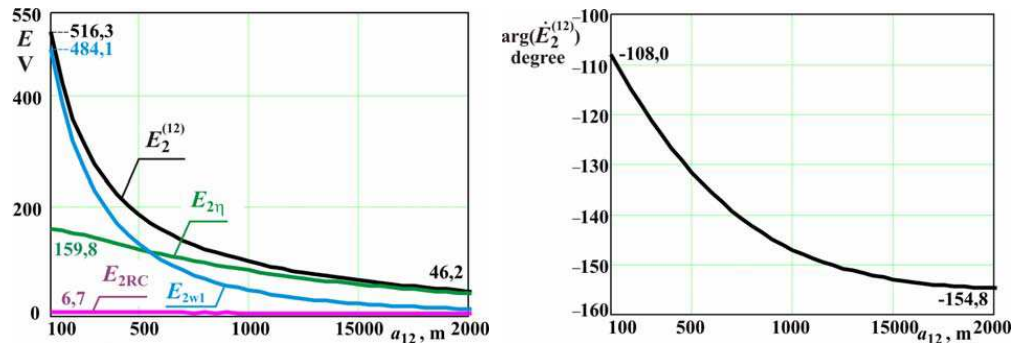


Figure 3: Voltage module \dot{E}_{2w1} , $\dot{E}_{2\eta}$, \dot{E}_{2RC} , $\dot{E}_2^{(12)}$ and $\arg(\dot{E}_2^{(12)})$ values for $100 \leq a_{12} \leq 2000 \text{ m}$.

Figure 4 shows superposed curves of voltage modules $\dot{E}_2^{(6)}$, $\dot{E}_2^{(7)}$, \dot{E}_{2w1} and arguments $\dot{E}_2^{(12)}$ under a_{12} distance increase from 100 m to 2000 m (in “dead space” by r parameter).

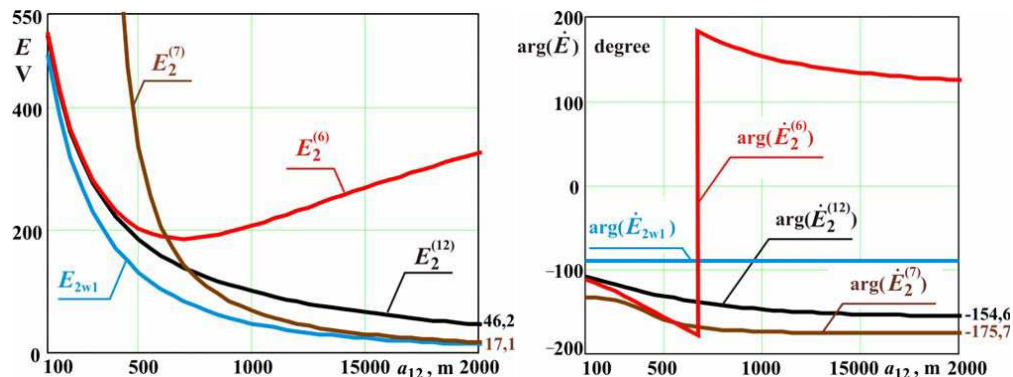


Figure 4: Voltage module and argument changes $\dot{E}_2^{(6)}$, $\dot{E}_2^{(7)}$, \dot{E}_{2w1} and $\dot{E}_2^{(12)}$ under a_{12} distance increase from 100 m to 2000 m.

There is good match of voltage module and argument curves $\dot{E}_2^{(6)}$ and $\dot{E}_2^{(12)}$ up to “dead space”. Voltage curves $\dot{E}_2^{(6)}$ and $\dot{E}_2^{(7)}$ by r parameter in “dead space” (see Fig. 4) are in the area of large errors while voltage module $\dot{E}_2^{(12)}$ curve changes smoothly without going into the zone of large errors. After “dead space” by r parameter ($a_{12} > 1800 \text{ m}$) voltage module $\dot{E}_2^{(7)}$ curve continuous to approach and coincide with voltage module \dot{E}_{2w1} curve practically, but remain in lower position than voltage module $\dot{E}_2^{(12)}$ and $\dot{E}_{2\eta}$ curves.

Discussed mathematical model allows to define voltage induced in parallel lines its current magnetic field more exactly, especially in Carson' integral “dead space” by r parameter.

REFERENCES

1. Carson, J. R., “Wave propagation in overhead wires with ground return,” *Bell Syst. Techn. J.*, Vol. 5, No. 4, 1926.
2. Kostenko, M. V., L. S. Perelman, and Y. P. Shkarin, “Wave processes and electrical disturbances in multi-wire high voltage lines,” *M.: Energy*, 272, 1973 (in Russian).

Effective HF Modeling of Passive Devices Based on Frequency Dependent Hodge Operators and Model Order Reduction

Gabriela Ciuprina, Daniel Ioan, and Mihail-Iulian Andrei

Politehnica University of Bucharest

Spl. Independentei 313, Bucharest 060042, Romania

Abstract— The paper proposes an efficient method for the modeling of high frequency electromagnetic field effects, such as skin or proximity effects, inside on-chip metallic conductors. Compact sub-models are extracted by using an electromagnetic field discretization approach based on the Finite Integration Technique with frequency dependent Hodge operators and model order reduction based on fitting and adaptive frequency sampling.

1. INTRODUCTION

Designers of radio-frequency (RF) integrated circuits (ICs) require models for passive components which describe all relevant electromagnetic (EM) field effects at high frequency (HF). These effects are quantified by the Maxwell's equations for the the full wave (FW) electromagnetic field, which are, from mathematical point of view, partial differential equations to which uniqueness conditions are added (e.g., boundary conditions in the case of frequency domain simulations).

From the design point of view, the quantities of interest are S -parameters defined at intentional terminals placed on the domain's boundary and that is why results in most papers dealing with EM modeling of RF effects finally refer to computations and measurements of quantities defined at terminals and not to details on the field distribution inside the computational domain [1].

Thus, not only the correct definition of the physical and mathematical model of the field problem is important, but also the state/semi-state space system associated to the continuous model above. This model has distributed parameters, an infinite dimension state space, and a finite number of input and output signals, related to the device terminals. In order to obtain a model with a finite number of degrees of freedom (DoFs), that can be implemented and solved, a numerical method is needed to discretize the continuous model. If this numerical approach is a classical variant of a mesh or cell based method such as the finite element method (FEM) or the finite integration technique (FIT) then, in order to describe field effects at high frequency such as skin and proximity effects, the element or cell dimensions have to be much less than the skin depth, leading to models that require an extremely intensive computational effort to be solved or reduced, both with respect to the computational time and the memory needed. That is why special approaches were proposed to deal with these high frequency effects.

Usually, the conductor is excluded from the computational domain and an impedance type boundary condition is imposed on the new boundary created due to this exclusion. Depending on how these impedance boundary conditions are defined, various methods with different degrees of accuracies are obtained, that can be implemented into numerical methods such as FDTD, FEM, DG-FEM [2]. In [3] a surface impedance concept is used as well, that rely on a surface current sheet model. This time, the thick conductors are not excluded from the domain, but are divided into two equal sheet conductor layers for which the currents are not enforced, yielding improved models for resistance and inductance. Another approach is to use asymptotic expressions for the solution in the conducting subdomains [4].

In this paper, we propose an alternative technique to model the high frequency magneto-quasi-static (MQS) field effects inside metallic conductors with high conductivity ($\sigma \gg \omega\epsilon$). The conductors are not excluded from the computational domain, but their material dependencies are defined by means of frequency dependent Hodge (FredHO) operators so that HF effects are taken into account. The approach was integrated into an in-house developed FIT code, and its validation was carried out for a problem consisting of two coupled inductors, for which measurements are available. This paper continues the research in [5,6] and it is similar, but not identical with Transmission Line Matrix (TLM) and Surface Impedance (SIBC) methods.

2. HODGE OPERATORS IN CLASSICAL FIT

In FIT, two Yee type staggered orthogonal grids are used as discretization mesh [7]. The centers of the primary (electric) cells are the nodes of the secondary (magnetic) cells and the secondary grid

is not extended outside the primary grid. The degrees of freedom are global variables, i.e., vectors of electric and magnetic voltages \mathbf{u}_e , \mathbf{u}_m and magnetic and electric fluxes φ , ψ assigned to the grid elements (edges and faces), in a coherent manner. By applying the global form of electromagnetic field equations on the mesh elements (elementary faces and their borders), a system of differential algebraic equations (DAE), called Maxwell Grid Equations (MGE) is obtained. The MGE set of equations are completed with the equations that describe the material behavior. For instance, these constitutive relationships required in the MQS regime are:

$$\mathbf{J} = \sigma \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}. \quad (1)$$

Their correspondent in the FIT variables are given by relationships that relate the conduction currents flowing through the edges of the primary grid to the voltages associated to the same edges and the magnetic fluxes flowing through the edges of the secondary grid to the magnetic voltages associated to the same edges:

$$i_k = G_k u_{e_k}, \quad \varphi_k = R_{m_k} u_{m_k}, \quad (2)$$

where G_k are electric conductances associated to the edges of the primary grid and R_{m_k} are magnetic reluctances associated to the edges of the secondary grid. By writing these constitutive relationships for all the edges, the compact matrix form is obtained

$$\mathbf{i} = \mathbf{M}_\sigma \mathbf{u}_e, \quad \varphi = \mathbf{M}_\mu \mathbf{u}_m = \mathbf{M}_\nu^{-1} \mathbf{u}_m, \quad (3)$$

where \mathbf{M}_σ and \mathbf{M}_μ are known as Hodge operators which are constant diagonal matrices in classical FIT, built by independent averaging of material constants $\nu = 1/\mu$ and σ over each cell. Typically, when the primary grid is generated according to the material interfaces, every primary edge has 4 neighboring cells that may have different materials, and every secondary edge passes through exactly two electric cells so that it may cross at most 2 different materials (Fig. 1). That is why, typical expressions of the components of the Hodge operators are given by

$$M_{\sigma_k} = \sum_{j=1}^4 G_{k,j} = \frac{1}{l_k} \sum_{j=1}^4 \sigma_j A_j, \quad M_{\nu_k} = \sum_{j=1}^2 R_{m_{kj}} = \frac{1}{A_k} \sum_{j=1}^2 \frac{l_j}{\mu_j}, \quad (4)$$

and represent a parallel connection of 4 d.c. conductances and a series connection of 2 magnetic reluctances. Finally, MGE augmented with material relationships are completed with the discrete representation of appropriate boundary conditions, such as the ones of electric circuit element type (ECE), as used in [8].

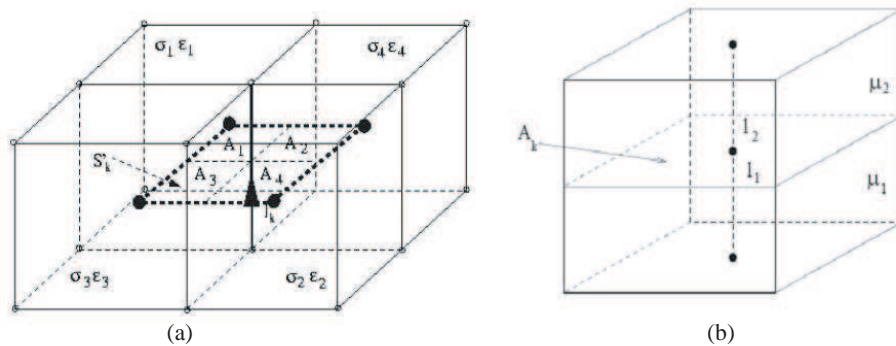


Figure 1: Notations useful for the definition of Hodge operators: an electric edge (a) of length l_k has 4 neighboring electric cells of conductivities $\sigma_1, \dots, \sigma_4$; the cells contribute with A_1, \dots, A_4 , which are 1/4 of the areas faces, in the formula of HO in classical FIT; a magnetic edge (b), of length $l_1 + l_2$ is included in only two electric cells of permeabilities μ_1 and μ_2 , going through an electric face of area A_k .

3. FREQUENCY DEPENDENT HODGE OPERATORS

In order to describe field effects at high frequency such as skin and proximity effects, the cell dimensions have to be much less than the skin depth $\delta = \sqrt{2/(\omega\mu\sigma)}$, which is $6.7 \mu\text{m}$ for Cu at

100 GHz and $15\ \mu\text{m}$ at 20 GHz. In order to keep the number of cells at a reasonable level, non-uniform grids could be used, with peripheral cells smaller than internal ones. Even so, the number of cells required by a reasonable accuracy can be relatively high.

This drawback can be avoided if the frequency independent (d.c.) Hodge operators described in the previous section are replaced with appropriate ones, i.e., frequency dependent (computed in a.c.), that describe high field effects in conductors. In order to derive this appropriate expression, the complex Helmholtz equation for the electric field in a rectangular homogeneous cell (one cell of Fig. 1(a)) having the conductivity σ and the permeability μ , of dimensions: a (along the Ox axis), $2b$ and $2c$ (along Oy and Oz , respectively) is solved by the method of variable separation. Its complex admittance of the cell along the Ox direction is

$$Y = \frac{8}{\pi^2 R_0} \sum_{m=1}^{\infty} \frac{1}{(2m-1)^2} \left[\frac{\tanh(\alpha_m b)}{\alpha_m b} + \frac{\tanh(\beta_m c)}{\beta_m c} \right], \quad (5)$$

where $R_0 = a/(4\sigma bc)$ is the d.c. resistance of the analyzed cell along the Ox direction, and the complex numbers α_m and β_m are given by

$$\alpha_m = \sqrt{\gamma^2 + \left[\frac{(2m-1)\pi}{2c} \right]^2}, \quad \beta_m = \sqrt{\gamma^2 + \left[\frac{(2m-1)\pi}{2b} \right]^2}, \quad (6)$$

where $\gamma^2 = i\omega\mu\sigma$ is the complex diffusion constant in the conductor. Relation (5) does a smooth connection between the d.c. value R_0 and the value given by a strong skin depth formula $a/(4(b+c)\delta)$.

The admittance Y in (5) is a complex quantity and, in the FredHO approach we propose, its formula divided by 4 replaces the simple expression (4) of G_{kj} used in classical FIT, where $A_j = bc$ and $l_k = a$.

The validation of the FredHO code was carried out on a simple test case (straight conductor with rectangular cross-section, placed in air) having an analytical solution taken as reference. The result obtained with classical FIT with a uniform grid and 770 DoFs was at an error of 31% far away from the analytical solution, whereas a classical FIT with a non-uniform grid and 667 DoFs was at an error of 3% far away from the analytical solution. When using FIT with FredHO, the conductor does not need to be discretized at all and that is why a relative error of 0.0006% was obtained with only 5 DoFs. The error was computed for the frequency dependence of both the a.c. resistance and the inner inductivity as shown in Fig. 2.

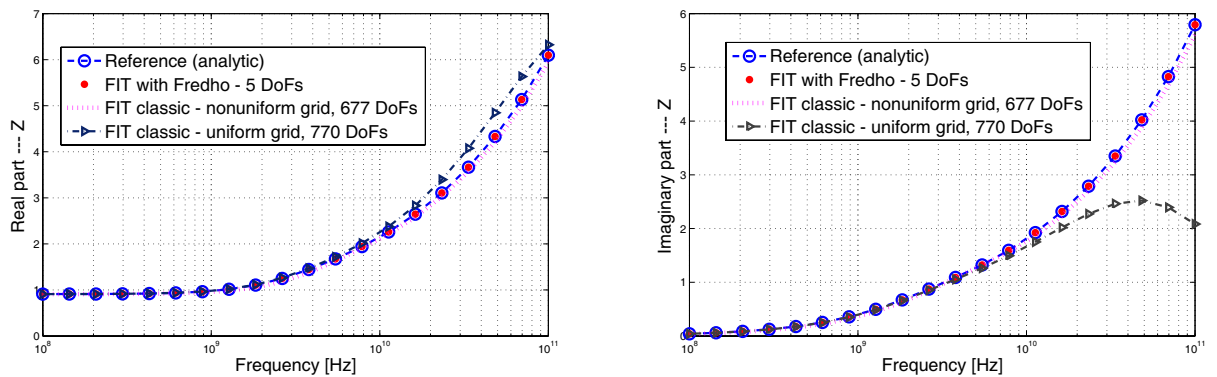


Figure 2: Frequency characteristic (real and imaginary parts of impedance) — with or without FredHO.

4. MODEL ORDER REDUCTION

By applying the classical FIT with ECE boundary conditions, a linear time invariant system is assembled:

$$\mathbf{C} \frac{d\mathbf{x}(t)}{dt} + \mathbf{G}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{L}\mathbf{x}(t), \quad (7)$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state space vector of size n , $\mathbf{u}, \mathbf{y} \in \mathbb{R}^m$ are the vector of inputs and outputs signals respectively, $\mathbf{B} \in \mathbb{Z}^{n \times m}$ and $\mathbf{L} \in \mathbb{Z}^{m \times n}$ are selection matrices and $\mathbf{C}, \mathbf{G} \in \mathbb{R}^{n \times n}$ state

matrices with very sparse structures. The number of degrees of freedom n is the state space size, related to the number of grid cells. The quantity m is the number of terminals.

A very efficient model order reduction (MOR) algorithm for this class of problems is Vector Fitting (VF) [9] that starts from the values of the transfer function

$$\mathbf{Y}_{\text{FIT}}(\omega) = \mathbf{L}(\mathbf{G} + j\omega\mathbf{C})^{-1}\mathbf{B}, \quad (8)$$

in a set of given frequency samples and finds the best rational approximation of this frequency characteristic. The VF order reduction procedure uses as input a set of values $(\omega_k, \mathbf{H}(\omega_k))$, $k = 1, F$, where F is the number of frequency samples and it identifies the poles p_i , the residual matrices \mathbf{K}_i and the constant terms \mathbf{K}_∞ , \mathbf{K}_0 of the rational approximation

$$\mathbf{H}_{\text{VF}}(\omega) = \sum_{i=1}^q \frac{\mathbf{K}_i}{j\omega - p_i} + \mathbf{K}_\infty + j\omega\mathbf{K}_0, \quad (9)$$

for $\mathbf{H}_{\text{FIT}}(\omega)$. Here q is the order of the reduced system, which is automatically found. The accuracy of the models extracted by VF depends on the number and distribution of frequency samples. For the problems we consider, there is no prior knowledge of these frequency characteristic samples and that is why a method to generate an optimal list of samples, by Adaptive Frequency Sampling (AFS), aiming to minimize the approximation error was implemented. Thus, a robust and efficient VF-MOR procedure was obtained [10]. It is very important to note that the last version of the VF algorithm does a passivity check and enforce it, if necessary.

The contributions from the Hodge operators in MQS affect only the state space matrix \mathbf{G} and that is why in the FIT with FredHO implementation, the matrix \mathbf{G} is recomputed for every frequency sample, and the correct expression of the complex admittance is

$$\mathbf{Y}_{\text{FIT+FredHO}}(\omega) = \mathbf{L}(\mathbf{G}(\omega) + j\omega\mathbf{C})^{-1}\mathbf{B}. \quad (10)$$

For an efficient implementation of the matrix assembly, only the contribution of the conducting cells have to be recomputed for each frequency to be solved for, by using (5) for each metallic cell.

5. RESULTS FOR A REAL BENCHMARK

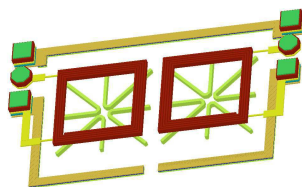


Figure 3: Two coupled spiral inductors test.

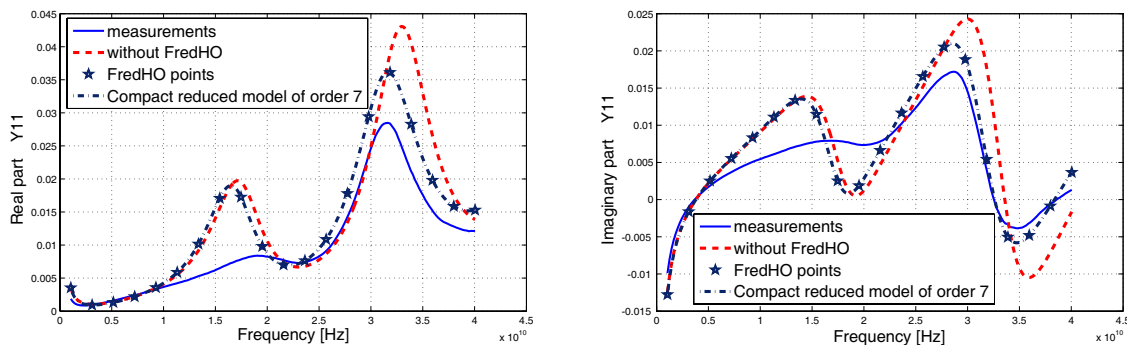


Figure 4: Frequency characteristic — with or without FredHO for two coupled spiral inductors.

The final validation was carried out for a more complex problem consisting of two coupled inductors (Fig. 3), for which measurements are available. Simulations with and without the use of

FredHO were done (Fig. 4). As expected, at low frequencies the simulations are almost identical, whereas at high frequencies the use of FredHO improves the solution behavior, the global relative error with respect to the measurements decreasing from 29% to 19%.

The computational effort required by FredHO is mainly spent into the preprocessing step (matrix assembly), and depends on the volume occupied by high conducting parts, for which MQS with FredHO is used. In the case of problems for which a low reduced order model is expected, only a small number of frequency points have to be computed (e.g., we have used 20 points, marked with stars in Fig. 4), computation followed by an automatic model order reduction based on vector fitting. Thus, a final reduced model of order 7 is obtained, for which a compact netlist is synthesized with a differential equation macromodel procedure of Palenius and Roos (see reference in [8]). The final reduced compact netlist consists of lumped circuit elements with constant parameter values, whose behavior is able to catch the frequency dependence of the original field problem.

6. CONCLUSION

This paper proposed the modeling of high frequency electromagnetic field effects in highly conductive materials by using frequency dependent Hodge operators. This alternative to methods discussed in the literature and based on the surface impedance concepts avoids the removal of metallic subdomains from the computational domain, thus simplifying the geometric modeling. Validation of the FredHO operators integrated into a finite integration technique implementation, combined with model order reduction based on vector fitting was carried out for a real benchmark. The use of FredHO itself can be seen as an a-priori model order reduction technique, since it decreases drastically the number of degrees of freedom that are otherwise needed in models that do not leave the high conductive parts outside the computational domain.

ACKNOWLEDGMENT

The authors thank the Joint Research Project PN-II-PT-PCCA-2011-3, of UEFISCDI, No. 5/2012.

REFERENCES

1. Stavarakis, K., et al., "Fast parameter sweeps for the calculation of S -parameters in electromagnetic field simulations," *Int. Conf. on Electromagnetics in Advanced*, 2012.
2. Hiltunen, I., et al., "Broad band surface impedance boundary conditions for higher order time domain discontinuous Galerkin method," *Scientific Computing in Elec. Eng., Zurich*, 2012.
3. Agilent EEs of EDA, "Accurate modeling of spiral inductors on silicon using planar EM simulation," <http://cp.literature.agilent.com/litweb/pdf/5989-9526EN.pdf>, 2008.
4. Schmidt, K. and A. Chernov, "Robust families of transmission conditions of high order for thin conducting sheets," INS Preprint No. 1102, 2011.
5. Ioan, D. and M. Piper, "FIT models with frequency dependent Hodge operators for HF effects in metallic conductors," *Progress In Electromagnetics Research Symposium*, Pisa, 2004.
6. Ciuprina, G., et al., "Frequency parameterized models for planar on-chip inductors," *Scientific Computing in Elec. Eng., Zurich*, 2012.
7. Weiland, T., "A discretization method for the solution of Maxwell's equations for 6 component fields," *AE, Electronics and Communication*, Vol. 31, 116–120, 1977.
8. Ioan, D. and G. Ciuprina, "Reduced order models of on-chip passive components and interconnects, workbench and test structures," *Model Order Reduction: Theory, Research Aspects and Applications*, Vol. 13, 447–467, W. Schilders, et al., editors, Springer, 2008.
9. Gustavsen, B. and A. Semlyen, "Rational approximation of frequency responses by vector fitting," *IEEE Trans. Power Delivery*, Vol. 14, 1052–1061, 1999.
10. Lazăr, I. A., et al., "Effective extraction of accurate reduced order models for HF-IC using multi-CPU architectures," *Inverse Problems in Science and Engineering*, Vol. 20, No. 1, 15–27, 2011.

Image Restoration of Two-dimensional Signal Sources with Superresolution

B. A. Lagovsky and A. B. Samokhin

Moscow State Institute of Radio Engineering and Automation (Technical University), Russia

Abstract— The inverse two-dimensional problem of restoration the angular amplitude distribution of the signal source is investigated. A new method of digital processing allows reconstructing the image of objects with superresolution substantiated. Algebraic methods for digital processing of received signals by special algorithms allow to obtain additional information about the angular distribution of the intensity. It is equivalent to increasing the effective angular resolution of the system. We propose a pre-symmetrization method for two-dimensional signals, which can increase the achieved level of angular superresolution.

1. INTRODUCTION

One of the main ways to improve monitoring and goniometric systems in different ranges of electromagnetic waves — from radio to optical — is to increase the angular resolution. Generalization of the known methods to obtain superresolution [1, 2] for two-dimensional image restoration problem is essentially complicates the algorithms, greatly increases the processing time of signals.

Considered promising for two-dimensional problems of image restoration with superresolution radiation sources are algebraic methods for solving one-dimensional problems. They were to present the approximate solutions in the form of finite expansions given by a sequence of functions with unknown coefficients. Presented in [3–5] algebraic methods for solving one-dimensional problems are considered perspective for two-dimensional image reconstruction problems of signal sources. They consist in the presenting approximate solutions in the form of finite expansions given by a sequence of functions with unknown coefficients. A generalization of algebraic methods to two-dimensional problem does not lead to a serious complication of algorithms. The numerical solution time varies slightly.

2. PROBLEM STATEMENT

Let the scan sector of angle measuring system is the solid angle Ω_0 . Use a narrow radiation pattern $f(x, y)$, x, y — the deflection angles from zero direction in the Cartesian coordinate system. The object of observation with unknown angular size Ω is located within the scan sector.

Denote the required two-dimensional distribution of the amplitude of the reflected or emitted by the source signal $I(x, y)$. The received signal while scanning the sector $U(x, y)$ is a two-dimensional convolution:

$$U(x, y) = \int_{\Omega} f(x - x', y - y') I(x', y') dx' dy' \quad (1)$$

The problem is to restore the angular distribution $I(x, y)$ based on the analysis of the received signal $U(x, y)$ and the known pattern with maximum possible angular resolution exceeding the Rayleigh criterion:

$$\delta\theta \cong \lambda/d \quad (2)$$

where d — the size of the aperture, λ — wavelength.

Posed problem belongs to a class of inverse. It is ill-posed because not satisfy the second and third requirement of well-posed problem (Hadamard). Attempts to increase the resolution in comparison with the classical quantity (2) by solving the Equation (1) lead to the appearance of instabilities in the solutions.

3. MAIN RESULTS

To find an approximate solution of the two-dimensional problem by algebraic methods in the beginning, based on the type of $U(x, y)$, the position and size of the solid angle Ω , in which is located the source of the signals are evaluated. In the future, based on preliminary solutions obtained during the iteration process, the size Ω and the location of source are refined.

The required distribution $I(x, y)$ will be found in the form of an expansion in a finite system of orthogonal in Ω two-dimensional functions. Only separable systems of functions are used, i.e., in the form of $G_{n,m}(x, y) = g_n(x)g_m(y)$. This restriction is justified because any two-dimensional set with a finite number of non-zero readings can be decomposed into a finite sum of separable sequences.

Introducing the $I(x, y)$ as an expansion in a finite system of N^2 functions:

$$I(x, y) \cong \sum_{n=1}^N \sum_{m=1}^N b_{n,m} g_n(x) g_m(y) \quad (3)$$

we obtain a received signal:

$$U(x, y) \cong \sum_{j,k=1}^N b_{j,k} \varphi_{j,k}(x, y), \quad (4)$$

$$\varphi_{j,k}(x, y) = \int_{\Omega} f(x - x', y - y') g_j(x') g_k(y') dx' dy'. \quad (5)$$

The coefficients $b_{j,k}$, which minimize the mean square deviation of the signal on the basis of (3), are found by solving a system of linear equations:

$$\int_{\Omega} U(x, y) \varphi_{j,k}(x, y) d\alpha = \sum_{n,m=1}^N b_{n,m} \int_{\Omega} \varphi_{j,k}(x, y) \varphi_{n,m}(x, y) dx dy \quad j, k = 1, 2, \dots, N \quad (6)$$

Thus, carried out parameterization of the inverse problem and the solution are reduced to solving a system of Equation (6).

The principal feature of the resulting systems of the type (6) is their ill-conditioning, which is a consequence of attempts to solve the inverse problem. With the increasing number of functions used in the expansion (3) increases the angular resolution and, at the same time sharply increase — according to the exponential law — the condition number of algorithms. So solutions become less and less stable.

Algebraic methods allow gradually approach a limiting angular resolution for each problem to be solved by successive increase the number of used features. The maximum achievable angular resolution is limited by the level of random components in the signal and the accuracy of measurements.

4. SIMULATION RESULTS

Quantitative characteristics of increasing angular resolution and its boundaries were studied on a mathematical model. Considered a narrow radiation pattern formed by a planar square array size $30 \times 30d/\lambda$. The investigated object was described as the angular distribution of the reflected signal amplitude is identically equal to zero outside the solid angle Ω ($-\theta_{0.5}/2 \leq x \leq \theta_{0.5}/2, -\theta_{0.5}/2 \leq y \leq \theta_{0.5}/2$) and in the form of a complex function within that sector. Border Ω_0 are the angles x and y , in which the received signal is reduced by half compared with the maximum value.

On the next step of the simulation we solve the inverse problem — restoration using $U(x, y)$ the amplitude angular distribution of the signal source by means of solutions the system (6). The problems were solving with increasing resolution, i.e., sequential increasing the number functions used.

Figure 1 shows an example of a source image reconstruction with large gradients of the angular distribution of the signal amplitude, and localization of small areas compared to the width of the beam. Figure 1 shows the results of image reconstruction of the signal source, which consists of four sources of varying intensity with small angular size. Sources are not resolved under the direct supervision without signal processing (curve 1). Horizontal axes on Figure 1 — x and y -axis.

Prior information about the solution of the problem consisting of the fact that the source is a set of several closely spaced small-sized sources was used.

Maximum number of functions used in the solution, was 16. Achieved effective resolution on axes was found to be $1/4\theta_{0.5}$, i.e., 4 times higher than the Rayleigh criterion. It was also achieved good localization signal sources, which amounted to one quarter $1/4\theta_{0.5}$ for each of the coordinates.

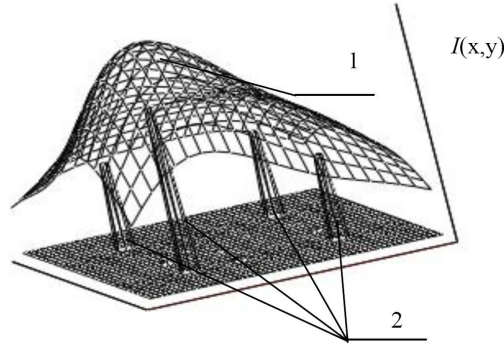


Figure 1: The reconstructed image of four point signal sources with different intensities. 1 — received signal $U(x, y)$; 2 — restored sources.

The correct number of objects and were found. Their locations were found almost exactly. In contrast to many known methods any false source does not appear. The intensities were found with a small error, which amounted to 2.5%.

5. SYMMETRIZATION OF THE INVERSE PROBLEM

Increasing the number of used functions in (3) can improve the angular resolution. This means an increase in the dimension of the matrices (6) and, as a consequence, decrease the stability of solutions. Thus, there is a maximum number of functions M , provide adequate solutions to obtain the maximum achievable resolution. However, there is possibility to increase the angular resolution without reducing the stability of solutions. It is proposed to carry out pre-symmetrization of the inverse problem. For one-dimensional problems, this method is based on the fact that it is always possible to provide the received signals U and I in the form of a solution even amount U_0 , I_0 and odd U_e , I_e parts:

$$\begin{aligned}
 I_o(\alpha) &= 1/2(I(\alpha) + I(-\alpha)), & I_e(\alpha) &= 1/2(I_o(\alpha) - I_e(-\alpha)), & I(\alpha) &= 1/2(I_o(\alpha) + I_e(\alpha)), & (7) \\
 U_o(\alpha) &= \int_{\Omega} f(\alpha - \phi)I_o(\phi)d\phi & U_e(\alpha) &= \int_{\Omega} f(\alpha - \phi)I_e(\phi)d\phi \\
 U_o(\alpha) &= 1/2(U(\alpha) + U(-\alpha)) & U_e(\alpha) &= 1/2(U(\alpha) - U(-\alpha)) & (8)
 \end{aligned}$$

As a result, due to the linearity, the initial problem of finding $I(\alpha)$ decomposes into two. First — find even part of $I_o(\alpha)$ for even part of the received signal based on the selected system even functions. Second — find the odd part of $I_e(\alpha)$ on the odd part $U_e(\alpha)$ on the basis of an odd system of functions. The general solution of the problem — the superposition of the even and odd solutions

$$I(\alpha) = 1/2(I_o(\alpha) + I_e(\alpha)). \quad (9)$$

If each of the problems can obtain a stable solution with N functions, the final sum solution includes $2N$ functions. At the direct solution of the problem as a superposition of $2N$ functions would be required to solve a system of $2N$ equations. In the present problem it is sharply — on the order, reduce the stability of solutions which is estimated by condition number.

Thus, for the one-dimensional problems symmetrization doubles the resolution without reducing stability of the solution. For two-dimensional problems, we introduce the concept of dual parity — if the function is even in x and y , then we call it even-even if is even in x and odd in y , let's call it even-odd, etc.. We represent the received signal and the solution as the sum of four components with different parity form $U_{o,o}$, $U_{o,e}$, $U_{e,o}$ and $U_{e,e}$.

$$\begin{aligned}
 U_{o,o}(x, y) &= 1/4(U(x, y) + U(-x, y) + U(x, -y) + U(-x, -y)) \\
 U_{e,o}(x, y) &= 1/4(U(x, y) - U(-x, y) + U(x, -y) - U(-x, -y)) \\
 U_{o,e}(x, y) &= 1/4(U(x, y) + U(-x, y) - U(x, -y) - U(-x, -y)) \\
 U_{e,e}(x, y) &= 1/4(U(x, y) - U(-x, y) - U(x, -y) + U(-x, -y))
 \end{aligned} \quad (10)$$

To represent the solutions of the system now requires 4 systems of functions with different parity.

For example, based on 4 products of trigonometric functions:

$$\begin{aligned} & \cos\left(\frac{2\pi mx}{T_x}\right) \cos\left(\frac{2\pi my}{T_y}\right); \quad \sin\left(\frac{2\pi mx}{T_x}\right) \cos\left(\frac{2\pi my}{T_y}\right); \\ & \cos\left(\frac{2\pi mx}{T_x}\right) \sin\left(\frac{2\pi my}{T_y}\right); \quad \sin\left(\frac{2\pi mx}{T_x}\right) \sin\left(\frac{2\pi my}{T_y}\right), \end{aligned} \quad (11)$$

where T_x and T_y — the dimensions of the decision region. Due to the linearity the original problem is now divided into four problems of different parity. The complete solution is a superposition of all four solutions:

$$I(x, y) = 1/4(I_{e,o}(x, y) + I_{o,e}(x, y) + I_{o,o}(x, y) + I_{e,e}(x, y)). \quad (12)$$

On a mathematical model the problem of reconstructing 4 small size sources with smoothly inhomogeneous intensity distribution was investigated. In Figure 2, the received while scanning region signal is shown without described above digital processing.

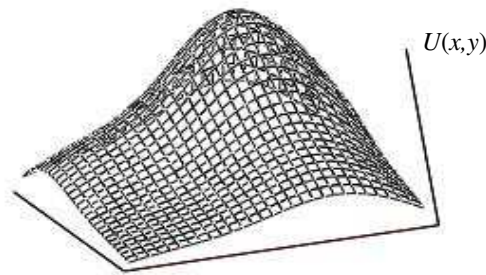


Figure 2: The received signal while scanning. Its digital processing allowed to restore the image sources.

Figure 3 shows the source (Figure 3(a)) and the reconstructed image of source in solving the inverse problem on the basis of symmetrization (Figure 3(b)).

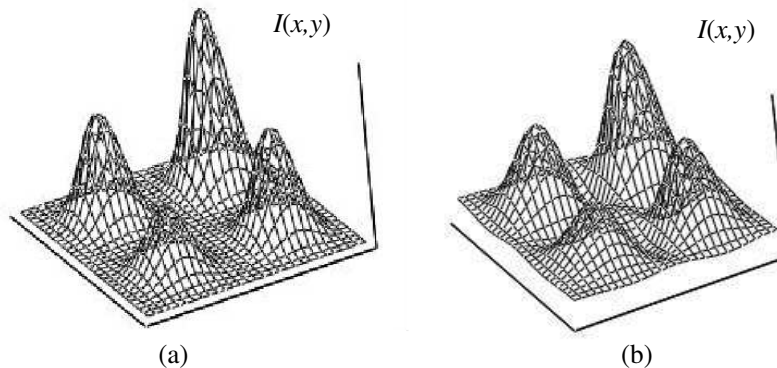


Figure 3: Image sources. (a) The real source of the signal. (b) Reconstructed image.

Four functions (11) were used as the systems of orthogonal functions. The resulting solution (12) enabled to allow all 4 sources. The location and character of the distribution of intensity were found correctly. Amplitude values were reconstructed with minor errors — within 7%.

6. CONCLUSION

Algebraic methods of processing received signals allow reconstructing two-dimensional image signal sources with relatively small errors. The use of the algebraic methods greatly increases the effective angular resolution and provides an opportunity to achieve superresolution. Achieved resolution is 4 times higher than the Rayleigh criterion by using simple algorithms and 5–8 times or more by using more complicated algorithms based on the symmetrization of the problem.

REFERENCES

1. Quinquis, A., E. Radoi, and F. Totir, "Some radar imagery results using superresolution techniques," *IEEE Trans. Antennas and Propagation*, Vol. 52, No. 5, 1230–1244, 2004.
2. Herman, M. A. and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Processing*, Vol. 57, No. 6, 2275–2284, 2009.
3. Lagovsky, B. A., "Superresolution: Simultaneous orthogonalization of function systems describing the received signal and its source," *PIERS Proceedings*, 993–996, Moscow, Russia, Aug. 19–23, 2012.
4. Lagovsky, B. A. "High performance angular resolution algorithm for radar systems," *PIERS Proceedings*, 1637–1641, Moscow, Russia, Aug. 18–21, 2009.
5. Lagovsky, B. A., "Image restoration of the objects with superresolution on the basis of spline — Interpolation," *PIERS Proceedings*, 989–992, Moscow, Russia, Aug. 19–23, 2012.

A Direct Spectral Domain Method for Near-ground Microwave Radiation by a Vertical Dipole above Earth in the Presence of Atmospheric Refractivity

R. Bhattacharjee^{1,2}, C. R. Anderson², and G. D. Durgin¹

¹Georgia Institute of Technology, USA

²United States Naval Academy, USA

Abstract— A spectral domain method is presented to calculate the potential of a vertical dipole in a multilayered medium as a model of long-range radio propagation. The spectral domain Green's function (SDGF) for structures with up to hundreds of thousands of layers is calculated using an efficient matrix formulation, thus enabling simulation of continuously stratified media. The SDGF is sampled to perform pole/residue-extraction using contour quadratures and to perform a novel asymptotic Filon-Clenshaw-Curtis (FCC) quadrature to calculate the far-field Green's functions. The near-fields are directly calculated using an adaptive Clenshaw-Curtis quadrature without pole-extraction. Results of numerical simulations are presented in the case of a graded-index waveguide and an atmospheric gradient-layer above a realistic lossy ground.

1. INTRODUCTION

Electromagnetic (EM) modeling of refractive and terrain effects over long links has been modeled by ray tracing [1, 17, 20] or by using the parabolic/paraxial wave equation [3, 4, 7]. These methods are approximations to the wave propagation physics of Maxwell's equations, which can limit their applicability. Other approaches have included the moving-window finite-difference time-domain (MWFDTD) method [14, 16, 21, 22]. While capturing the relevant physics, MWFDTD leaves something to be desired with respect to intuition about the fields; a list of field strength values is not readily understood in terms of propagation mechanisms such as direct waves, reflected waves, interface/ground waves, and guided propagation modes.

On the other hand, the theory of Sommerfeld integrals (SIs) and spectral domain Green's functions (SDGFs) in multilayered media has been used in the simulation of printed circuit board (PCB) structures such as transmission lines and antennas [8]. The SDGF/SI technique expresses the EM fields as SIs, which must be evaluated numerically using a variety of methods [2, 13, 15, 18]. With certain approaches, the solutions decompose into terms that directly correspond to the aforementioned propagation mechanisms. The key contribution of the present work is the application of these rigorous full-wave techniques to long range radio frequency (RF) propagation. Quadrature methods used for PCBs are not directly applicable because of the differences in length scales involved. We therefore use novel asymptotic quadrature to calculate far-fields, an approach that has not, to the best of our knowledge, been attempted for RF propagation before.

2. ATMOSPHERIC AND EM MODEL

The problem geometry is a non-magnetic, multilayered dielectric material bounded by two half-spaces with interfaces parallel to the xy -plane that contains a z -directed, time harmonic ($e^{-j\omega t}$ time dependence) dipole radiator at a height z' above the origin, as in Figure 1. Each layer is homogeneous with index of refraction $n_\ell = \sqrt{\varepsilon_0 \varepsilon_{r\ell}}$, where $\varepsilon_{r\ell}$ is the relative permittivity of the ℓ^{th} medium and $\varepsilon_0 \approx 8.85418782 \times 10^{-12}$ F/m is the permittivity of free space. The media wavenumbers are $k_\ell = n_\ell k_0$, $k_0 = 2\pi f/c_0$, f is the frequency of the wave, and $c_0 = 299,792,458$ m/s is the speed of light in vacuum. The EM fields in this geometry can be expressed in terms of only the z component of the magnetic vector potential, $A_z(\rho, z, z')$, where ρ is the xy -distance away from the dipole and z is the height above the origin. A_z obeys a forced wave-equation in the source layer, and an unforced one in the others. The equation reduces to the 3D Helmholtz equation under harmonic time dependence, which is indicated in Figure 1.

The spatial domain potential can be shown to admit a spectral (Sommerfeld) integral representation given by

$$A_z(\rho, z, z') = \int_0^\infty \tilde{A}_z(k_\rho, z, z') J_0(k_\rho \rho) k_\rho dk_\rho, \quad (1)$$

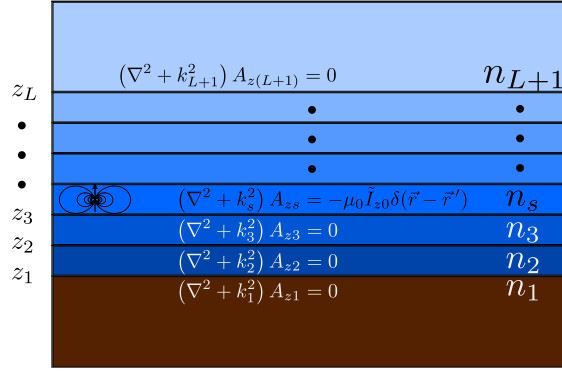


Figure 1: A dipole radiates inside layers of dielectric material. There are L layer interfaces, and $L + 1$ layers including the top and bottom semi-infinite half-spaces.

where J_0 is the Bessel function of the first kind and order zero and k_ρ is the radial wavenumber of a spectral component. The SDGF \tilde{A}_z is split in a piecewise manner into functions in each layer, noted as $\tilde{A}_{z\ell}$ for $\ell = 1, 2, \dots, L + 1$. These functions obey a 1D Helmholtz equation in each layer,

$$\frac{\partial^2 \tilde{A}_{z\ell}}{\partial z^2} + k_{z\ell}^2 \tilde{A}_{z\ell} = \begin{cases} -\mu_0 \tilde{I}_{z0} \frac{\delta(z-z')}{4\pi^2}, & \text{in source layer} \\ 0, & \text{in other layers,} \end{cases} \quad (2)$$

where \tilde{I}_{z0} is the complex magnitude of the current flowing in the dipole, $\mu_0 = 4\pi \times 10^{-7}$ H/m is the permeability of free space, and in each layer the z wavenumber is given by the auxiliary relationship $k_{z\ell} = \sqrt{k_\ell^2 - k_\rho^2}$. Equation (2) is to be solved in all layers simultaneously for a given k_ρ . Then Equation (1) is calculated by numerical quadrature over various values of k_ρ .

3. SPECTRAL DOMAIN GREEN'S FUNCTION SOLVER

The general solution to Equation (2) is

$$\tilde{A}_{z\ell} = j \frac{\mu_0 \tilde{I}_{z0}}{8\pi^2} \times \begin{cases} \frac{e^{jk_{z\ell}|z-z'|}}{k_{z\ell}} + R_\ell^+ e^{jk_{z\ell}(z-z_{\ell-1})} + R_\ell^- e^{-jk_{z\ell}(z-z_\ell)}, & \text{in source layer} \\ R_\ell^+ e^{jk_{z\ell}(z-z_{\ell-1})} + R_\ell^- e^{-jk_{z\ell}(z-z_\ell)}, & \text{in other layers,} \end{cases} \quad (3)$$

where R_ℓ^\pm are generalized reflection coefficients in each layer that have to be fixed by boundary conditions. The boundary conditions are that the tangential EM fields have to be continuous at the layer interfaces, and that there are no incoming waves ($R_1^+ = R_{L+1}^- = 0$, the Sommerfeld radiation condition). Applying these to Equation (3) reduces the problem to a linear system of equations for the R_ℓ^\pm coefficients, the exact details of which are to follow in a future publication. The system of equations is sparse: each equation involves only four of the $2L$ unknowns. The sparse system can be assembled into a $2L \times 2L$ pentadiagonal matrix with at most $(10L - 6)$ nonzero entries out of the full $4L^2$. The memory requirements for large numbers of layers in the structure are not prohibitive. Optimized algorithms exist for solving pentadiagonal systems; this leads to $O(L)$ time-complexity for solving for R_ℓ^\pm . At the time of writing, a typical computer can solve systems of $L \sim 100,000$ layers in ~ 1 second. Once the matrix equation is solved for the R_ℓ^\pm , the SDGF can be evaluated at arbitrary heights z using Equation (3). The matrix is formed and solved, and the SDGF is evaluated at multiple heights each time the quadrature routines of the next section require samples from the SDGF, which happens during both pole-extraction and asymptotic quadrature.

4. QUADRATURE OF THE SPECTRAL DOMAIN GREEN'S FUNCTIONS

SIs are considered to be difficult to numerically integrate due to their slow decay and oscillations [2, 12, 13, 15, 24]. The slow decay is from poles on or near the real k_ρ -axis in the SDGF, which degrade the ability of numerical integration algorithms to converge. The oscillations are from the the Bessel function $J_0(k_\rho \rho)$, which oscillates more rapidly as ρ increases. Conventional quadratures

have to calculate *more* points from the integrand/SDGF as the range increases to resolve the details of the Bessel function. Far-field calculations are therefore time-consuming and intensive. In this work, the slow decay and rapid oscillation problems are handled using two separate techniques, one well established, and the other novel.

4.1. Handling Slow-decay with Pole Extraction

Poles and residues are extracted from the SDGF and analytically integrated using the residue theorem. An effective search algorithm based on complex contour integrals, which automatically finds the pole wavenumbers and residues, has been described in the literature [9, 13]. In the present work, the poles are located by refining user-provided initial guesses with a direct search optimization and *then* contour integrals are used to calculate residues. These are approximated by an adaptive Gauss-Kronrod quadrature adapted from [19, 23]. The residues at all desired heights are calculated in parallel by the chosen Gauss-Kronrod algorithm. The poles are known to come in opposing pairs, $\pm k_p$, so the residue theorem is applied to SIs of the form $\int_0^\infty \frac{1}{k_\rho^2 - k_p^2} J_0(k_\rho \rho) k_\rho dk_\rho = \frac{j\pi}{2} H_0^{(1)}(k_p \rho)$, where the pole is at $k_\rho = k_p$, and $H_0^{(1)}$ is the Hankel function of the first kind and order zero.

4.2. Handling Oscillatory Integrands with Filon-like Asymptotic Quadratures

The standard approach to quadrature of oscillatory integrands is to sample them on enough points to resolve the oscillations and then approximate the integral as a weighted sum of the sampled values. For the present case this would mean that the number of SDGF samples would increase in proportion to the range ρ . Calculation of far-fields would require *more* computational effort than the near-fields. However, there are several families of *asymptotic* quadratures that *increase* in accuracy as the oscillation increases [10, 11]. The method used presently is the Filon-Clenshaw-Curtis (FCC) rule [5, 6]. Although designed for complex exponential oscillation, far-field SIs can be adapted for the FCC in the following way. Since we're interested in asymptotic results, the Bessel function is replaced by its asymptotic expansion in Equation (1), giving

$$A_z(\rho, z, z') \approx \sqrt{\frac{1}{2\pi\rho}} \int_0^\infty \tilde{A}_z(k_\rho, z, z') \left(e^{jk_\rho\rho - j\frac{\pi}{4}} + e^{-jk_\rho\rho + j\frac{\pi}{4}} \right) \sqrt{k_\rho} dk_\rho, \quad (4)$$

which can be broken up into two integrals, each of which has a form that is directly amenable to FCC quadrature after the semi-infinite interval is truncated to a large finite value. Finite truncation is justified physically because the pole-extracted SDGF decays exponentially beyond the largest material wavenumber and the rapid spatial oscillations of large wavenumbers are known to contribute only to the near-field singularity; the corresponding waves do not propagate into the far-field.

5. NUMERICAL EXPERIMENTS

The method presented is general, and could be applied to any multilayered problem. As a proof-of-concept, consider a PCB-type example in which a copper substrate has a graded-index material coated on it, with a linear gradation of 10% over ten wavelengths. A plot of the extracted guided mode potentials as a function of height and range appears in the top panel of Figure 2. There are seven poles corresponding to guided propagation in the structure. It is clear from the plots that the structure acts as a waveguide, trapping the energy in the refractive gradient. In this example, the source dipole is 3.1123 wavelengths above the substrate. The lower panel of Figure 2 is a plot of ray paths that propagate in the same gradient over a reflective ground plane. The two plots compare favorably in terms of the distance scale of the modal “skip zones.”

Another numerical example appears in Figure 3, with the top panel representing a dipole radiating above earth with $\epsilon_r = 15$ and a conductivity of $\sigma = 12 \times 10^{-3}$ S/m in a strong gradient, with $n(z) = 1 + e^{-z}/10$ above ground. While this is an unusually strong gradient for atmospheric cases, it is an exaggerated example to illustrate the effects of propagation through refractive gradients at ranges that can be easily visualized. For more realistic gradients, the effects are apparent farther away from the dipole. Because the FCC quadrature becomes more accurate farther away, and there is no additional computational cost of increasing range, there is no inherent problem with calculating extremely far fields, only one of visualizing the long length scales in the kinds of plots presented. Log-scaled line plots can be appropriate, but we find the visualization of the figures presented more illustrative. The visualized potential is calculated by the sum of guided modes and the asymptotic FCC quadrature result for $\rho > 20\lambda_0$ and by a direct Clenshaw-Curtis quadrature for $\rho \leq 20\lambda_0$. The

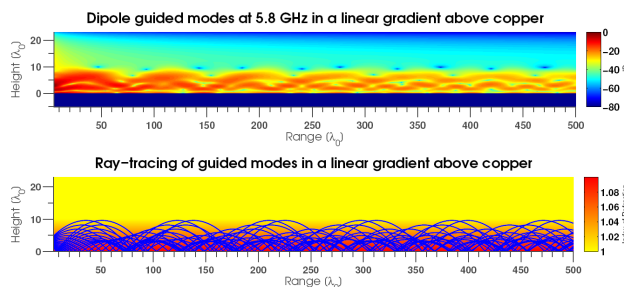


Figure 2: A dipole radiates in a layered coating on a copper substrate at 5.8 GHz. The potential $A_z(\rho, z)$ is visualized on a decibel scale.

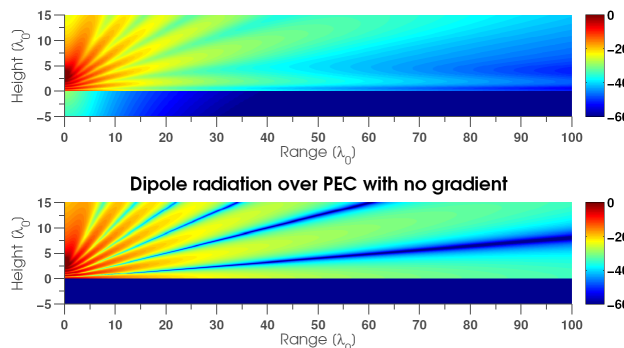


Figure 3: A dipole radiates in two similar scenarios. The top panel is using the SDGF/SI method in an actual scenario and the bottom panel is using image theory in an idealized scenario.

fact that there is no discernible discontinuity in the fields at $\rho = 20\lambda_0$ is evidence of the accuracy of the far-field approximation and FCC quadrature. The lower panel of Figure 3 is a comparison with a simplified environmental model that can be solved using image-theory, a technique that is widely used by communications engineers as the “two-ray” model. The differences between the two are apparent at the surface level, where the two-ray model can be said to break-down.

6. CONCLUSION

A direct method for calculating EM fields in multilayered media based on SDGFs and quadrature of SIs has been presented. Application of asymptotic quadratures to SI problems is novel to our knowledge. Although the technique is conceived of as a method for simulating long range radio propagation near ground level, it is applicable to general EM propagation in multilayered media. Initial testing indicates agreement with the physics; indeed, the calculated fields do solve Maxwell’s equations and boundary conditions. Future directions include careful validation against other models and measurements, exploring enhancements to speed convergence, formulating the solution for horizontal dipoles, and finally handling terrain through boundary-integral/scattering methods.

ACKNOWLEDGMENT

We’d like to thank Dr. Víctor Domínguez for his work on the FCC quadrature and willingness to answer questions via email. This work was funded by the Office of Naval Research.

REFERENCES

1. Akbarpour, R. and A. R. Webster, “Ray-tracing and parabolic equation methods in the modeling of a tropospheric microwave link,” *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 11, 3785–3791, 2005.
2. Aksun, M. and G. Dural, “Clarification of issues on the closed-form Green’s functions in stratified media,” *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 11, 3644–3653, 2005.
3. Dockery, G., R. Awadallah, D. Freund, J. Gehman, and M. Newkirk, “An overview of recent advances for the TEMPER radar propagation model,” *2007 IEEE Radar Conference*, 896–905, Apr. 2007.

4. Dockery, G. D., “Modeling electromagnetic wave propagation in the troposphere using the parabolic equation,” *IEEE Transactions on Antennas and Propagation*, Vol. 36, No. 10, 1464–1470, Oct. 1988.
5. Domínguez, V., I. Graham, and T. Kim, “Filon-Clenshaw-Curtis rules for highly-oscillatory integrals with algebraic singularities and stationary points,” *SIAM Journal of Numerical Analysis*, 2013, to Appear.
6. Domínguez, V., I. G. Graham, and V. P. Smyshlyaev, “Stability and error estimates for Filon-Clenshaw-Curtis rules for highly oscillatory integrals,” *IMA Journal of Numerical Analysis*, Vol. 31, No. 4, 1253–1280, 2011.
7. Donohue, D. and J. Kuttler, “Propagation modeling over terrain using the parabolic wave equation,” *IEEE Transactions on Antennas and Propagation*, Vol. 48, No. 2, 260–277, Feb. 2000.
8. Fang, D. G., *Antenna Theory and Microstrip Antennas*, CRC Press, 2010.
9. Hu, B. and W. C. Chew, “Fast inhomogeneous plane wave algorithm for electromagnetic solutions in layered medium structures: Two-dimensional case,” *Radio Science*, Vol. 35, No. 1, 31–43, 2000.
10. Iserles, A. and S. Nøsett, “On quadrature methods for highly oscillatory integrals and their implementation,” *BIT Numerical Mathematics*, Vol. 44, No. 4, 755–772, 2004.
11. Iserles, A., S. Nøsett, and S. Olver, “Highly oscillatory quadrature: The story so far,” *Numerical Mathematics and Advanced Applications*, A. de Castro, D. Gómez, P. Quintela, and P. Salgado, Eds., 97–118, Springer Berlin Heidelberg, 2006.
12. Kaifas, T., “Direct rational function fitting method for accurate evaluation of sommerfeld integrals in stratified media,” *IEEE Transactions on Antennas and Propagation*, Vol. 60, No. 1, 282–291, 2012.
13. Ling, F. and J.-M. Jin, “Discrete complex image method for Green’s functions of general multilayer media,” *IEEE Microwave and Guided Wave Letters*, Vol. 10, No. 10, 400–402, 2000.
14. Luebbers, R., J. Schuster, and K. Wu, “Full wave propagation model based on moving window FDTD,” *2003 IEEE Military Communications Conference, MILCOM’03*, Vol. 2, 1397–1401, 2003.
15. Michalski, K., “Extrapolation methods for Sommerfeld integral tails,” *IEEE Transactions on Antennas and Propagation*, Vol. 46, No. 10, 1405–1418, 1998.
16. Ohs, R. R., J. W. Schuster, and T. Y. Fung, “Full wave simulation of radiowave propagation in unattended ground sensor networks,” *Proc. IEEE Radio and Wireless Symp. RWS’09*, 183–186, 2009.
17. Ozgun, O., “Recursive two-way parabolic equation approach for modeling terrain effects in tropospheric propagation,” *IEEE Transactions on Antennas and Propagation*, Vol. 57, No. 9, 2706–2714, Sep. 2009.
18. Quinto, M., S. Boutami, and J. Hazart, “Fast computation of dipole radiation in stratified background using graphics processing unit,” *Progress In Electromagnetics Research M*, Vol. 20, 115–126, 2011.
19. Shampine, L. F., “Vectorized adaptive quadrature in MATLAB,” *J. Comput. Appl. Math.*, Vol. 211, No. 2, 131–140, Jan. 2008.
20. Valtr, P. and P. Pechac, “Analytic tropospheric ray-tracing model for constant refractivity gradient profiles,” *First European Conference on Antennas and Propagation, EuCAP 2006*, 1–4, 2006.
21. Wu, K., J. Schuster, and R. Luebbers, “Full wave modeling of RF propagation between low-to-the-ground antennas,” *2005 IEEE Antennas and Propagation Society International Symposium*, Vol. 2B, 711–714, 2005.
22. Wu, K., J. Schuster, R. Ohs, and R. Luebbers, “Application of moving window FDTD to modeling the effects of atmospheric variations and foliage on radio wave propagation over terrain,” *2004 IEEE Military Communications Conference, MILCOM 2004*, Vol. 3, 1515–1521, 2004.
23. Wyatt, A., QUADVGK: Vectorised adaptive G7,K15 Gaussian quadrature on vector of integrals, Feb. 2008.
24. Yuan, M., T. Sarkar, and M. Salazar-Palma, “A direct discrete complex image method from the closed-form Green’s functions in multilayered media,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, No. 3, 1025–1032, 2006.

A Variational Approach to Compute Singular Axisymmetric Electromagnetic Fields

F. Assous¹ and I. Raichik²

¹Ariel University, Israel

²Bar-Ilan University, Israel

Abstract— We propose a new variational approach to solve the axisymmetric Maxwell equations in singular domains, as for example in non convex polygonal domains. We focus on the computation of the electric field $E = (E_r, E_z)$. We show that the key point is to solve axisymmetric Laplace-like operators in the singular domain. This can not be performed by a standard finite element method, which would give a solution identically equal to zero. To get the true non-vanishing solution, we decompose the computational domain into several subdomains, then we derive an *ad hoc* variational formulation, in which the interface conditions are imposed through a method deduced from a Nitsche approach. Numerical examples are shown.

1. INTRODUCTION

Many structures that are to be modeled present a surface with reentrant edges or corners, called geometrical singularities, since they can generate very strong fields. Various approaches have been suggested for solving Maxwell equations in a singular domain. Among others, theoretical works of Birman and Solomyak [1], or characterizations of the singularity of the electromagnetic fields (see [2] and references therein), called electromagnetic singularities. Let us also mention comparisons of different existing approaches for solving the 2D Maxwell equations in [3].

In this paper, we are interested in the two-dimensional axisymmetric singular geometry case (r, z) . Since geometric singularities of the domain (like reentrant corners) have basically an influence on the space part of the equations, we will restrict ourselves to the static problem for the electric field. Then, it will be shown that the key point to compute the singular electric field is related to solve a Laplace operator in a singular domain. The core of this paper will be devoted to derive an *ad hoc* numerical method to solve this Laplace problem. This method will be based on a domain decomposition method, so that it may be supported by the huge bibliography of the well-known domain decomposition methods. It will require the introduction of extended Nitsche method (see [4, 5]) with an “exchange” approach in order to handle the transmission conditions.

2. MATHEMATICAL MODEL

2.1. Maxwell Equations in Axisymmetric Geometry

Let Ω be a bounded axisymmetric domain, limited by the surface of revolution Γ . We denote by ω and γ_b their intersections with a meridian half-plane (see Fig. 1). One has $\partial\omega := \gamma = \gamma_a \cup \gamma_b$, where either $\gamma_a = \emptyset$ when γ_b is a closed contour (i.e., Ω does not contain the axis), or γ_a is the segment of the axis lying between the extremities of γ_b . We denote ν is outward normal, and by τ the unit tangential vector such that (τ, ν) is direct. The natural coordinates for this domain are the cylindrical coordinates (r, θ, z) . A meridian half-plane is defined by the equation $\theta = \text{constant}$,

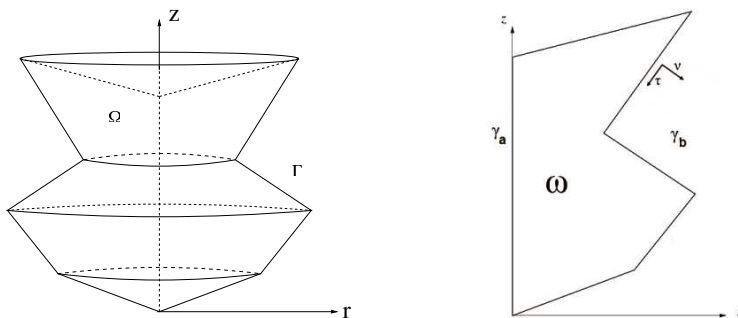


Figure 1: The axisymmetric domain Ω and the corresponding meridian domain ω .

and (r, z) are Cartesian coordinates in this half-plane. Assuming symmetry of revolution, namely $\partial/\partial\theta = 0$, means that the fields are entirely determined by their trace in ω , that is the datum of their value in a meridian half-plane. With this geometry, it is possible to decouple the Maxwell system into two systems of equations. In this paper, we consider only the electric field problem. Let us denote by $\mathbf{E} = (E_r, E_z)$, $\mathbf{J} = (J_r, J_z)$ the meridian component of the electric field and of the current density. Also denote by B_θ the azimuthal component of the magnetic induction, the axisymmetric Maxwell equations for the vector electric field can be written with a system in (\mathbf{E}, B_θ)

$$\begin{cases} \frac{\partial \mathbf{E}}{\partial t} - c^2 \mathbf{curl} B_\theta = -\frac{1}{\varepsilon_0} \mathbf{J}, \\ \frac{\partial B_\theta}{\partial t} + \mathbf{curl} \mathbf{E} = 0, \\ \mathbf{div} \mathbf{E} = \frac{\rho}{\varepsilon_0}. \end{cases} \quad (1)$$

where the perfect conductor boundary condition is expressed $\mathbf{E} \cdot \boldsymbol{\tau} = 0$ on γ_b , and the symmetry conditions on the axis γ_a becomes $\mathbf{E} \cdot \boldsymbol{\nu} = E_r = 0$, $B_\theta = 0$ on γ_a . The operators \mathbf{div} , \mathbf{curl} and \mathbf{curl} above are obtained simply by applying $\partial/\partial\theta = 0$ in the classical formulae for the divergence, and \mathbf{curl} operators in cylindrical coordinates (see for instance [6]).

Let us now introduce some Sobolev functional spaces and their respective norms, which will be useful for what follows. Note that, as we consider an axisymmetric domain ω , the weight r plays a important role in these definitions. We define

$$\begin{aligned} L_r^2(\omega) &= \left\{ f : \omega \rightarrow \mathbb{R}; \int_\omega f^2 r dr dz < +\infty \right\}, & H_r^1(\omega) &= \left\{ f \in L_r^2(\omega); \int_\omega |\mathbf{grad} f|^2 r dr dz < +\infty \right\}, \\ H_m^1(\omega) &= \left\{ f \in H_r^1(\omega); \int_\omega \frac{f^2}{r} dr dz < +\infty \right\}. \end{aligned}$$

This allows to define the usual vector Sobolev spaces

$$\mathbf{L}_r^2(\omega) = L_r^2(\omega) \times L_r^2(\omega), \quad \mathbf{H}^1(\omega) = \{ \mathbf{v} = (v_r, v_z) \in \mathbf{L}_r^2(\omega); v_r \in H_m^1(\omega) \text{ and } v_z \in H_r^1(\omega) \}$$

and the classical spaces for the Maxwell equations

$$\mathbf{H}(\mathbf{curl}, \omega) = \{ \mathbf{v} \in \mathbf{L}_r^2(\omega); \mathbf{curl} \mathbf{v} \in L_r^2(\omega) \}, \quad \mathbf{H}(\mathbf{div}, \omega) = \{ \mathbf{v} \in \mathbf{L}_r^2(\omega); \mathbf{div} \mathbf{v} \in L_r^2(\omega) \}.$$

We also define the corresponding subspace with the vanishing tangential trace $\mathbf{H}_0(\mathbf{curl}, \omega)$ and finally the space of solutions

$$\mathbf{X} := \mathbf{H}_0(\mathbf{curl}, \omega) \cap \mathbf{H}(\mathbf{div}, \omega).$$

Considering the azimuthal component B_θ , one can easily show that it is solution to a scalar wave equation. Indeed, it is sufficient to take the time derivative of the second equation of (1), to apply the \mathbf{curl} operator to the first one, and to use the identity $\mathbf{curl} \mathbf{curl} = -\Delta$. In these conditions, even in a singular domain, B_θ is always regular, and there is no difficulty to compute it.

For this reason, we focus on the computation of electric field \mathbf{E} . Moreover, as we handle the geometrical singularity of the domain, we are basically concerned with the space-dependent part of the model (not the time-dependent one). Thus, we simplify the problem by considering the stationary problem associated with Equation (1), by performing $\frac{\partial}{\partial t} = 0$. The obtained equation appears as a particular case of the following problem, that we will examine next:

For given scalar functions f and g , find $\mathbf{u} = (u_r, u_z)$ solution to

$$\begin{cases} \mathbf{curl} \mathbf{u} = f & \text{in } \omega, \\ \mathbf{div} \mathbf{u} = g & \text{in } \omega, \\ \mathbf{u} \cdot \boldsymbol{\tau} = 0 & \text{on } \gamma_b, \\ \mathbf{u} \cdot \boldsymbol{\nu} = 0 & \text{on } \gamma_a. \end{cases} \quad (2)$$

2.2. A Decomposition in Regular and Singular Parts

As proved in [6], problem (2) is singular in the sense that for a non-convex domain ω — typically ω containing a reentrant corner or a conical point with a sufficient large vertex angle — the space

of solutions \mathbf{X} is *not a subspace* of $\mathbf{H}^1(\omega)$. Nevertheless, \mathbf{X} can be decomposed into two subspaces, where \oplus denotes a direct sum

$$\mathbf{X} = \mathbf{X}_R \oplus \text{grad } \Psi_S,$$

where $\mathbf{X}_R = \mathbf{X} \cap \mathbf{H}^1(\omega)$ is a regular subspace of \mathbf{X} . Above, Ψ_S denotes the space of the primal singularities of the Laplacian, i.e., solutions of a Laplace problem that belong to $H_r^1(\omega)$ but lack an H^2 -style regularity. More precisely, $\psi_S \in \Psi_S$ is characterized by

$$-\Delta \psi_S = P_S \quad \text{in } \omega, \quad (3)$$

$$\psi_S = 0 \quad \text{on } \gamma_b, \quad (4)$$

$$\frac{\partial \psi_S}{\partial \boldsymbol{\nu}} = 0 \quad \text{on } \gamma_a. \quad (5)$$

Here, the right-hand side P_S is singular in the sense that it belongs to $L_r^2(\omega)$, but not to $H_r^1(\omega)$, and satisfies the same equation as ψ_S but with a vanishing right-hand side, namely

$$\Delta P_S = 0 \quad \text{in } \omega, \quad (6)$$

$$P_S = 0 \quad \text{on } \gamma_b, \quad (7)$$

$$\frac{\partial P_S}{\partial \boldsymbol{\nu}} = 0 \quad \text{on } \gamma_a. \quad (8)$$

In an another way, \mathbf{X}_R being a *regular* subspace, as a subspace of $\mathbf{H}^1(\omega)$, one can compute a numerical approximation by a standard method, for instance a P_1 -conforming finite element method. The difficulty comes from the *singular* subspace $\text{grad } \Psi_S$, that has been proved to be a finite-dimensional subspace (see [6]), the dimension being the number of reentrant corners of the domain ω plus the number of conical points for a sufficient large vertex angle. The key point is so to compute P_S , which can not be solved by a standard finite element method, which would give $P_S = 0$. Our aim is to propose a numerical method to compute P_S from which ψ_S and \mathbf{E} will be computed.

3. THE NUMERICAL DOMAIN DECOMPOSITION METHOD

For the sake of simplicity, we will consider a non convex domain ω with only one reentrant corner. The principle of a domain decomposition method is to split the computational domain into several subdomains. Here, we split the domain ω into two subdomains ω_1 and ω_2 . The subdomain ω_1 is taken as the “external” domain and the subdomain ω_2 is the vicinity of the reentrant corner, with the interface $\gamma = \omega_2 \cap \omega_1$. For simplicity, we choose ω_2 to be an open angular sector in the neighborhood of the reentrant corner. This leads use to us local polar coordinates centered at the reentrant corner. In these conditions, ω_1 is the “regular” subdomain, since it does not contain the singularity, whereas ω_2 is the “singular” subdomain that contains the singularity.

Our aim is now to solve the problem in each subdomain separately, taking into account the continuity of the solution and of its normal derivative at the interface γ

$$P_S|_{\omega_1} = P_S|_{\omega_2} \quad \text{and} \quad \frac{\partial P_S}{\partial \nu_1}|_{\omega_1} + \frac{\partial P_S}{\partial \nu_2}|_{\omega_2} = 0 \quad \text{across the interface } \gamma,$$

where ν_1 and ν_2 are respectively the outgoing normals of ω_1 and ω_2 at the interface γ . To simplify, let us denote by P_1 (resp. P_2) the “external” (resp: “internal”) solution, that is the restriction of P_S to ω_1 (resp: ω_2). In the regular subdomain ω_1 , P_1 is regular everywhere and can be solved by standard finite element method. However, the problem set in ω_2 remains singular, since ω_2 still contains the reentrant corner, and a standard numerical method would give $P_2 = 0$. To overcome this difficulty, one uses (cf. [6]) that P_2 can be decomposed into

$$P_2 = Q_2 + P_2^S, \quad (9)$$

where Q_2 is a regular part everywhere in ω_2 of P_2 , that is Q_2 belongs to $H_r^1(\omega_2)$. P_2^S denotes the singular part of P_2 , that is the part that belongs to $L_r^2(\omega_2)$ but not to $H_r^1(\omega_2)$. We have an analytic local expression of the singular part P_2^S (cf. [6]), given by

$$P_2^S = \rho^{-\alpha} \sin(\alpha\phi), \quad (10)$$

where (ρ, ϕ) denote the local polar coordinates centered at the reentrant corner. Consequently,

$$\frac{\partial P_2}{\partial \nu_2} = \frac{\partial Q_2}{\partial \nu_2} + \frac{\partial}{\partial \nu_2} (\rho^{-\alpha} \sin(\alpha\phi)).$$

Moreover, we get after some algebra that $\Delta(\rho^{-\alpha} \sin(\alpha\phi)) = \frac{-\alpha}{r} \rho^{-\alpha-1} \sin((\alpha+1)\phi)$, and using that $\frac{\partial}{\partial \nu_2} = \frac{\partial}{\partial \rho}$, the system (6)–(8) can be decomposed in the following two systems

$$\begin{array}{ll} \text{Find } P_1 \in H_m^1(\omega_1) \text{ solution to} & \text{Find } Q_2 \in H_m^1(\omega_2) \text{ solution to} \\ \left\{ \begin{array}{ll} \Delta P_1 = 0 & \text{in } \omega_1, \\ P_1 = 0 & \text{on } \partial\omega_1 \cap \gamma_b, \\ \frac{\partial P_1}{\partial \nu_1} = 0 & \text{on } \partial\omega_1 \cap \gamma_a, \\ P_1 = P_2 & \text{on } \gamma, \\ \frac{\partial P_1}{\partial \nu_1} = -\frac{\partial P_2}{\partial \nu_2} & \text{on } \gamma. \end{array} \right. & \left\{ \begin{array}{ll} -\Delta Q_2 = \frac{-\alpha}{r} \rho^{-\alpha-1} \sin((\alpha+1)\phi) & \text{in } \omega_2, \\ Q_2 = 0 & \text{on } \partial\omega_2, \\ Q_2 = P_1 - \rho^{-\alpha} \sin(\alpha\phi) & \text{on } \gamma, \\ \frac{\partial Q_2}{\partial \nu_2} = -\frac{\partial P_1}{\partial \nu_1} + \alpha \rho^{-\alpha-1} \sin(\alpha\phi) & \text{on } \gamma. \end{array} \right. \end{array}$$

We have now to derive a variational formulation, that will be the basis of the finite element method. The first ingredient comes from a paper of J. Nitsche [5], where the author proposed a variational approach to enforce weakly Dirichlet boundary condition. We apply it here to impose the continuity of the solution P_S , that is $P_1 = P_2$ on γ . Note that in this case, P_2 is considered as a data when solving P_1 in ω_1 , and reciprocally. We have also to impose the continuity of the normal derivatives across γ . For that, we introduce an *exchange* method that consists in exchanging the term $\partial P_1 / \partial \nu_1$ with $-\partial P_2 / \partial \nu_2$ in the variational formulation on ω_1 , and reciprocally for a coupled variational formulation on ω_2 . We so obtain the variational formulation with (P_1, Q_2) as unknowns, that reads *Find* $P_1 \in H_m^1(\omega_1)$ *such that*

$$\begin{aligned} & \iint_{\omega_1} \nabla P_1 \nabla \varphi_1 r d\omega + \int_{\gamma} \frac{\partial Q_2}{\partial \nu_2} \varphi_1 r d\gamma - \int_{\gamma} \frac{\partial \varphi_1}{\partial \nu_1} P_1 r d\gamma + \beta \sum_{E \in \Sigma_h} \frac{1}{h_E} \int_{\gamma} P_1 \varphi_1 r dE \\ &= \int_{\gamma} \alpha \rho^{-\alpha-1} \sin(\alpha\phi) \varphi_1 r d\gamma - \int_{\gamma} \frac{\partial \varphi_1}{\partial \nu_1} Q_2 r d\gamma - \int_{\gamma} \frac{\partial \varphi_1}{\partial \nu_1} \rho^{-\alpha} \sin(\alpha\phi) r d\gamma \\ &+ \beta \sum_{E \in \Sigma_h} \frac{1}{h_E} \int_{\gamma} Q_2 \varphi_1 r dE + \beta \sum_{E \in \Sigma_h} \frac{1}{h_E} \int_{\gamma} \rho^{-\alpha} \sin(\alpha\phi) \varphi_1 r dE, \quad \forall \varphi_1 \in H_{m0}^1(\omega_1) \end{aligned}$$

Find $Q_2 \in H_m^1(\omega_2)$ *such that*

$$\begin{aligned} & \iint_{\omega_2} \nabla Q_2 \nabla \varphi_2 r d\omega + \int_{\gamma} \frac{\partial P_1}{\partial \nu_1} \varphi_2 r d\gamma - \int_{\gamma} \frac{\partial \varphi_2}{\partial \nu_2} Q_2 r d\gamma + \beta \sum_{E \in \Sigma_h} \frac{1}{h_E} \int_{\gamma} Q_2 \varphi_2 r dE \\ &= \int_{\gamma} \alpha \rho^{-\alpha-1} \sin(\alpha\phi) \varphi_2 r d\gamma - \int_{\gamma} \frac{\partial \varphi_2}{\partial \nu_2} P_1 r d\gamma - \int_{\gamma} \frac{\partial \varphi_2}{\partial \nu_2} \rho^{-\alpha} \sin(\alpha\phi) r d\gamma \\ &+ \beta \sum_{E \in \Sigma_h} \frac{1}{h_E} \int_{\gamma} P_1 \varphi_2 r dE - \beta \sum_{E \in \Sigma_h} \frac{1}{h_E} \int_{\gamma} \rho^{-\alpha} \sin(\alpha\phi) \varphi_2 r dE \\ &- \iint_{\omega_2} \frac{\alpha}{r} \rho^{-\alpha-1} \sin((\alpha+1)\phi) \varphi_2 r d\omega, \quad \forall \varphi_2 \in H_{m0}^1(\omega_2) \end{aligned}$$

Assuming that P_S has been obtained, one can similarly compute ψ_S of Ψ_S . The last step consists in computing a basis of the singular electric subspace $\text{grad } \Psi_S$ which is simply obtained by taking the gradient of the function ψ_S . A P_1 -conforming finite element method has been developed, based on the FreeFem++ package [7], to solve this problem. Numerical illustrations are given in the following section.

4. NUMERICAL RESULTS

We consider the 3-D top hat domain Ω with a reentrant circular edge, that corresponds to an L-shaped 2-D domain ω with reentrant corner. Our aim is to compute the singular electric field $\text{grad } \psi_S$ basis of $\text{grad } \Psi_S$. We introduce an unstructured mesh of the L-shaped domain ω made up of triangles, with no particular mesh refinement near the corner. We first compute the singular function P_S . The numerical result is depicted in Fig. 2(a). When representing functions or fields with a singular behavior, we have chosen to truncate the results in the singular (infinite) node. One can see that the method is able to compute P_S , whereas a usual method would give $P_S = 0$ as solution. Moreover, as expected, the continuity of the solution and of its normal derivative are well handled.



Figure 2: (a) Singular P_S solution and (b) singular electric field (z -component).

Similarly, we compute Ψ_s , from which one gets the electric basis $\text{grad } \psi_S$ of $\text{grad } \Psi_S$ by taking the grad of ψ . The longitudinal component obtained is also depicted in the Fig. 2(b). This still shows that the method captures well the singular electric field near the edge (and far away from it). A conforming P_1 Finite Element Method can not yield such a result.

5. CONCLUSION

We have presented a new variational method to compute the singular electric field in an axisymmetric singular domain. It is based on a decomposition of the computational domain into two subdomains: an internal one, close to the singularity, and an external one. This method uses the local expression of the singularity, and an extended version of the Nitsche method coupled with an “exchange” original approach. This allows us to handle continuity interface conditions, both of the Dirichlet and the Neumann type. Numerical results have been shown and illustrate the efficiency of the method.

REFERENCES

1. Birman, M. S. and M. Z. Solomyak, “L2-theory of the Maxwell operator in arbitrary domains,” *Russian Math. Surveys*, Vol. 42, 75–96, 1987.
2. Costabel, M., “A remark on the regularity of solutions of Maxwells equations on Lipschitz domains,” *Math. Meth. Appl. Sci.*, Vol. 12, 365–368, 1990.
3. Hazard, C., “Numerical simulation of corner singularities: A paradox in Maxwell-like problems,” *C. R. Acad. Sci. Paris, Ser. Iib*, Vol. 330, 57–68, 2002.
4. Becker, R., P. Hansbo, and R. Stenberg, “A finite element method for domain decomposition with non-matching grids,” *M2AN*, Vol. 37, 209–225, 2003.
5. Nitsche, J., “Ber ein variationsprinzip zur lsung von dirichlet-problemen bei verwendung von teilrumen, die keinen randbedingungen unterworfen sind,” *Abh. Math, Sem. Univ. Hamburg*, Vol. 36, 9–15, 1971.
6. Assous, F., P. Ciarlet, Jr., and J. Segré, “Numerical solution to the time-dependent Maxwell equations in two-dimensional singular domain: The singular complement method,” *J. Comput. Phys.*, Vol. 161, 218–249, 2000.
7. Hecht, F., *FreeFem++*, Laboratoire J. L. Lions, Université Pierre et Marie Curie, 2010.

Electromagnetic Sources and Observers in Motion X — Unification of Electromagnetism and Gravity

Selwyn E. Wright

Moor Lane Laboratory, ECASS Technologies Ltd, HD8 0QS, UK

Abstract— This tenth paper in the series of EM Sources and Observers in Motion is complimentary to paper IX dealing with the Nature of Gravity. Motional Electromagnetic (EM) and gravitational theories appear to have two inherent deficiencies that have prevented them from becoming a unified theory. (α) Not recognising that the propagation medium (ether) is the essential thread that runs through these developments. (β) Not realising that EM waves and gravity are two forms of the same field — unsteady electric and steady difference electric fields. A New Relativity theory (NR) re-establishes a preferred frame of reference, re-uniting classical and modern physics. It appears that all observed actions at a distance are based on a propagation medium, and that EM, gravitational and inertial fields use the same electrical medium. The only difference from classical theories, is that time and structure of EM systems shrink with motion through the medium, gravity compresses the medium, and residual gravity within the universe provides the inertial field. All predictions involving observations have to satisfy the wave equation through its propagation medium. All causal observations, which currently support Einstein's Special Relativity (SR) [4] and General Relativity (GR) [7], are predicted using the medium based theory. The ether-less aspects of SR and GR are non causal (not predictable). NR predicts measured SR and GR properties in a straight forward physical manner with a medium, without ambiguity or paradox.

1. INTRODUCTION

Measurements, based on Einstein's SR and GR, such as [16] on binary pulsars, are causal, (predictable). They are not ether-less as Einstein claimed. Einstein used Lorentz's [2] medium based Transform (LT) and field equations based on a propagation medium. Einstein did not appear to realise the equations he used were medium based, not believing in an ether. Einstein's ether-less concept of relativity is not supported by the medium. based LT, where interaction with the EM medium is essential. Ether-less claims and predictions, such as simultaneity, time travel and no absolute time and space, are non causal, they cannot be measured. Einstein not accepting the medium's presence, believed that the general wave equation could not distinguish between source and observer motion, nor distinguish between stationary and moving medium reference frames. The new theory accommodates these additional features, predicts the medium based measured aspects of SR and GR and identifies the ether-less, non causal aspects of relativity.

[18–25] has shown that Einstein's ether-less relativity, based on his inertial frame, is non causal, it cannot predict measured observations. Without a medium relativity cannot explain how the simple Doppler effect occurs, how a moving more dense medium than a vacuum can convect light, how an impulsive wave is formed in [10] radiation, how two systems can move apart physically, greater than the speed of light, but not relative to each other. Nor distinguish between light propagation on Earth, around the Earth or through the Solar System and beyond. Einstein's invariant ether-less inertial frame cannot predict propagation time asymmetry or explain the inconsistency of predicting ether-less and medium based motional properties, both at the same time. To account for these measured observations, the medium's presence has to be accepted, Einstein's concept of relative motion rejected, and the medium based LT extended for both source and observer motion. None of the observations can be observed without a medium.

Those who believe there is nothing wrong with Einstein's relativity ignore fundamental unanswered questions such as; how does light propagate, or how can one solve the wave equation for propagating waves, without a propagation medium? Removing the medium is against basic wave theory, creating a discontinuity between classical and modern physics. There is no evidence, or causal model that can support Einstein's ether-less universe. Further, after a century of SR and GR, there have been no significant developments to explain what gravity is, how it is propagated and how it is related to EM waves. This theory identifies and establishes the existence of a well defined measured propagation medium and unifies it with gravity. There is now sufficient evidence to re-think our basic EM model by reinstating the EM medium for the propagation of EM waves and gravity, and recognising that gravity is the attraction between dissimilar charges. The following aspects support the medium based New Relativity (NR) theory:

1.1. Basic Physics

- 1 NR challenges SR as a more consistent and comprehensive theory. It is derived from Lorentz's medium based motional transform; it restores the continuity between classical and modern physics.
- 2 All observed and measured motional effects, both classical and relativistic, are with respect to the propagation medium. All predictable EM motional and gravitational theories are based on the medium.
- 3 EM, gravitational and inertial fields propagate using the same medium, they are causal being solutions of the wave equation. Whereas, Einstein's ether-less predictions are non causal, they cannot be measured.
- 4 Relativists who claim to have verified Einstein's ether-less relativity have usually verified Lorentz's medium based transform with its time and structural contraction through motion relative to the medium.
- 5 Lorentz's rectangular transform axes predictions are causal; they predict motional properties. Oblique transform axes, which simulate Einstein's ether-less properties are non causal, they cannot predict reality.

1.2. Moving Systems and Media

- 1 Einstein's invariant inertial frame, predicting ether-less properties, is not a solution of the wave equation. All causal theories are variant; they predict Propagation Time Asymmetry (PTA) around moving systems.
- 2 At Earth speeds, gravitational strengths and small integration times, relativistic and gravitational effects cannot contribute significantly to the dominant instantaneous classical PTA.
- 3 Besides predicting the same measured observations as SR, NR also predicts additional measured observations that SR cannot predict, and exposes ether-less aspects of SR that are non causal.
- 4 NR shows that not only does the medium exist over all space, but it is not homogeneous. It is attracted to and compressed, both time and space, around gravitational bodies. according to Schwarzschild [8].
- 5 Einstein's homogeneous ether-less universe is not supported by measurement. NR supports a confirmed Gravitational Entrainment Model (GEM), describing various media motions around the Earth and heavens.

1.3. Motional Properties

- 1 NR allows distinction to be made between moving medium frames, and distinction between moving sources and observers with respect to the propagation medium, not possible without a medium.
- 2 Observer motion contracts its time and structure, relatively expanding the medium's time and space, thus having the ability to relatively reduce gravity's time and space compression of the medium.
- 3 NR shows time travel is non causal, it is not a solution of the wave equation. It is possible to travel to the past, causally, but not to interfere, it has happened. It is not possible to travel to the future it has not occurred.
- 4 NR shows that it is possible to travel faster than the speed of light across hybrid frames, making intergalactic exploration possible. Distance is measured in the medium and time in the moving frame.
- 5 NR shows that LT is fundamental, SR is inconsistent, incomplete and its ether-less interpretation non causal. Attempts to explain observations without a medium lead to non causal predicaments.

1.4. Gravitational Attraction

- 1 NR extends van der Waals (1873) near field electrostatic theory of attraction between molecules to account for the far field attraction between matter, constituting gravity in General Relativity (GR) [7].
- 2 NR shows unsteady electrical fields are electromagnetic, steady dissimilar difference electrical fields are gravitational and residual difference electrical fields within the universe, are inertial.

- 3 Medium provides attractive ‘dark’ energy attracting gravitational mass and possibly repulsive ‘dark’ energy through expansion of medium overcoming the global gravitational attraction.
- 4 All EM wave theories can be accounted for using a propagation medium. Einstein’s predictions that are measurable are medium based, those that cannot be measured are ether-less.
- 5 NR provides a medium link between LT, SR, GR, PTA, inertial and accelerating frames. It also provides an optical description of the Equivalence Principle, and a link to the unification theory of the universe.

2. EINSTEIN’S INERTIAL FRAME

Einstein’s ether-less universe provides no answer on how light and gravity are transmitted across the cosmos and cannot identify any authority allowing the order between cause and effect to be reversed. Attempting to remove the medium appears to be the root cause of confusion. Eliminating the medium is irrational, it violates well established physical principles, preventing a causal solution of the wave equation, and removing the means of propagating waves that produce the observed effect (transmission of energy and information). There is no justification for removing the medium. *Lorentz’s motional theory extended by NR’s wave equation solution, is the legitimate description of wave propagation with respect to the reference medium.* It is shown through extensive measurements to be the case, there is no other causal description.

If a system moves relative to the medium it does not affect the propagation speed in the medium according to Lorentz’s basic wave theory. But it does result in PTA and variance in the moving frame, which is against Einstein’s ether-less, inertial frame predictions. Einstein’s concept of an inertial frame, where there is absolutely no difference between a stationary and constantly moving frame i.e., the physics, propagation speed and propagation time are all considered invariant, cannot provide a solution of the wave equation. *The frame with no medium cannot predict observations, it is non causal, it is purely an internal moving frame description, which cannot account for external observed wave propagation.*

3. MOTIONAL EFFECTS ARE CLASSICAL

Not only are the EM disturbances medium based, but relativistic arguments cannot explain them. Lorentzian contraction and Schwarzschild’s gravitation compression effects, discussed in Section 5, Paper IX, are small at Earth speeds and gravitational strengths. This requires a large summation period for their effects to be appreciable. Whereas, the classical PTA is the fundamental instantaneous disturbance. It is shown that the basic EM motional effects: [1, 2, 9, 12, 14], cited to support the relativistic effect at these low speeds, are in fact explained using only the PTA part of Lorentz’s motional theory. *If this variant PTA effect, for EM systems in motion, had been established in 1905, it could have prevented Einstein’s ether-less, invariant relativity from developing.*

Einstein’s inexplicable concept of relative motion, offering no alternative mechanism to that provided by the medium, and insisting that the observations were transmitted without a medium, are not possible. Thus, Einstein’s belief that his Special Relativity (SR) [4] and General Relativity [7] had no medium (substance) to propagate EM waves and gravity is in error. They are based on a medium, as Lorentz predicted, not on relative motion between systems as Einstein believed. *All EM motional observations can be predicted directly (without Einstein’s SR) by simply using the medium based classical wave equation. Modified additionally, for high speed motion, by the medium based Lorentz time and structural contraction.*

4. EINSTEIN’S ERROR

It is true that the propagation speed and physics are invariant in a constantly moving frame, constituting Einstein’s two postulates. But to observe these effects a third postulate is required: a medium is needed to transmit the observations and accommodate PTA. According to causal wave theory the medium can:

- I. Either move with the system, where there will be no PTA ahead or behind the moving system.
- II. Or systems move relative to the medium, where there will be PTA, there is no other possibility.

But propagation defined through the causal solution of the wave equation, must always be relative to the medium, not relative to the moving frame as Einstein believed. This is Einstein’s basic error; wave propagation must always be relative to the preferred reference frame — the propagation medium. Its propagation speed remains invariant in the moving frame only because both time and

structure contract by exactly the same ratio, through motion with respect to the medium. This makes the propagation speed invariant, but not the PTA laid down in the stationary medium. It is Lorentz's radiation theory, extended by NR's general wave equation, based on a medium, which correctly predicts the variant motional properties for source and observer motion. Einstein's invariant inertial frame responsible for predicting the ether-less properties does not support reality.

Einstein's supportive error was the null effect on light propagation on Earth as it moves through space, according to the Michelson and Morley Experiment (MMX) [1]. Einstein believed that the MMX supported his inertial frame, where there is no difference between a moving and stationary frame. This again is not possible, from (I) above waves without PTA must mean that there is no relative motion with respect to the medium, i.e., the MMX is explained quite naturally with the medium moving with the Earth, as in the confirmed Gravitational Entrainment Model (GEM), described in Section 6.

5. LORENTZ FUNDAMENTAL

It is not generally realised that [2], with help from [3], developed the fundamental EM motional wave theory *based on a medium*. Through the Lorentz Transform (LT) the complete measured motional properties are predicted. This is dominated by the motional distortion of the original disturbance, shown in Figure 2, Paper IX. It is characterised by the classical wave PTA surrounding the system, moving relative to the medium, evident in all wave theories. In EM theories, there is the additional modifying Lorentzian time and structural contraction (LC) at high speed.

Time and space of the medium do not change, it's the time and structure of physical objects (atoms and molecules) that contract (not dilate) passing through the medium. Time slows and structures shrink in the direction of motion by the same fraction making the propagation speed of light invariant. *Spacecraft shrink, both time and structure in the direction of motion, but not the distances they travel.* LC is not instantaneous, at low speeds it needs to be integrated over a considerable time for its effect to be appreciable. Notionally, the LT can be expressed by the incremental equation:

$$\Delta LT = \Delta PTA + \Delta LC \quad (1)$$

The above expression is determined through the causal solution of its wave equation, using a medium with finite properties, the same as any other wave propagation theory. There is no other rational outcome, Lorentz predicts no ether-less properties. He predicts measured events according to the system's instantaneous motional interaction with the medium. The theory is causal, the cause (source event) must always occur before the effect (observed event), which is not the case for the ether-less theory.

6. MEASURED DATA

The acceptance of a new theory is by measurement confirmation. The above theory is vindicated through all known measured data. *Michelson and Morley Experiment (MMX) [1], [5, 6] and Michelson and Gale (M&G) [9], sometimes thought to support the lack of a propagating medium, are all explained through the medium based classical PTA*, as illustrated in the medium based Gravitational Entrainment Model (GEM) in Figure 1. The MMX (no motional effect on light propagation on Earth, relative to the stationary medium moving with the Earth through space), has no PTA because the measuring equipment is stationary relative to the medium. Sagnac (mirror motion in and against the light propagation, relative to the stationary medium on Earth), demonstrates PTA. M&G (medium on Earth clinging to the Earth's surface, increasing its speed towards the equator relative to the surrounding medium) relies on PTA for its effect.

Similar arguments apply to more modern data, [12, 13] and Global Positioning Systems (GPS) [14]. Reasenberg (medium moving with the Solar System at high speed through the universe in Mars-Earth communication), causes effectively no relative motion, negligible PTA. GPS (satellite and Earth station rotating relative to stationary medium surrounding the Earth), creates PTA causing a predicted measured surface positional displacement, where from Equation (9), paper IX, M is the Earth's surface Mach number ($M = 440/300 \cdot 10^6 \approx 1.5 \times 10^{-6}$) and d is now the satellite distance above the Earth ($d \approx 20 \cdot 10^6$ m), giving $\Delta d \approx vt = v/d/c = Md = 1.5 \times 10^{-6} \times 20 \cdot 10^6 = 30$ m, which is what is measured. Saburi et al. (communications across the pacific via satellite and Earth rotating relative to stationary medium surrounding the Earth), causes a predicted measured PTA $\approx 2\Delta t = 2Mt = 300$ ns (100 m).

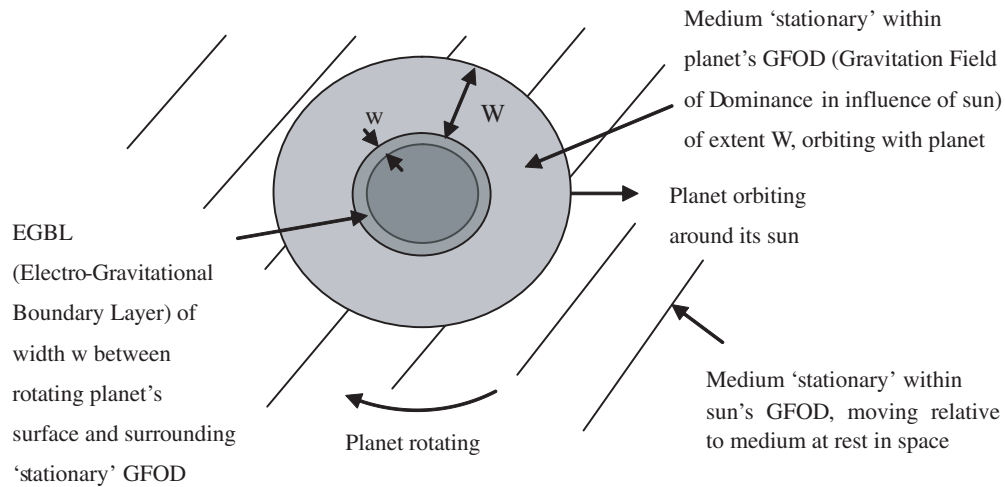


Figure 1: Gravitational Entrainment Model (GEM). Medium profiles around a rotating and orbiting planet, according to [1, 5, 6, 9, 11], Hafele and Keating, [12–15].

All six of these basic experiments are accounted for through the PTA using a propagation medium. Whereas, claims that relativistic effects account for these motional changes are not possible at Earth speeds and short integration times. In the absence of gravitational matter, the medium in the universe appears to be stationary, on average, providing a universal reference for motion. The cosmic microwave background (CMB), detected by [11], is shown to be EM radiation, propagating uniformly in all directions, throughout the universe, relative to the propagation medium basically at rest in space. The stationary medium has also been confirmed through the Cosmic Background Explorer COBE [15]. Here the CMB energy collection increases with system motion relative to the stationary medium, similar to trawling fish nets catch more fish than stationary ones. This model, with the 'stationary' medium surrounding the Earth is supported by the latest results from NASA's Gravity Probe B [17].

[16] on binary pulsar measurements is a modern data set cited as supporting Einstein's SR and GR. These predictions are founded on the medium based aspects of SR and GR, using Lorentz medium based transform, not on the ether-less aspects of SR and GR that Einstein believed. They studied observations of pulses from a distant double pulsar system. In this case, the pulsar's orbital time slowing and orientation change is attributed to the high speed of the pulsar relative to the medium. And the propagation time delay caused through the intense gravity of the heavy pulsars (neutron stars) compressing the medium and retarding the propagation in the vicinity of the pulsars, in complete agreement with NR.

7. CONCLUSION

The propagation medium (ether) exists. Einstein's ether-less beliefs are false. NR solves the wave equation for both source and observer motion relative to the propagation medium. Light and gravity are two aspects of the same electric field and are propagated by the same propagation medium providing the basis for a unification theory of the universe.

REFERENCES

1. Michelson, A. A. and E. W. Morley, "On the relative motion of the Earth and the luminiferous ether," *American Journal of Science*, Vol. 34, No. 203, 333–345, 1887.
2. Lorentz, H. A., "Simplified theory of electrical and optical phenomena in moving systems," *Proc. Acad. Science Amsterdam*, Vol. 1, 427–442, 1899.
3. Poincaré, H., "La théorie de Lorentz et le principe de réaction," *Archives Néerlandaises des Sciences exactes et naturelles*, 253–278, 1900.
4. Einstein, A., "On the electrodynamics of moving bodies," *Annalen der Physik*, Vol. 17, 891–921, 1905.
5. Sagnac, G., "L'éther lumineux démontré par l'effet du vent relatif d'éther dans un interféromètre en rotation uniforme," *Comptes Rendus*, Vol. 157, 708–710, 1913.
6. Sagnac, G., "Sur la preuve de la réalité de l'éther lumineux par l'expérience de l'interféromètre tournant," *Comptes Rendus*, Vol. 157, 1410–1413, 1913.

7. Einstein, A., "Die feldgleichungen der gravitation (The field equations of gravitation)," *Königlich Preussische Akademie der Wissenschaften*, 844–847, 1915.
8. Schwarzschild, K., "Über das gravitationsfeld eines massenpunktes nach der Einstein'schen theorie," *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, Vol. 1, 189–196, 1916.
9. Michelson, A. A. and H. G. Gale, "The effect of the Earth's rotation on the velocity of light," *The Astrophysical Journal*, Vol. 61, 140–145, 1925.
10. Cerenkov, P. A., "Visible emission of clean liquids by action of γ radiation," *Doklady Akad. Nauk SSSR*, Vol. 2, 451, 1934.
11. Penzias, A. A. and R. W. Wilson, "A measurement of the flux density at 4080 Mc/s," *Astrophysical Journal Letters*, Vol. 142, 1149–1154, 1965.
12. Yamamoto, M. and K. Harada, "High precision time comparison via satellite and observed discrepancy of synchronization," *IEEE Transactions on Instrumentation and Measurement*, Vol. 25, No. 4, 473–477, 1976.
13. Reasenberg, R. D., I. I. Shapiro P. E., MacNeil, R. B. Goldstein, J. C. Breiden-Thal, J. P. Brenkle, D. L. Cain, T. M. Kaufman, T. A. Komarek, and A. I. Zygielbaum, "Viking relativity experiment. Verification of signal retardation by solar gravity," *Astrophysical Journal*, Vol. 234, L219–L221, 1979.
14. Logsdon, T., *The NAVSTAR Global Positioning System*, Van Nostrand Reinhold, GPS, 1992.
15. "Cosmic background explore," NASA Goddard Space Flight Center, COBE, 1992.
16. Kramer, M., I. H. Stairs, R. N. Manchester, M. A. Mc Laughlin, A. G. Lyne, R. D. Ferdman, M. Burgay, D. R. Lorimer, A. Possentti, N. D. Amico, J. M. Sarkissian, G. B. Hobbs, J. E. Reynolds, P. C. C. Freire, and F. Camilo, "Tests of general relativity from timing the double pulsar," *Science*, Vol. 314, 97–102, 2006.
17. Gravity Probe B, <http://www.npr.org/2011/05/06/136057344/proof-that-einstein-got-it-right?> and <http://www.sciencedaily.com/releases/2011/05/110504150655.htm>, 2011.
18. Wright, S. E., "Electromagnetic sources and observers in motion I — Evidence supporting the EM propagation medium for the transmission of light," *PIERS Proceedings*, 71–76, Xi'an, China, March 22–26, 2010.
19. Wright, S. E., "Electromagnetic sources and observers in motion II — Einstein's ether-less relativity versus Lorentz's medium based theory," *PIERS Proceedings*, 77–81, Xi'an, China, March 22–26, 2010.
20. Wright, S. E., "Electromagnetic sources and observers in motion III — Derivation and solution of the electromagnetic motional wave equation," *PIERS Proceedings*, 1151–1155, Cambridge, USA, July 5-8, 2010.
21. Wright, S. E., "Electromagnetic sources and observers in motion IV — The nature of gravity and its effect on the propagation medium," *PIERS Proceedings*, 1156–1161, Cambridge, USA, July 5–8, 2010.
22. Wright, S. E., "Electromagnetic sources and observers in motion V — A revised theory of relativity," *PIERS Proceedings*, 367–372, Marrakesh, Morocco, March 20–23, 2011.
23. Wright, S. E., "Electromagnetic sources and observers in motion VI — New motional optics," *PIERS Proceedings*, 362–366, Marrakesh, Morocco, March 20–23, 2011.
24. Wright, S. E., "Electromagnetic sources and observers in motion VII — Medium support for a new relativity theory," *PIERS Proceedings*, 1024–1029, Moscow, Russia, August 19–23, 2012.
25. Wright, S. E., "Electromagnetic sources and observers in motion VIII — New relativity theory establishes Einstein's ether-less aspect of relativity as irrational," *PIERS Proceedings*, 1030–1035, Moscow, Russia, August 19–23, 2012.

Performance Analysis of Blind Calibration in Five-port Receivers Based on Blind Source Separation

F. Vidal¹, A. Duarte², A. Medeiros³, and B. Huyart⁴

¹Department of Environmental and Technological Sciences, UFERSA, Brazil

²Department of Computer Engineering and Automation, UFRN, Brazil

³Department of Electrical Engineering, UFRN, Brazil

⁴Department of Communications and Electronics, Telecom ParisTech, France

Abstract— In this paper, we evaluate the performance of a blind calibration procedure based on blind source separation (BSS) for five-port direct conversion receiver, recovering the in-phase, $s_I(t)$, and quadrature, $s_Q(t)$, components from observed output signals, taking advantage of the fact that no prior knowledge about the signals $s_I(t)$ and $s_Q(t)$ is necessary. An experimental prototype has been fabricated using microstrip components operating at 2.4 GHz frequency band. Simulation and measurement results are presented to validate the proposed approach for the algorithms JADE, FastICA as well as an iterative clustering approach called k-Means algorithm.

1. INTRODUCTION

The diversity of currently available wireless services (GSM, UMTS, WiMax, LTE and IEEE 802.15) requires that multi-mode transceivers are able to support many wireless communication standards that operate in several frequency bands with low cost, low power dissipation and high integration. The conventional superheterodyne receiver makes it difficult to achieve these goals, mainly due to the need for expensive external analog filters to eliminate image response. The direct conversion receiver, also called zero-IF or homodyne receiver, is one of the candidates to meet these needs.

In order to solve these problems and satisfy the requirements for small size, low cost and wide-band, designers have been directing their efforts to design other receivers architecture. A new direct conversion receivers architecture based on the six-port reflectometer have been proposed as multimode and multiband or software receiver operating with digital signal processors (DSPs) [1–3]. Advantages such as phase and amplitude demodulations, compatibility with DSP and broadband specification, have made of five-port architecture a great choice for software defined radio application. However, the use of the five-port as a communication receiver requires an I/Q calibration (or regeneration) procedure in order to generate the in-phase (I) and quadrature (Q) components of the transmitted baseband signal.

2. FIVE-PORT RECEIVERS

The direct conversion receiver based on the five-port system is shown in Fig. 1. The receiver consists of an RF linear circuit with two inputs, three outputs, three RF power detectors, three analog-to-digital converters (A/C) and a digital signal processor (DSP). This system generates a signal $s(n)$ in the digital domain representing the baseband signal.

A block diagram of the internal structure of the RF five-port linear circuit is presented in Fig. 2, with the ϕ_i phase-shifted by 120 degrees.

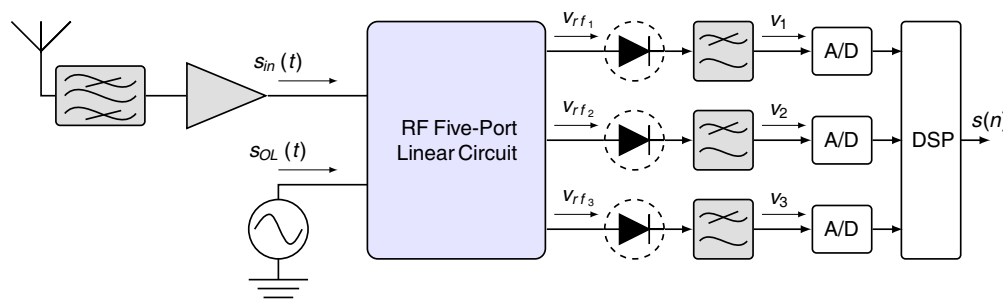


Figure 1: Homodyne five-port receiver.

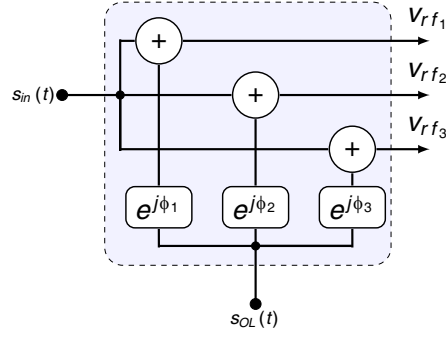


Figure 2: Linear five-port circuit.

The linear circuit performs three vectorial additions between the two input signals $s_{in}(t) = \Re[s(t)e^{j2\pi f_c t}]$ and $s_{OL}(t) = \Re[A_{OL}e^{j2\pi f_c t}]$, where A_{OL} is the local oscillator's amplitude, f_c is the carrier frequency and $s(t) = a(t)e^{j\theta(t)}$ is the complex envelope of the $s_{in}(t)$, $a(t)$ and $\theta(t)$ are respectively the time dependent envelope and phase of $s_{in}(t)$. The three out coming RF signals (v_{rf_i} for $i = 1, 2, 3$) resulting from the three vectorial additions of the signals $s_{in}(t)$ and $s_{OL}(t)$ can be expressed as [4]:

$$v_{rf_i} = \Re \left[A_{OL}e^{j(2\pi f_c t + \phi_i)} + s(t)e^{j2\pi f_c t} \right] = A_{OL} \cos(2\pi f_c t + \phi_i) + a(t) \cos[2\pi f_c t + \theta(t)] \quad (1)$$

The output signal $v_i(t)$ in Fig. 1, after the diode detection and ideal low-pass filtering, can be expressed as

$$v_i(t) = [v_{rf_i}(t)]^2 * h_{lp}(t) \quad (2)$$

Substituting Eq. (1) into Eq. (2), we obtain

$$v_i(t) = K_{lp_i} \frac{A_{OL}^2}{2} + K_{lp_i} \frac{a^2(t)}{2} + K_{lp_i} A_{OL} a(t) \cos[\theta(t) - \phi_i] \quad (3)$$

After some trigonometric manipulations and removing the dc-offset term $K_{lp_i} A_{OL}^2/2$, by subtracting the average of the signals $v_i(t)$, the Eq. (3) can be rewritten as:

$$v_i(t) = L_i s_n(t) + N_i \cos \phi_i s_I(t) + N_i \sin \phi_i s_Q(t) \quad (4)$$

where $s_n(t) = a^2(t)$, $s_I(t) = a(t) \cos \theta(t)$ and $s_Q(t) = a(t) \sin \theta(t)$ are respectively, the in-phase and quadrature components of the baseband signal $s(t)$. Stacking Eq. (4) for $i = 1, 2, 3$, we obtain

$$\begin{bmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{bmatrix} = \begin{bmatrix} L_1 & N_1 \cos \phi_1 & N_1 \sin \phi_1 \\ L_2 & N_2 \cos \phi_2 & N_2 \sin \phi_2 \\ L_3 & N_3 \cos \phi_3 & N_3 \sin \phi_3 \end{bmatrix} \begin{bmatrix} s_n(t) \\ s_I(t) \\ s_Q(t) \end{bmatrix} \quad (5)$$

From (5), we can see that the three output signals $v_i(t)$ represent a linear combination of the three time-variant terms $s_n(t)$, $s_I(t)$ and $s_Q(t)$.

Denoting $\mathbf{v}(t) = [v_1(t), v_2(t), v_3(t)]^T$ and $\mathbf{s}(t) = [s_n(t), s_I(t), s_Q(t)]^T$, where T denotes transpose, and the matrix \mathbf{A} as

$$\mathbf{A} = \begin{bmatrix} L_1 & N_1 \cos \phi_1 & N_1 \sin \phi_1 \\ L_2 & N_2 \cos \phi_2 & N_2 \sin \phi_2 \\ L_3 & N_3 \cos \phi_3 & N_3 \sin \phi_3 \end{bmatrix} \quad (6)$$

Mathematically, the relationship between the base-band signal components, $s_I(t)$ and $s_Q(t)$, and the three output signals of the five-port down-conversion, $v_1(t)$, $v_2(t)$ and $v_3(t)$, can be expressed as:

$$\mathbf{v}(t) = \mathbf{A}\mathbf{s}(t) \quad (7)$$

Considering that the matrix \mathbf{A} is nonsingular, we could solve the linear equation in Eq. (7) by simply inverting the linear system

$$\mathbf{s}(t) = \mathbf{A}^{-1}\mathbf{v}(t) \quad (8)$$

where

$$\mathbf{A}^{-1} = \begin{pmatrix} \gamma_1 & \gamma_2 & \gamma_3 \\ \alpha_1 & \alpha_2 & \alpha_3 \\ \beta_1 & \beta_2 & \beta_3 \end{pmatrix} \quad (9)$$

Substituting Eq. (9) into Eq. (8), we obtain the expressions for the base-band $s_I(t)$ and $s_Q(t)$ signals, as follows:

$$s_I(t) = \alpha_1 v_1(t) + \alpha_2 v_2(t) + \alpha_3 v_3(t) \quad (10)$$

$$s_Q(t) = \beta_1 v_1(t) + \beta_2 v_2(t) + \beta_3 v_3(t) \quad (11)$$

However, we do not know anything about the matrix \mathbf{A} . Recover the signals $s_I(t)$ and $s_Q(t)$ from only the observation of the three output signals, $v_1(t)$, $v_2(t)$ and $v_3(t)$, is a typical *blind source separation (BSS)* problem.

A typical instantaneous linear mixtures problem that benefits from an *Independent Component Analysis (ICA)* solution consists of N observations that are linear combinations of M mutually independent source signals. The observed data $\mathbf{v} = \mathbf{A}\mathbf{s}$ is a linear nonsingular mixture of source signals $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_M)^T$, independent of each other with no information about the sources or the mixing matrix \mathbf{A} [5]. The problem consists of estimating a separating matrix \mathbf{W} , inverse of \mathbf{A} , such that $\mathbf{Y} = \mathbf{W}\mathbf{v}$, we can extract the independent components \mathbf{s} from the output vector \mathbf{y} , such that the components y_i are statistically as independent from each other as possible and are approximately equal to s_i , i.e., $\mathbf{Y} = \mathbf{s}$. To estimate the coefficients in matrix \mathbf{W} , from a finite set of measurements, its necessary to define the optimization criterion based on some objective function, also called the cost function or contrast function, whose global maxima corresponds to a separation of all sources. Once we have defined the optimization criterion, the next step is to define a numerical algorithm to maximize it. Many kinds of blind signal separation algorithms are found in the literature [8]. In this paper, we evaluate the performance of JADE [6], FastICA [7] and K-means [8] algorithms as a blind calibration procedure to five-port receiver.

3. RESULTS

In order to evaluate the performance of the FastICA as a blind source separation technique to regenerate the I/Q signal, we have built a five-port receiver manufactured in microstrip technology designed for 2.4 GHz ISM band, as showed in Fig. 3.

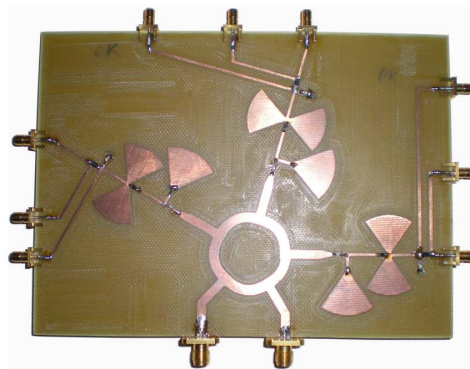


Figure 3: Five-port receiver.

As expected, the measured S -parameters results, in Fig. 4, show that the circuit is matched to 2.4 GHz.

Using the five-port modeling described in [4], we compare, by simulation, the performance, each algorithm, in terms of the probability of bit-error for a 16-QAM modulated RF signal in the AWGN channel, illustrate in Fig. 5.

We observe that the bit error performance of the JADE algorithm provided a performance improvement over FastICA algorithm, result for k-means is very close to the JADE.

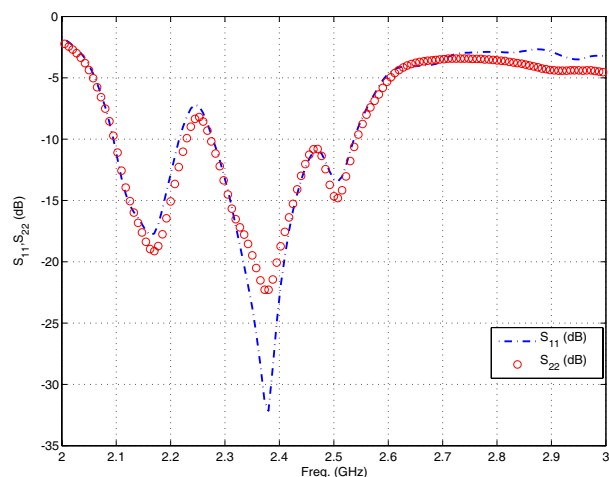


Figure 4: Input reflection coefficient.

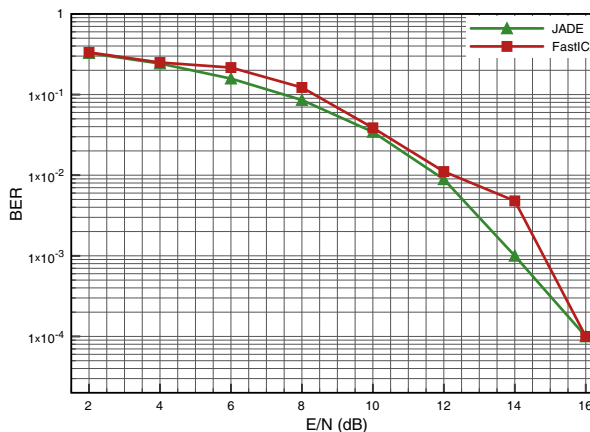


Figure 5: BER results.

4. CONCLUSION

A calibration procedure for five-port receive based on blind source-separation takes advantage of the fact that no prior knowledge about the signal sources is necessary. The success of the BSS algorithms as a blind calibration method for a five-port receiver has been demonstrated by the numerical and experimental results illustrated in Fig. 5. A blind calibration method based on independent component analysis together with a wide-band five-port receiver allow us to design a low cost and low power direct conversion receiver for software defined radio applications.

REFERENCES

1. Haruyama, S., R. Morelos-Zaragoza, and Y. Sanada, "A software defined radio platform with direct conversion: Soprano," *Wireless Personal Communications*, Vol. 23, 67–76, Oct. 2002.
2. Luy, J., T. Mueller, T. Mack, and A. Terzis, "Configurable rf receiver architectures," *IEEE Microwave Magazine*, Vol. 5, No. 1, 75–82, 2004.
3. Neveux, G., B. Huyart, and G. J. Rodriguez-Guisantes, "Wide-band RF receiver using the five-port technology," *IEEE Transaction on Vehicular Technology*, Vol. 53, 1441–1451, Sep. 2004.
4. Sousa, F. R., "Nouvelles applications du corrélateur Cinq-Port: PLL, récupération de porteuse et un MODEM de télécommunications dans la bande 1.8–5.5 GHz," Vol. 1, Éditions Universitaires Européennes, 2010.
5. Principe, J. C., *Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives*, 1st Edition, Springer Publishing Company, Incorporated c 2010, 2010.
6. Cardoso, J. and A. Souloumiac, "Blind beamforming for non gaussian signals," *IEE Proceedings — F*, Jan. 1993.
7. Aapo Hyvärinen, E. O. and J. Karhunen, "Independent component analysis," *Adaptive and Learning Systems for Signal Processing, Communications, and Control*, John Wiley & Sons, Inc., May 2002.
8. Linde, Y., A. Buzo, and R. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, Vol. 28, 84–94, 1980.

Propagation Measurement of a Partially Open Drain: An Examination of the LOS and NLOS Sections

Anwuna Kingsley Awelemdy, Soo Yong Lim, and Kah Phooi Seng
Sunway University, Malaysia

Abstract— This paper presents field measurement results collected from inside a partially open drain at 2.4 GHz, a frequency band that is commonly deployed for wireless local area networks. The motivation of this paper stems from the need to better understand the behaviour of radio waves in partially open drain environments, in terms of signal coverage and attenuation characteristics. Partially open drains are a common sight in several Asian countries. Their design and shapes are similar to a certain degree to that of tunnels, except that they have one less reflecting wall on top. This is one major difference compared to the drainage systems in the US and Europe because the drainage systems in those regions are primarily below ground and covered. Previous research in this area has shown that the existence of the partially open drains makes up an additional channel in which the radio signal can travel. In this work, we are interested to further scrutinize the partially open drain environment in two sections, namely the line-of-sight (LOS) and non-line-of-sight (NLOS) sections using an empirical approach. Towards this end, one Data Acquisition Unit (DAU) has been assembled and tested to be used for data collection. Field measurements were taken at intervals of 30 cm using two dipole antennas that were oriented for vertical-vertical polarization. In this paper, the repeatability of the measurement results that was verified by consecutive measurements is reported. Besides, discussion and insights are offered to analyze the measurement results collected from inside a partially open drain at LOS and NLOS sections. The direct research output of this work is highly relevant to the wireless communications systems design in many Asian cities, where partially open drains dominate.

1. INTRODUCTION

In South-East Asia, it is common to see partially open drains along most buildings and roads. They are a major part of most South-East Asian cities. Wireless Communications is a rapidly growing field, users demand faster data transmission rates and better coverage [1]. To keep up with this growth, it is important to study all aspects of the environment [2]. The propagation characteristics of drain environment is not fully known, it resembles the tunnel environment aside from the absence of the top wall. It is expected theoretically that drains will have some wave-guiding effect on the signals and resemble the tunnel propagation. Wireless system designers need to have reliable information regarding the characteristics of all the structures in an environment for accurate and efficient propagation prediction. In retrospect, ample knowledge of the partially open drain propagation channel is needed for optimal wireless systems design.

Previous research on partially open drains [3] concluded that they provide an additional channel for the signals to travel. This conclusion was drawn in [3] after a series of measurements were performed in the line-of-sight (LOS) portion of the drain under study at various frequencies. In this paper, we are interested to explore further the propagation environment of the non-line-of-sight (NLOS) portion of the partially open drains. Towards this end, field measurements were conducted at one selected location on the Sunway University Campus at 2.4 GHz.

Empirical analysis provides a fairly accurate estimation of the propagation characteristics of any environment [4], although it is usually site-specific. Empirical Analysis aid designers by giving them a general idea of the propagation channel characteristics without having to use complex theories which take a lot of time to compute. It is relatively fast and can also aid the design of better transmitters and well located mobile terminals with reliable and possibly reduced transmission power, hence saving cost.

2. EXPERIMENTAL PROCEDURE

2.1. Measurement Environment

The majority of drains in South East Asia are partially open, this means that they have one less reflecting wall on top, unlike other parts of the world where the drains are usually covered. The dimensions of the drains vary with respect to drainage requirements of the environment. The drain at which the measurement was performed for this paper is approximately 61 m in length, 0.61 m in width and has a depth of 1.52 m. The environment is illustrated in Fig. 1. The drain is

divided into two sections, namely, the LOS section which is 38.4 m in length and the NLOS section which is 22.6 m long. The wall of the drain is slightly rough as it is not plastered or painted. The measurements were conducted when the drain was in a relatively dry condition.

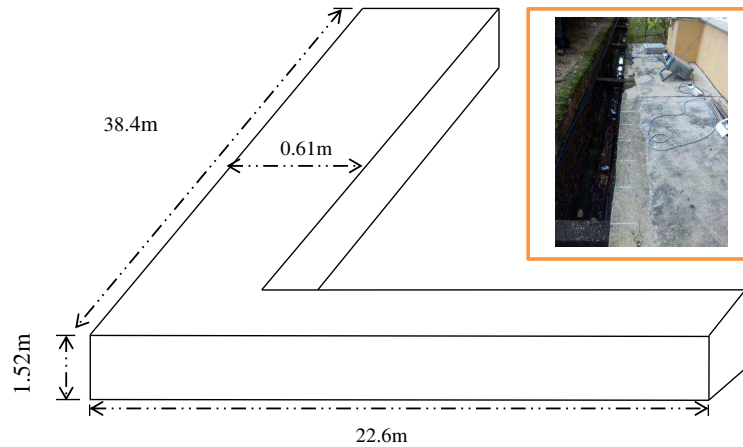


Figure 1: Layout of the measurement environment (not to scale). The inset shows one photo of the actual partially open drain at which measurements were conducted.

2.2. Measurement Setup

One Data Acquisition Unit (DAU) was setup for this project at 2.4 GHz, operating in the zero span mode of the spectrum analyzer. The transmitter unit (Tx) was made up of an Agilent E4428C ESG Analog Signal Generator that can generate signal up to 6 GHz, and an L-Com HGV-2409U dipole antenna that functions at 2.4 GHz. The signal generator was set to transmit a continuous wave at the maximum output of 25 dBm. On the other hand, the receiver unit (Rx) consists of an Agilent E4404B Spectrum Analyzer that works between 9 kHz to 6.7 GHz, and an L-Com HGV-2409U dipole antenna that functions at 2.4 GHz. The antennas were connected to the signal generator and the spectrum analyzer using N-type connecting cables from Mini-Circuits. The antennas have a gain of 8 dBi and were orientated for vertical-vertical polarization. They were placed on two tripods on the transmitting and the receiving ends respectively and their heights were approximately 1.25 m including the tripods. During the measurements, the antennas were placed inside the identified partially open drain with initial Tx/Rx separation distance of 90 cm. Subsequent measurements were taken at 30 cm intervals for a total of 200 data points. The transmitter was stationary, while the receiver was moved step by step. For easy movement, the receiver was placed on a trolley. The received data was sampled at 5 seconds, with 303 points, and then averaged to yield a single value at each interval. The dynamic range of this measurement setup was 90 dB, with the reference level set at 0 dBm and an attenuation of 10 dB. The Resolution and Video Bandwidths were both manually set to 1 kHz. Fig. 2 presents a diagram of the measurement setup.

2.3. Measurement Procedure

The signal was measured and saved at 30 cm intervals. The spectrum analyzer was set to a single sweep of 5 seconds for each measurement location, after which the data was stored. The data for each measurement location was stored in .csv formats which contain the max hold, min hold, state and average power value. The data can be acquired by two ways. One way is to use a GPIB cable to remotely acquire the .csv files which are stored in the C:\ drive of the spectrum analyzer. This can be done with VI programming codes. Another way is to store the files in a floppy disk drive. We opted for the second option, which is the floppy disk drive option. For that, a Matlab script was written to import the files sequentially using the *xlsread* command, find the mean of the data in each file, assign a distance value to each file, and plot a graph of the signal strength versus the distance. The measurements were performed under nice weather conditions. The repeatability of the measurement results was examined when the second measurement was taken a week after the first measurement was performed.

2.4. Measurement Results

It is expected that the signal strength will decrease with increasing distance. The plot in Fig. 3 shows that the LOS section attenuates with increasing distance, and the range of the attenuation

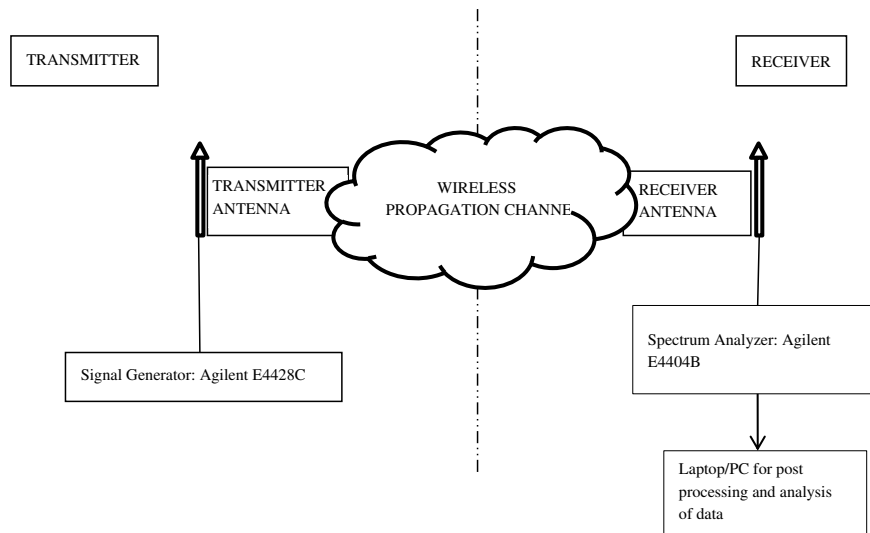


Figure 2: Measurement setup.

is about 50 dB. There is a noticeable signal attenuation of about 30 dB at the transition from the LOS to NLOS sections. Since the measurable dynamic range is about 90 dB, we can observe that the signal fluctuates close to the noise floor when the dominant LOS signal was lost, e.g., from 40 m to 60 m. The measurements were repeated after a week and the two measurements results were compared and plotted together. The standard deviation between the two measurements is 4.9083 dB, which shows good agreement between the two sets of results.

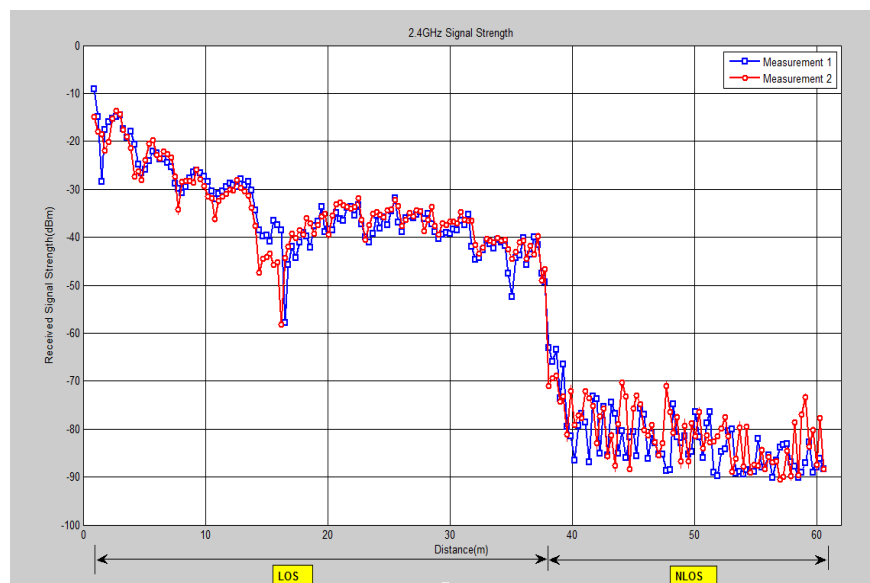


Figure 3: The first and second sets of the measurement results are plotted with indications of the LOS and NLOS sections.

3. CONCLUSION

This paper shows the empirical procedures and results of field measurements conducted inside a partially open drain at 2.4 GHz. The repeatability of the measurement results was proven to be satisfactory, with a standard deviation of 4.9083 dB. The measurement site was carefully selected to contain both a LOS and NLOS sections of the partially open drain as the goal of this paper is to scrutinize the propagation environment of the partially open drain in these two distinct conditions. From the obtained measurement results, we can conclude that the signal propagation is significantly stronger at the LOS section compared to that at the NLOS section. When the

dominant propagation mechanism (LOS) is lost, signal in the NLOS section dropped drastically by about 30 dB, after which it fluctuates up and down near the noise floor.

ACKNOWLEDGMENT

This work was supported in part by the Malaysian government, under the MOSTI (Ministry of Science, Technology and Innovation) E-Science Grant 06-02-16-SF0023.

REFERENCES

1. Iskander, M. F. and Z. Yun, "Propagation prediction models for wireless communication systems," *IEEE Trans. Microwave Theory and Tech.*, Vol. 50, No. 3, 662–673, Mar. 2002.
2. Rappaport, T. S., J. N. Murdock, D. G. Michelson, and R. Shapiro, "An open-source archiving system: RF-propagation measurements and radio-channel modeling and simulation," *IEEE Vehicular Technology Magazine*, 24–32, Jun. 2011.
3. Lim, S. Y. and C. C. Pu, "Measurement of a tunnel-like structure for wireless communications," *IEEE Antennas and Propagation Magazine*, Vol. 54, No. 3, 148–156, Jun. 2012.
4. Rappaport, T. S., *Wireless Communications: Principles and Practice*, 2nd Edition, Prentice Hall PTR, 2002.

Analysis of the Benefits of Diversity Polarization in Wireless Networks

I. Cuiñas and M. García Sánchez

Department of Teoría do Sinal e Comunicaci3ns, Universidade de Vigo, Spain

Abstract— Radio networks have been proposed for multiple applications. Some of them are focused on linking, tracking or surveying people or goods in emergency situations, when wired or even standard wireless networks could collapse. The contents of this contribution are centered in the improvement the use of diversity polarization could provide in terms of availability, which must push these techniques to be implemented in those applications with critical connection abilities. The use of measured co-polar components and synthesized cross-polar ones is proposed, allowing the computation of possible improvements due to polarization diversity in systems operating below 10 GHz, where an advantage of the existing multipath phenomena could be obtained.

1. INTRODUCTION

Fading effects acquire vital importance in the process of designing and planning a radio link in systems operating at frequencies below 10 GHz. Particularly, an due to multipath phenomenon, fast fading appear as a key factor. This kind of fading could represent the propagation dominant factor in the digital radio link systems that operate in the quoted spectral range. The analysis of techniques to mitigate or eliminate fading events becomes a must when deploying radio links, and principally when the final application of the link is focused on assuring communications during rescue or emergency occasions.

This study addresses one of the solutions designed to minimize these adverse effects, which are diversity techniques; and among them, the polarization diversity at reception, by assessing the suitability of their use. Other advantages of diversity techniques, in addition to reducing the percentage of time a given fading influence, are the increased reliability (the existing redundancy), or, depending on the combination of signal processing and subsequent reception, to achieve improvements in quality parameters as signal to noise relation or the error rate. Therefore, the mechanisms of diversity constitute a fundamental contribution to combat the effects of multipath propagation. Previous works report polarization diversity gain at indoor channels, which could be estimated in up to 15 dB at UHF band [1], or lower at higher frequencies (10.5 dB at 2050 MHz [2]), and also outdoors, reporting gains up to 11 dB at 1800 MHz [3]. A comparison of different diversity techniques could be read in [2], where the authors conclude that polarization diversity could be as good as spatial diversity, depending on the environments. However, less research has been done at higher frequencies, over 5 GHz [4].

Thus, this contribution deals with the application of one of such techniques, the polarization diversity, in order to assure an improvement in the disposability of a radio link for communications during emergency events. As most of these situations are related to indoors actuations (fires at houses, gas explosions at inhabited places, terrorism attacks in train stations or airports), various indoors scenarios have been considered to host a large measurement campaign that looked for obtain the radio channel behavior in rooms with different sizes and dimensions, and both furnished and empty. The measurements are described at Section 2.

The outcomes of such campaigns have been analyzed, and they present several fading events. These results, combined with the electromagnetic behavior of the materials at walls, floor and ceiling, have been processed to compute depolarization indexes in order to obtain the depolarized field generated by multipath phenomena within these various environments. The depolarization due to multipath is the topic of Section 3. Once values from both orthogonally linear polarized fields, one measured and the other synthesized, have been obtained, an analysis of polarization diversity at reception can be developed. This analysis is developed at Section 4.

The improvements in propagation conditions obtained as a result of the use of polarization diversity (i.e., reduction of percentage of fading event duration) are presented for indoor environments in a frequency range below 10 GHz, which has direct application in the planning process of emergency radio links, as well as wireless local area network (WLAN).

2. MEASUREMENTS

A large measurement campaign had been performed at frequencies below 10 GHz, involving five different environments and situations. Those are: medium size room line of sight (LoS); large room LoS and brick-wall obstructed LoS (OLOs); and furnished and empty office [5].

Maintaining the transmitter at a fixed location within each environment, the receiver was moving along an automated positioner, following 2.5 meter paths, obtaining the frequency response of the radio channel at each of the receiving points, which was separated eighth wavelengths. Thus, the frequency swept method has been used for getting the initial data by means of a vector network analyzer. After the measurement campaign, a set of 420 files were obtained at each environment, with 801 frequency points each, in module and phase format.

Besides, reflection and transmission mechanisms at different wall obstacles (brick wall, plaster-board, plywood) have been also studied by means of another measurement campaign [6], which included both co-polar and cross-polar contributions. So, information is available to study the depolarization induced by different propagation mechanisms induced by different material walls.

3. DEPOLARIZATION DUE TO MULTIPATH

In the reception end, both co-polar and cross-polar waves arrive. The co-polar wave is mainly what was intended to transmit from the transmission end. Although it is not forced, the cross-polar wave could be useful for the proposal of this paper.

The received orthogonally depolarized wave is the result of several contributions. The initial contribution is the residual emission of the transmitting antenna in the orthogonal polarization. Several contributions are added to this wave, as the interaction of the incident wave at any obstacle forces a change in the wave polarization: a fraction of the incident energy is reflected, transmitted or scattered in the orthogonal polarization. After several interactions, the total amount of depolarized wave could be also large than that organized in the “correct” polarization.

Following that analysis, the depolarized wave reaching the receiver could be synthesized from data measured at only co-polar situation and depolarization indexes [7] obtained from transmission and reflection experimental studies. Thus, it is possible to re-build the cross-polar contribution: the depolarization indexes represent the percentage of electric field strength that changes its polarization. Then, once identified the different paths following by the signals received at each reception point (which, combined, provided the total multipath contribution), it is possible to reconstruct the propagation mechanisms involved at each path: the number of reflections on walls, and the type of material on which the reflection has been occurred; the number of transmission across walls and the obstacle type; and the distance covered in free space. Then, at each transmission or reflection could be applied the depolarization indexes to compute the amount of electric field that has changed its polarization. An example of these indexes is summarized in Table 1.

Table 1: Depolarization indexes in the considered rooms.

Environment	Direct path	One reflection	Two reflections	Three reflections
Empty office, LoS	1.8%	8.2%	16.4%	24.6%
Furnished office, oLoS	1.8%	8.2%	16.4%	24.6%
Medium room	1.8%	0.75%	1.5%	24.6%
Large room, LoS	1.8%	0.75%	1.5%	24.6%
Large room, NLoS	9.4%	15.8%	24%	32.2%

Besides, information on depolarization due to transmitting antenna has to be added. Finally, the results are a pair of electric field strengths at each receiving location: one of them is the measured co-polar component, and the other is the synthesized cross-polar component. Both must be used to analyze the behavior of implementing a polarization diversity technique in reception.

4. ANALYSIS OF POLARIZATION DIVERSITY

The reception, at the same point, of two orthogonally polarized waves at the same frequency could be analyzed as the availability of two different propagation paths. Although the “physical” path (i.e., the “ray”) is the same, as both waves reach the receiver after travelling the same route and beating the same obstacles, virtually there are two signals with low correlation.

Up to three schemes have been tested to implement the polarization diversity technique at reception: the sum, the mean, and the switching. Using a sum scheme, the diversity device provides as an output the sum of both branches signals. If mean scheme is implemented, the result is the average of both inputs. Finally, when switching scheme is applied, the device swaps between both channels getting the maximum strength at each instant.

The estimation of the mean improvement in terms of received power at different points along the measurement path within the various considered scenarios is indicated in Table 2.

Table 2: Mean improvement in terms of received power due to polarization diversity in reception.

Environment	Improvement (%)
Empty office, LoS	18.5
Furnished office, oLoS	21.2
Medium room	10.2
Large room, LoS	17.9
Large room, NLoS	23.3

Analyzing the values at Table 2, the application of polarization diversity seems to provide advantages at all considered scenarios. Looking carefully, the improvement is observed to be large at those situations in which the direct path is obstructed. This obstruction could be partial, as in the furnished office, or complete, as when transmitting from a different room and the line of sight is cut by a brick wall.

The size of the room has also influence, as the larger the room is, the stronger the improvement induced by polarization diversity.

5. CONCLUSIONS

An estimation of the improvement obtained by the use of polarization diversity at reception devices in a wireless network has been presented. Measured data at indoor environments below 10 GHz has been the basis for synthesizing the cross-polar components that would be received within different locations at five specific environments.

We considered frequencies below 10 GHz as it is known that multipath phenomena appear at such bands. Relating to multipath, concretely to transmission and reflection mechanisms, depolarization could be generated. The depolarization indexes related to transmission or reflection mechanisms have been obtained to compute the depolarized waves. And these depolarized waves open the door to the implementation of polarization diversity in reception.

Improvements up to 23% in terms of received power have been estimated, which represent an important increment in the allowable power and permits a better quality of reception.

This study has been developed indoors as multipath phenomena results dramatically highlighted in such environments; but also because many emergency situation could appear at scenarios as considered, and this work contributes to improve the definition of radio communication systems to provide support during emergency tasks.

ACKNOWLEDGMENT

This work was partially funded by Spanish Ministry of Economy and Competitiveness, State Secretary for Research under project TEC2011-28789-C02-02, that was co-funded using European Regional Development Funds (ERDF), as well as the Galician Regional Government under project CN 2012/260 “Consolidation of Research Units: AtlantTIC”.

The authors would like to acknowledge Mr. José Carlos Fernández Ribao for his help during the data processing.

REFERENCES

1. Sánchez, M. and M. G. Sánchez, “Analysis of polarization diversity at digital TV indoor receivers,” *IEEE Trans. Broad.*, Vol. 46, 233–239, 2000.
2. Dietrich, C. B., K. Dietze, J. R. Nealy, and W. L. Stutzman, “Spatial, polarization, and pattern diversity for wireless handheld antennas,” *IEEE Trans. Antennas Propag.*, Vol. 49, 1271–1281, Sep. 2001.

3. Turkmani, A. M. D., A. Arowjolu, P. A. Jefford, and C. J. Kellett, "An experimental evaluation of the performance of two-branch space and polarization diversity schemes at 1800 MHz," *IEEE Trans. Veh. Technol.*, Vol. 44, 318–326, May 1995.
4. Cuiñas, I., M. G. Sánchez, and A. V. Alejos, "Depolarization due to scattering on walls in the 5 GHz band," *IEEE Trans. Antennas Propag.*, Vol. 57, No. 6, 1804–1812, Jun. 2009.
5. Cuiñas, I. and M. G. Sánchez, "Wide-band measurements of nondeterministic effects on the BRAN indoor radio channel," *IEEE Trans. Veh. Technol.*, Vol. 53, No. 4, 1167–1175, Jul. 2004.
6. Cuiñas, I., D. Martínez, M. G. Sánchez, and A. V. Alejos, "Modelling and measuring reflection due to flat dielectric surfaces at 5.8 GHz," *IEEE Trans. Antennas Propag.*, Vol. 55, No. 4, 1139–1147, Apr. 2007.
7. Cuiñas, I., M. G. Sánchez, and E. Amoedo, "Modelling and measuring depolarisation by building obstacles in the 41.5 GHz band," *Microw. Opt. Tech. Lett.*, Vol. 24, 34–36, 2000.

Conception of Radio Frequency Link in the Radio of 3.5 GHz in Consonance with the National Brazilian Rules and International Legislation

T. M. Sanchez Otobo^{1,2}, H. Tertuliano Filho^{1,2}, C. A. Dartora^{1,2},
E. Nascimento Júnior^{1,2}, and E. Cherubini Rolin^{1,2}

¹Department of Electrical Engineering, Federal University of Parana
C.P: 19011, Centro Politécnico, Jardim das Américas, Curitiba, PR 81531-990, Brazil

²Department of Electrical Engineering, Federal Institute of Education
C.P: 19011, Centro Politécnico, Jardim das Américas, Palmas, PR 81531-990, Brazil

Abstract— In Brazil, most of the research works in radio propagation found in the literature are based on unlicensed bands of 900 MHz, 2.4 GHz, 5.4 GHz and 5.8 GHz, because they offers a wide range of applications of interest with the advantage of being free of charge to any user, as far as they do not interfere with other licensed services and if their operation level do not exceed 1 Watt of power transmission. However, for those who use a licensed band (e.g., 2.5 GHz, 3.5 GHz, 7.5 GHz and 8.5 GHz), despite the priority of band use, the Brazilian law for frequency range and service, controlled by the Agency of the Brazilian Telecommunications Services — ANATEL (in accord with international organizations) provides a strong impositions of rules, precedents, decrees and laws that cause, in some cases, projects that are carried out by transposition of frequency or even a complete non-observance. Research centers develop their projects based on theoretical logic but the Brazilian companies do it in complete dissonance with the academic world. Under this scene, the regulatory agency uses successive and constant decrees, to regulate these licensed bands because of the growing number of new applications and services. Due to the huge number of legislative parameters, the radio link designer often neglects important standards or lanes modification, because of lack of knowledge. Within this universe, this article relates how a licensed frequency band specifies a project conducted in strict observance of these standards, where only technical design constraints are considered. After a careful analysis of these projects, a comparative analysis is made with the aim of investigating the impact of the legislation in the design of radio link. Thus, the designer who is subject to these varying norms will be aware of the principal aspects to be considered.

1. INTRODUCTION

The design of a radio link in different frequency bands such as, for example, VHF (Very High Frequency — metric waves — 30 to 300 MHz), HF (High Frequency — decametric waves — 3 to 30 MHz) or UHF (Ultra High Frequency — decimetric waves — 300–3000 MHz) passes through the understanding of the various factors that comprise it. In this work, it is understood that these factors are primarily:

- An engineering plan (challenge, solution and result);
- A legislative plan (application of law);
- A financial plan (feasibility and needs);
- A business plan (to whom and under what conditions).

A designer needs, beyond the sight of these factors also whether the connection will be made via radio, via VPN, via fiber optic interconnection only voice and so on, so that it can establish how many and what are the stages of the project engineering, relevant standards, the financial plan and the business plan will then be developed. It is understood by engineering plan to analyze the design of a radio link from two points of view: A macroscopic and microscopic. From the macro point of view it is understood that the following issues should necessarily be answered: what challenges are addressed and their whys, what are the possible solutions to these challenges and what are the expected results.

From the microscopic point of view, once defined the objectives to be achieved, the concern of the designer should be then parameterize the whole set of factors that will carry out the purposes for which the radio link was created and primarily ensure there is sufficient power available at the receiver, so that communication is established with acceptable signal levels. One cannot expect

an answer to the challenges brought by macro objectives of designing a radio link alone. The optical design macro of a link implies a careful reading and understanding of the marketing plan, financial plan and the legislative plan. But also and especially the radio link — project object — be able to solve the challenges of connecting companies or education centers, interconnect offices, deploy links lan to lan or clear channel, point to point connection, linking enterprise branch, interconnect businesses or only enable the installation of point to point link. Consequently modes of interconnection available in a radio project would:

- **Interconnection via Radio:** This type of link could possibly be designed for users in regions close to neighborhoods or towns where the interconnection between them could be via digital radio system. To solve this challenge, we can undertake the project by connecting with direct visibility (point to point) or through points of repetition. This type of radio link to interconnect on average and depending on the project can reach distances ranging from 50 to a maximum of 100 km with rates ranging from 8.5 Mbps to 10 Mbps speed for any application [1–3].
- **Connection via VPN (Virtual Private Network):** As a radio link using the service interconnection with radio technology achieves speed and traffic capacity, satisfactory interconnection points with VPN technology, depends on the internet link ends exists in order to obtain speed and traffic capacity suitable for each of the points of interconnection. For this type of connection is used, a hardware that communicates between secure and encrypted remote units (offices) and also allows to create access for remote users via VPN Client solution. The VPN Client feature allows any user in any location of the internet, a hotel, for example, in an Internet cafe or a cybercafé work safely as if it were physically in place, accessing all your files on the network, systems documents, among others [4, 5].
- **Connection via Fiber Optics:** The interconnection service via fiber optic is ideal for users who require high data rates at Gigabit speeds, with units in nearby areas, especially large industrial plant extension. This type of link known as corporate link requires additional expertise, because it is broadband project that does not use the same procedures for a project narrowband, and the use of media converters and fiber optic format air (through poles), internally or even underground [6–9].
- **Interconnection of Voice:** Liaison with interconnection of voice is a type of link in which communication between users is the only and exclusively for the use of a landline or a telephone exchange (PABX) that communicates via internet with another unit where the communication is established. Thus it becomes possible to reduce the cost of connection, while it uses a direct communication channel between servers. The Voice Interconnection Service also allows the use of the mobile phone as a branch company which provides mobility and convenience [10, 11].

In the sequence it will presented a specific radio link project at 3.5 GHz to illustrate the case study made without the national and international legislation and a comparison between then in order to verify the constraints and important differences between then.

2. CHALLENGES SOLUTIONS AN D RESULTS

Challenges: those that enable the project or not, only from the viewpoint of engineering. Several would be the challenges to be faced when starting the design of a radio link. So summarized stands out [11]:

- Availability of spectrum;
- Equipment to be used;
- Propagation of the signal;
- Amount of information;
- Distribution of information;
- Location of terminals.

It is noted that any item cannot be handled separately. All of them are iterative for all treatment and if isolated from one or the other can result in binding or poor interference from the point of view of signal or unfeasible from the point of view of infrastructure.

Solutions: whatever may be the solution, they must meet not only the items listed above in the challenge, as they must also meet the objectives macros. The solutions do not usually derail

the project of a radio link, but it can substantially burdens due to inappropriate actions. It is suggested that the following question could be applied early in the design of the loop in order to provide solutions for some problems that may arise in the design stage:

- What type of service and what your rate?
- Who are the solution providers?
- What is the spectrum available for such application?
- What are the legislative restrictions?
- What database is available and what is its quality?

Results: the best possible outcome for the designer is to have the link operating in the best and most strict quality standards that in this case translates it to the need of solving the challenges, application of the national and international current legislation and one specific financial plan. All of that in a perfect working condition [14].

3. IMPORTANT LEGISLATION APLLIED

To start a radio link project, a number of national and international relevant legislation must be applied in consonance with ANATEL — Agency of the Brazilian Telecommunications Services and the ITU (International Telecommunication Union). For a frequency of 3.5 GHz (case of study) which is relevant:

- Transmission power cannot exceed the value of 44.47 dBm;
- Emission level of spurious signals limited to -26 dBm;
- Canalization of the frequency band and limitations of sub-band;
- Analyze transmission rates;
- Class service;
- Analyze primary use of the service;
- Analysis service in secondary use;

For a frequency of 3.5 GHz which are the applicable law:

- Anatel Resolution Nos. 164,166, 309, 295, 416, 537;
- Radio Regulations of the ITU — International Telecommunication Union — Section S1-20;

For the design of a radio link in the range of 3.5 GHz which the applicable law:

- Attenuation in free space: ITU-R Recommendation ITU-R 525-2 and 341-5;
- Propagation attenuation: ITU-R 341-5;
- Attenuation due to sand and dust: CCIR Report 721-3;
- Attenuation due to vegetation: Recommendation ITU-R 833-3;
- Transmission losses: Recommendation ITU-R 341-5;
- Diffraction losses: Recommendation ITU-R 526-5;
- Attenuation by atmospheric gases: ITU-R Recommendation ITU-R 676-3 and 676-4;
- Factor of earth curvature and refraction indices: Recommendation ITU-R 453-6 and 453-7 ITU-R and ITU-R 369-6;
- Release Criteria signal: Recommendation ITU-R 530-8;
- Power Levels: Anatel Resolution 537;
- Levels of spurious signals: Anatel Resolution 537;
- Polarization: Recommendation ITU-R 838;
- EIRP: Anatel Resolution 164, Resolution 537, Resolution 88 and 323 Precedent;
- Signal Noise Minimum Recommendation ITU-R 526-5;
- Attenuation System: Recommendation ITU-R 341-5;
- Global attenuation: ITU-R 341-5;
- Use of antennas: Anatel Resolution 367 and Recommendation ITU-R 699-5;
- Use of cables: Anatel Resolution 399 and Resolution 470;
- Use of towers: Practice Telebras 240-410-600;
- Use transmitting equipment: Anatel Resolution 303;
- Use of receiver equipment: Anatel Resolution 303;
- Plan of the City of Curitiba: Law 11.535, Decree 606, Anatel Resolution 533, 529 and 571.

4. SIMULATIONS AND RESULTS

In order to compare and analyze the national and international effects of the actual legislation in a conception procedure of a radio link it was choose the technical data for a real link operating at 3.5 GHz and the impositions to make a project are summarized in Table 1:

Table 1.

	Point A (Transmitter)	Point B (Receiver)
Latitude	25°26'52.34"	25°25'30.60"
Longitude	49°16'33.24"	49°17'39.66"
Altitude	912 (m)	942 (m)
Height of Antennas	40 (m)	95.5 (m)
Diameter Antennas	0.6 (m)	0.5 (m)
Free Space	Friis formula	
Total loss	≤ 135 dB	
Transmit Power	33 W	
Maximum Signal level	-40	
Type of Antenna	Directional	
Antenna Efficiency	72%	
Transmission Line	Coaxial Cable RGC58	
Cable Loss	3.2 dB	
Connector Loss	0.8 dB	
Loss in Filter Duplex	2.5 dB	
VSWR at the output of filter	≤ 1.5	
Máximum E.I.R.P	-106.4 dB	
Reliability	92%	

The procedure for the accomplishment of the simulations was based on real life data collected by mine of one drive-test that has been supplied by a major local operator of cellular telephony, where by through a simulation software and a signal prediction we have obtained, with the integration of the data of drive-test, a level of signal considered for such tool. The Figure 1 choose the signal path of the points of transmission and reception.

The CelPlanner of the CelPlan Global Wireless Technologies available in our laboratories in version 7.4 was initially configured with the significant data of the relief, the morphology and the topography, beyond images of the city of Curitiba where drive-test was become fulfilled [13]. A Figure 2 below, it is possible to see the image of the route which the data was collected in the downtown of Curitiba in the urban area and also in the dense urban area.

Below in Figure 3, the morphology of the region where it was made the drive-test can be



Figure 1.

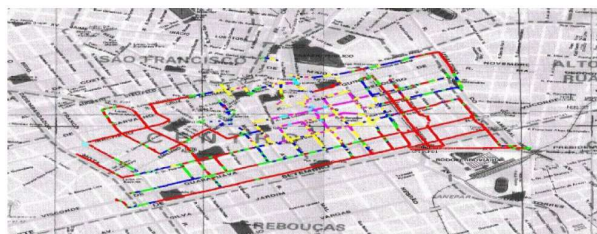


Figure 2.

observed.



Figure 3.

The antennas used in the simulations in its majority are of KATHREIN Mobilcom Brazil model 742212, the diagrams of irradiation for each TILT were supplied by the proper manufacturer. For the visualization of the height, tilt, azimuth and geographic coordinates of the ERB one software of maps was used. The coordinates of each used ERB was supply by the major operator local company. In the Table 2, it are summarized the obtained results for simulation of the real life data for the downtown city of Curitiba.

The used propagation model for simulation was a free adjustable model that can be configured by the user. The canalization of the frequency range for this frequency is presented in Figure 4.

After this procedure it is presented in Table 3 below the obtained results for simulation when the legislation is applied.

Table 2.

	TRASMITTER	RECEIVER	UNITIES
Latitude	25.447	-25.425	Degree
Longitude	-49.275	-49.294	Degree
Operation Range	3500 a 3600		MHz
Receiver Azimuth (za)	143.732		Degree
Transmitter Azimuth (zb)	323.72		Degree
Elevation angle	1.71		Degree
Great Circle Distance	3.13		km
Transmitter Antenna Gain	19.5		dBi
Receiving Antenna Gain	18.7		dBi
Total Gain	38.2		dB
propagation loss	113.2		dB
Transmission loss	75		dB
Wavelength	0.0856		M
Power Received	-29.82		dBm
Margim Signal at reception	10.01		dB
E.I.R.P	60.69		dBm
Total Path Loss	117.2		dB

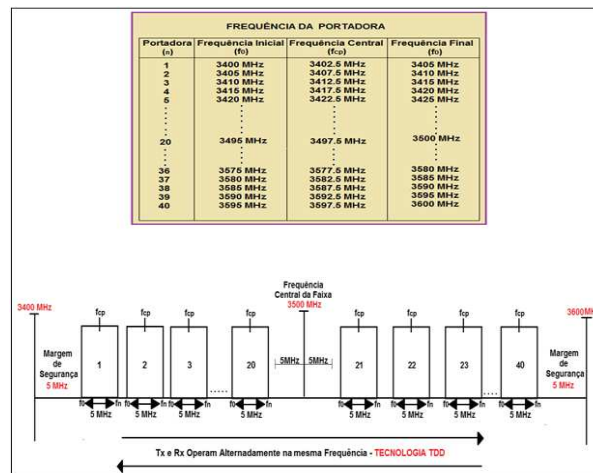


Figure 4.

Table 3.

Parameter	Results
The system gain value;	119.47 dB
Lever of thermal noise;	-102.39 dBm
Reception threshold;	-75 dBm
Fading margin for interference;	56.5 dB
Carrier signal/noise;	27.39 dB
Signal level at the reception;	-17.5 dBm
Carrier signal/noise in the absence of fading;	84.89 dB
Fading margin compound;	99 dB
Dispersive fading margin;	42.5 dB
Thermal fading margin;	57.5 dB
Interference level co-channel	-106 dBm
Carrier signal/interference	88.5 dB
Degradation due to the interference	1 dB
Netloss on route	61.97 dB

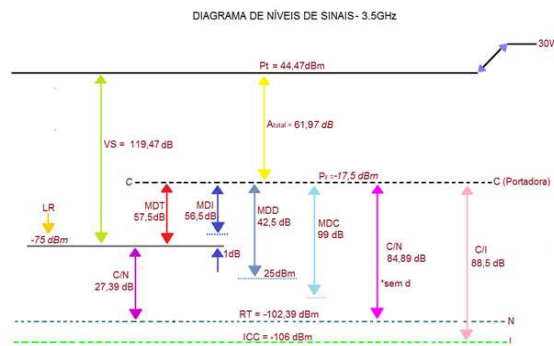


Figure 5.

The Figure 5 summarizes the obtained signals level obtained for the radio link in according with the legislation.

5. DISCUSSION

This paper presented a radio link conception at 3.5 GHz based in the national Brazilian laws. The analyses of the level of signal demonstrate that the radiate power, level of received power, the system gain value in according with the rules are all of them in according with the permitted

values.

6. CONCLUSION

In this article the steps considered important to make a conception of a radio link in consonance with the Brazilian national laws and the international ITU regulations was presented and discussed. The current laws were observed and analyzed for the frequency range of 3.5 GHz. It was observed that there is a significant difference in values when the resolution is applied to the conception of a radio link in spite the other one. The link performance evaluation depends on the factors listed here for a greater or lesser extent concerning the same performance.

ACKNOWLEDGMENT

The authors are grateful to Celplan Technologies, Campinas, SP, Brazil for the permission to use the software under their responsibility to produce these important results, their scientific, support and technical cooperation and also to TIM — Telecom Italia Mobile for the furnished drive-test real-life data.

REFERENCES

1. Halliday, C., *The Telecommunications Secrets*, Berkley Brasil Editoras, Rio de Janeiro, 1993.
2. Courbis, P., *Voyage ao Centre du Monde des Telecommunications*, Editions Angkor, Paris, 1993.
3. Tunc, C. A., A. Altintas, and V. B. Erturk, “Examination of existent propagation models over large inhomogeneous terrain profiles using fast integral equation solution,” *IEEE Transactions on Antennas and Propagation*, Vol. 53, No. 9, Sep. 2005.
4. Zimmermann, D., J. Baumann, A. Layh, F. Landstorfer, R. Hoppe, and G. Wolffe, “Database correlation for positioning of mobile terminals in cellular networks using wave propagation models,” *IEEE 60th Vehicular Technology Conference*, 2004.
5. Athanasiadou, G. E., I. J. Wassell, and C. L. Hong, “Deterministic propagation modeling and measurements for the broadband fixed wireless access channel,” *IEEE 60th Vehicular Technology Conference*, 2004.
6. Sarkar, T. K., Z. Ji, and K. Kim, “A survey of various propagation models for mobile communication,” *IEEE Antennas and Propagation Magazine* Vol. 45, 51–82, Jun. 2003.
7. NETO, Vicente Soares, et al., *Telecomunicações: Sistemas de Propagação e Rádio Enlace*, Érica, Ed., São Paulo, 1999.
8. Freeman, L. R., *Radio Systems Design for Telecommunications*, 2nd Edition, Wiley, 1997.
9. Bertoni, H., *Radio Propagation for Modern Wireless Systems*, Prentice Hall, 2000.
10. Rappaport, T. S. and L. B. Milstein, “Effects of radio propagation path loss on DSCDMA cellular frequency reuse efficiency for the reverse channel,” *IEEE Transactions on Vehicular Technology*, Vol. 41, No. 3, 231–242, Aug. 1992.
11. Tertuliano Filho, H., *Metodologia de Equacionamento e Parametrização de Enlaces de Rádio Frequência*, Editora UFPR, Curitiba, PR.
12. Oh, S.-H., “Prioritized channel assignment in a cellular radio network,” *IEEE Transactions on Communication*, Vol. 40, 1259–1269, 1992.
13. Felice, F., “Análise do desempenho de enlaces ponto-a-ponto utilizando a faixa de frequência não licenciada de 2.4 GHz em tecnologia spread spectrum,” Dissertação de Mestrado, Departamento de Engenharia Elétrica, Universidade Federal do Paraná, Curitiba, PR, 2008.

Some Insights on the Amount of Fading in Radio Channels

Hassan El-Sallabi¹, Khalid Qaraqe¹, and Erchin Serpedin²

¹Wireless Communications Laboratory, Electrical Engineering Department
Texas A&M University at Qatar, Qatar

²Electrical and Computer Engineering Department, Texas A&M University
College Station, TX, USA

Abstract— In this work we present an experimental investigation on the amount of fading as a measure of fading severity in radio channels. The presented results are for indoor environments and assume a multi-dimensional ray based model. The amount of fading of Rayleigh channel, which is equal to one, is used as a reference to know if the channel is less or more severely affected than Rayleigh radio channel. For the tested environment, it has been found that placing the antenna of indoor cell at a height close to the pedestrian height leads to less fading than placing it on the ceiling and at this preferred height. The amount of fading for line of sight propagation conditions exhibits higher correlation with coherence time than that corresponding to non-line of sight propagation.

1. INTRODUCTION

Wireless radio channels are known of their fading phenomena, where received signal fluctuates randomly due to the multipath arrival of radio signal at the receiver point. Fading phenomena can be fast fading and slow fading. Slow fading describes the signal variation slowly over tenths of wavelength distance. It is usually modeled as lognormal distribution. The fast fading results due to the phase difference of multipath components and user equipment (UE) speed defined by locations of scatterers and their electrical parameters as well as the position and antenna height of the transmitter and receiver and their antenna patterns. The geometry of the propagation scenarios defines the angular and delay domain properties of the channel. These properties are directly related to fast fading. The fast fading has been modeled in literature with many models based on the existence or non-existence of the line of sight (LOS) component. The widely known models are the Rayleigh model for non-line of sight (NLOS) component and Rician model when the LOS component exists. There are already many models that describe the fading channel subject to different fading patterns. A survey of various propagation models for mobile communications is presented in [1]. A real channel could be a combination of scenarios and propagation conditions described by different models. Instead of describing the channel with a particular model, we simulate the channel with a physics based channel model that represents how the signal might propagate in a multipath environment. The severity of fading is described in this work by a measure called the amount of fading. This measure is used to characterize different models that describe the fading conditions in radio channels. In this work we present some detailed insights on characterizing the amount of fading as a measure of fading severity in radio channels.

2. AMOUNT OF FADING

Due to the fading nature of radio channels, which results in signal fluctuations, diversity methods were introduced to improve performance of communications system. The focus of these methods is to reduce the variability of the received signal power to overcome the unreliability of the communication system due to poor fading conditions. Average signal-to-noise power ratio (SNR) at diversity combiner output ignores the variation of the received SNR, and presents the diversity benefit without increasing the transmit power. In order to study the impact of SNR variance, a measure that accounts its variability is needed. This measure has to account for higher moments of SNR at the diversity combiner output. The selected measure of severity of fading in this work is the amount of fading (AF), which can be computed using the first and second central moments of SNRs at diversity output. The AF is defined in [2] as

$$AF = \frac{var\{\alpha^2\}}{E\{\alpha^2\}^2}$$

where α is the instantaneous fading amplitude of a complex fading channel, $E\{\cdot\}$ and $var\{\cdot\}$ are the statistical mean and variance, respectively. To quantify the probability distribution of fading, it

is mentioned in [2] that for Nakagami- m fading distribution of α , $AF = 1/m$, whose range is $[0, 2]$. When $m = \infty$, $AF = 0$, which corresponds to the situation of “no fading”. When $m = 1$, $AF = 1$, which corresponds to Rayleigh fading, and when $m = 0.5$, $AF = 2$, which corresponds to the one-sided Gaussian distribution and the severest fading assumed by the Nakagami- m fading channel. The AF has been generalized in [3] to describe the performance of the diversity combining system over correlated lognormal channels. The inverse of AF might be considered as diversity factor. As noted in [4], the amount of fading relates directly with diversity order of the symbol error probability for Nakagami- m fading distribution and is not linked directly to the average symbol error probability of Nakagami- q (Hoyt), Nakagami- n (Rice) fading distributions. The AF is used to quantify the level of fading experienced at the output of a multiple-input multiple-output (MIMO) system in [5]. In [6], different formulations for AF for different fading distributions are given for Nakagami- q (Hoyt), Nakagami- n (Rice), Nakagami- m , Weibull, and log-normal shadowing fading distributions.

3. CHANNEL MODEL

The investigation of this work is simulation based using multi-ray an RF multi-dimensional propagation model. The RF model is for indoor cubical shaped environments that can be used to represent a corridor, office, lecture hall, convention center, etc., as a function of the indoor environment geometry. The model formulation in its essence is very similar to that presented in [7]. The input parameters of the model include operating frequency, system bandwidth, signal polarization, number of wall, floor and ceiling reflections, and electrical properties of reflecting surfaces. The ray parameters are derived from the environment geometry using electromagnetic image theory and vector mathematical operations to extract the needed path length and angles for every ray in terms of angle of arrivals and departures as well as surface reflection angles. Each ray is defined by its parameters such as its path length, azimuthal and co-elevation angle of arrival, azimuthal and co-elevation angle of departure. Its amplitude is computed with electromagnetic formulations for free space loss and loss due to interaction with scatterers in the environment. The interaction with scatterers takes place in different propagation mechanisms: reflection, transmission, diffraction and scattering. In this work only the reflection propagation mechanism is included. Different coefficients can be used for interaction losses that depend on waveform, plane wave, cylindrical wave or spherical wave. The most commonly used reflection coefficient is the Fresnel plane wave reflection coefficient, which is valid for flat surfaces and is a function of the incidence angle and the electrical properties of the reflecting surface. The received signal is obtained as a sum of multi-ray components as vector superposition of N individual rays, and it can be represented as follows:

$$h(t) = \sum_{n=1}^N A_n \delta(t - \tau_n) e^{-jk(r_n - \mathbf{V} \cdot \boldsymbol{\Psi}_n t)},$$

where k is the wave number expressed as $k = \frac{2\pi}{\lambda}$, λ is the wavelength of operating frequency, \mathbf{V} denotes the velocity vector of the UE, which is assumed at the receiver in this notation, and defined by $\mathbf{V} = v_x \vec{x} + v_y \vec{y} + v_z \vec{z}$ and $\boldsymbol{\Psi}_n$ stands for the arrival direction vector defined for ray n as

$$\boldsymbol{\Psi}_n = \cos(\phi_n) \sin(\theta_n) \vec{x} + \sin(\phi_n) \sin(\theta_n) \vec{y} + \cos(\theta_n) \vec{z}$$

where ϕ_n represents the horizontal arrival angle relative to the x -axis of ray n , θ_n denotes the elevation arrival angle relative to the z -axis of ray n , and r_n is the path length of ray n defined as:

$$r_n = \sum_{p=1}^{P_n} d_{n,p}$$

Notation $d_{n,p}$ denotes the distance traversed by the specular wave between the $(p-1)$ and p -th boundary intersections and the complex amplitude A_n is defined as

$$A_n = \frac{\lambda}{4\pi r_n} \sqrt{G_{tx}(\varphi_n, \vartheta_n) G_{rx}(\phi_n, \theta_n)} \prod_{p=1}^n \Gamma_p e^{-jk d_{n,p}}.$$

Parameter Γ_p denotes the surface reflection coefficient for the p -th wave-interface intersection, while the term $\frac{\lambda}{4\pi r_n}$ represents the free space path loss that accounts for the wave spreading loss, and finally, $G_{tx}(\varphi_n, \vartheta_n)$, $G_{rx}(\phi_n, \theta_n)$ are the transmitter and receiver antenna gain, respectively.

4. NUMERICAL RESULTS

The simulated results represent a lecture hall of a convention center indoor environment. The simulation results show the impact of antenna height and operating frequency of indoor cell. The receiver antenna height was 1.7 m. The lecture hall room dimensions are as follows: width is equal to 10 m, length is 15 m and height is 10 m. The receiver speed was 3 km/hr which is defined as the pedestrian speed in 3GPP standard. The fast walking speed of 10 km/hr is tested for AF comparison. Reflecting surfaces have permittivity of 5 and conductivity of 0.02. Rays with surface reflection order of up to 6 are included in addition to the line of sight component. This reflection order is per surface. This means that when the signal bounces between two walls, then rays of up to 12 reflection order could result and so on for rays bouncing between three surfaces would result in rays of up to 18 reflection order. In order to study the impact of operating frequency on the amount of fading, two different frequency ranges have been tested 0.9 GHz and 1.8 GHz. These bands represent cellular and personal communications (PCs) bands. The simulated temporal range is for 2 seconds for every spatial location. The temporal sampling rate is 26000 samples/sec. The simulated spatial range is 5 m with a spatial resolution of 2.5 cm. Antenna polarization is vertical. The three dimensional antenna pattern is the well-known omnidirectional pattern, where signals propagating in the xy -plane experience no antenna loss irrespective of azimuthal angles but the attenuation takes place in the non-horizontal plane propagation, where the amount of antenna loss depends on the elevation angles. The antenna loss increases as the elevation angle of the arrived signal gets away from the 90-degree plane (i.e., xy -plane). The minimum value of AF is zero, which means no signal variation. For different channels of similar mean value, the AF is a good indicator for comparing the severity of fading between them. In order to get the feeling of the AF numbers, Figure 1 shows the cumulative distribution function of the fading patterns for Rayleigh, double Rayleigh and triple Rayleigh with their corresponding AF values as 1, 1.73, and 2.56, respectively. The double Rayleigh fading exhibits double severity of fading than the single Rayleigh, which means the channel varies twice faster. The Rayleigh fading channel is just one type of fading channels. Our model is a multi-ray model, whose ray parameters change with the movement of mobile receiver environment dimensions and communication link setup such as antenna heights, polarizations, etc. Rate of changes of ray parameters depend on mobility velocity. This work investigated the impact of indoor cell antenna height. Two antenna heights have been tested: 2 m representing a similar height to receiver and 9.9 m (presented as a 10 m height in the figures below), which represents an antenna mounted on the ceiling. Figure 2 depicts the empirical cumulative distribution function of AF computed from channel traces for 900 MHz and 1800 MHz frequency ranges generated with the channel model described earlier. The AF ranges from slightly greater than zero till AF is slightly greater than 2. This indicates that the channel does not follow one fading model and the severity of channel fading varies too. Figure 2 shows that for this particular tested line of sight propagation scenario, placing antenna on ceiling causes more severe fading than placing the antenna at a height close to the receiver antenna height. It can be read from the figure that when the transmitter antenna height is 2 m for both frequencies, 95% of channel traces are less severe than Rayleigh fading, while when the transmitter antenna height is 10 m

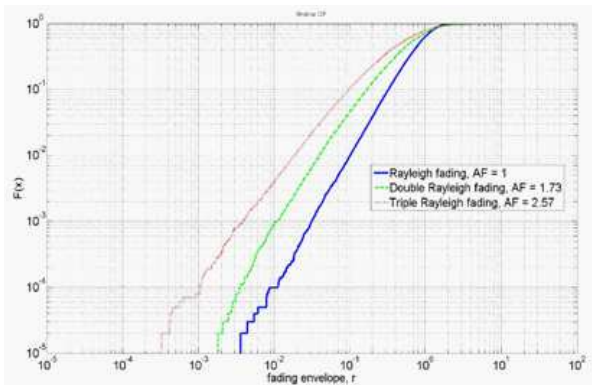


Figure 1: Different Rayleigh fading with their corresponding AF values.

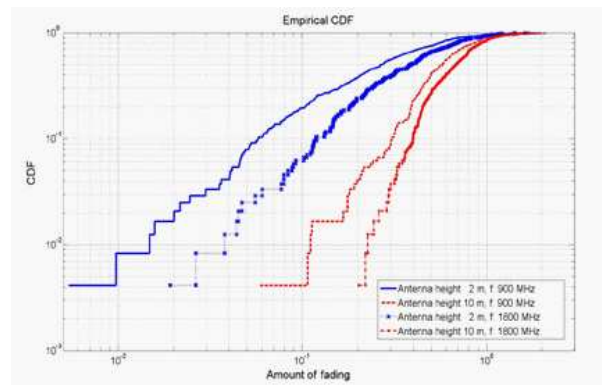


Figure 2: CDF of AF for moving MS with different transmitter antenna heights, $f = 900$ MHz.

10 m, about 90% and 85% of channel traces are subject to less fading severity than Rayleigh fading for 900 MHz and 1800 MHz frequency ranges, respectively. Furthermore, channel traces of both antenna heights have very low percentage of fading severity greater than double-Rayleigh fading. Samples of received signal strength due to different channel traces with different AF values for the tested frequency ranges are depicted in Figure 3. The presented AF values are well correlated with the fading patterns of the received signal fading profile. Figure 4 shows a comparison between the AF of LOS and NLOS propagation scenarios. The channel environment and communications link are the same. The only difference is the blocking LOS rays, which are not considered in the computation of channel traces in the NLOS propagation conditions. It is pretty clear that the AF for NLOS is higher than that of LOS as expected. The medians of AF for LOS at 900 MHz and 1800 MHz are 0.28 and 0.39 compared to the corresponding values of NLOS, which are 0.36 and 0.55, respectively. The AF of fading channel traces is well correlated with the coherence time of the channel. The computed correlation value of coherence time is 0.5. Figure 5 shows a scatter plot of coherence time of simulated channel traces with their AF for the LOS case with indoor cell antenna height of 10 m and frequency range of 1800 MHz. Table 1 shows the corresponding correlation values for LOS and NLOS propagation conditions for the two tested antenna heights and frequency ranges. Figure 6 shows the impact of mobile speed on AF for two different speeds 3 km/hr and 10 km/hr for indoor cells at a frequency range of 900 MHz and two different indoor cell antenna heights. The severe fading pattern is observed at higher speeds when the antenna of the indoor cell is placed very close to the ceiling. The relative increase in the median of AF for a low antenna height with the increase of mobile speed is higher than that of a higher antenna height.

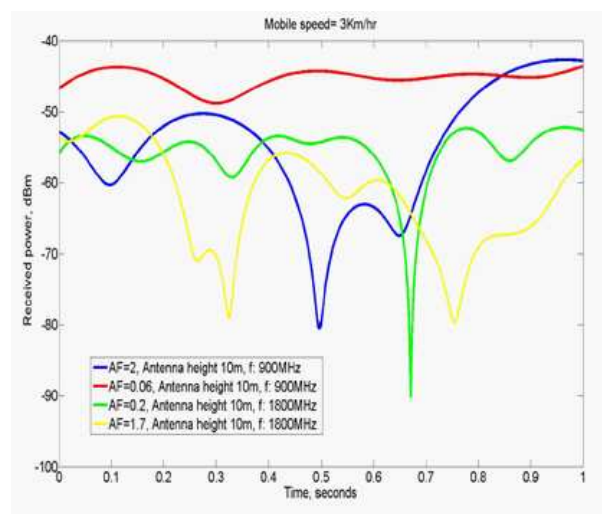


Figure 3: Impact of frequency range on AF.

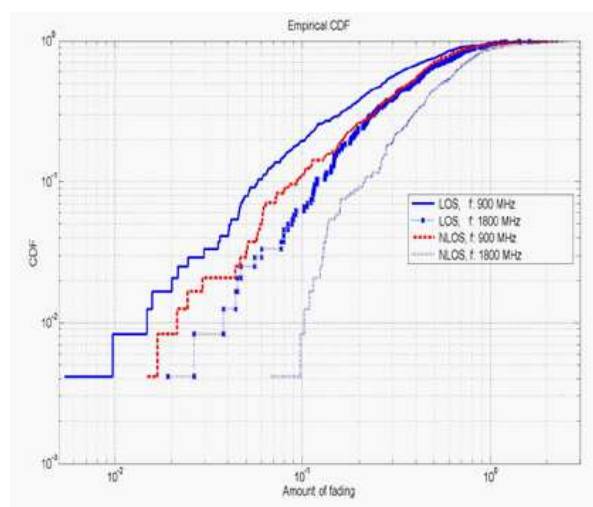


Figure 4: CDF of AF for moving MS with indoor cell antenna height of 2m for LOS and NLOS.

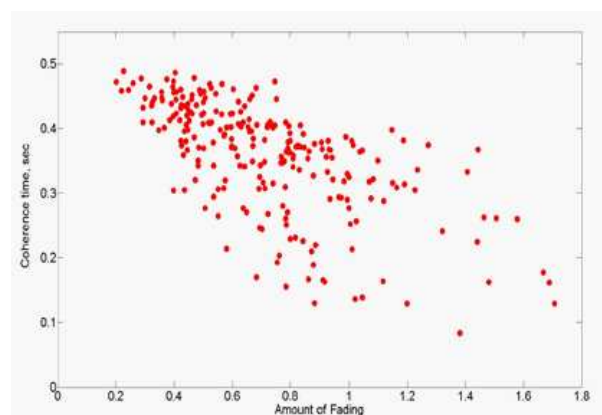


Figure 5: Coherence time versus amount of fading.

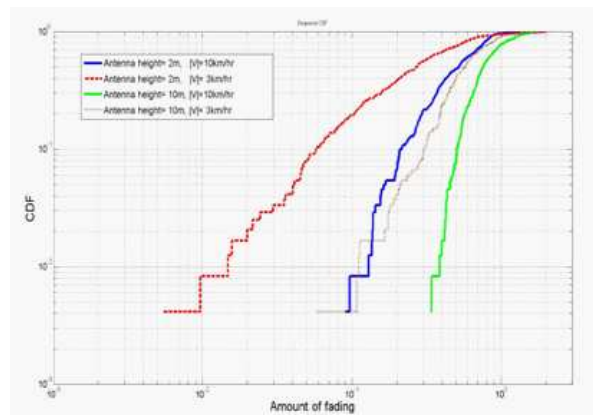


Figure 6: Samples of received signals for different channel traces with two different values of AF.

Table 1: Correlation values between AF and coherence time.

Frequency (MHz)	900		1800	
Antenna Height (m)	2	10	2	10
LOS	-0.81	-0.62	-0.82	-0.65
NLOS	-0.82	-0.65	-0.75	-0.54

5. CONCLUSION

Amount of fading is a parameter that describes the severity of channel fading conditions. Physics based modeling shows that the channel presents variable AF, which means that there is not a single statistical model such as Rayleigh but the latter represents only one possible case of channel realizations. According to the simulated environment, placing the antenna of the indoor cell on the ceiling causes more fading than if it is on a height close to the pedestrian height. As expected the higher operating frequency range and higher mobile speed lead to higher AF. The AF corresponding to LOS propagation conditions presents higher correlation with the coherence time than that of NLOS.

ACKNOWLEDGMENT

This publication was made possible by NPRP grant #: NPRP 09-341-2-128 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

1. Sarkar, T. K., et al., "A survey of various propagation models for mobile communication," *IEEE Antennas and Propagation Magazine*, Vol. 45, No. 3, 51–82, Jun. 2003.
2. Charash, U., "Reception through Nakagami multipath channels with random delays," *IEEE Trans. on Commun.*, Vol. 27, No. 4, 657–670, Apr. 1979.
3. Alouini, M. and M. Simon, "Dual diversity over log-normal fading channels," *Proc. IEEE Int. Conf. Commun., ICC'01*, Helsinki, Finland, Jun. 2001.
4. Wang, Z. and G. B. Giannakis, "A simple and general parameterization quantifying performance in fading channels," *IEEE Trans. on Commun.*, Vol. 51, No. 8, 1389–1398, Aug. 2003.
5. Holter, B. and G. E. Øien, "On the amount of fading in MIMO diversity systems," *IEEE Trans. on Wireless Commun.*, Vol. 4, No. 5, 2498–2507, Sep. 2005.
6. Simon, M. K. and M. S. Alouini, *Digital Communication over Fading Channels*, 2nd Edition, Wiley, New York, 2005.
7. Malik, W. Q., C. J. Stevens, and D. J. Edwards, "Spatiotemporal ultrawideband indoor propagation modelling by reduced complexity geometric optics," *IET Commun.*, Vol. 1, No. 4, 751–759, 2007.

RF Identification (RFID) Reader with Long Read Range in the UHF Band

Jae-Young Jung, Kyu-Won Han, and Chan-Won Park

Electronics and Telecommunications Research Institute (ETRI), Republic of Korea

Abstract— In this paper, the design of a long read range UHF RFID reader using a novel carrier leakage suppression method is presented. We developed and demonstrated through experiment in the laboratory environment in the carrier-to-tag data ratio by more than 36 dBc in the UHF RFID band, which is consistent with our simulation results. When inserted between the reader antenna and a Class-1 RFID reader, measurements show that the reader can successfully demodulate the generated tag signal for a power level of -80 dBm, which corresponds to the backscattered power at 20 m distance. The read range of the reader with the proposed carrier leakage suppression front-end is verified again in a real environment. The measured read range is found to be greater than 20 m for a commercial battery-assisted passive (BAP) tag.

1. INTRODUCTION

Recently, radio frequency identification (RFID) utilizing the 900 MHz frequency band has been widely used for a variety of applications such as inventory and supply chain management [1, 2]. A conventional passive RFID system, which is categorized as Class-1 [3], is composed of a reader and passive tags. The reader interrogates passive-tag attached items by sending an approximately 1W signal. The tag sends back the data stored in an internal memory to the reader using backscattering modulation. Backscattering modulation is formed by reflecting a received signal from the reader by altering the antenna load impedance in the passive tag according to the tag information, which results in modulating the phase and magnitude of the received signal. The reader then accesses the tag information from the backscattered signal. The read range is one of the key parameters in RFID systems and is frequently evaluated for RFID system assessment. The read range of about 20 m may be appropriate for the emerging application of long read range RFID systems, such as vehicle toll systems and the management of airline cargo containers and shipbuilding yard equipment [4]. Most Class-1 commercial readers show a receiving sensitivity of -60 to -70 dBm and the read range of commercial readers is limited to about 10 m. Carrier leakage power of the transmitter of a reader seriously degrades the reader sensitivity. The phase noise of this leakage causes a difficulty in the demodulation of a weak backscattered signal. Filters are not efficient in such leakage suppression because the spectrum of the leakage and that of the backscattered signal is so close. The difference is estimated to be about $40 \sim 640$ kHz according to ISO/IEC 18000-6C. Many researches for leakage suppression have been carried out using the method the leakage suppression is achieved by adding a 180° inverted signal, obtained using variable attenuator and phase shifter. In this paper, we propose the novel leakage suppression front-end using received signal alone. Transmitter signal is not used. Thus, the leakage suppression is expected to be less dependent on time and environmental changes.

2. DESIGN OF THE RF FRONT-END

The leakage power at the receiver consists of the leakage from the circulator due to a finite isolation (leakage 1) and that from the antenna mismatch (leakage 2). Typical isolation of a circulator is about 30 dB, and the leakage due to the isolation of the circulator may be treated as almost constant. However, leakage 2, the leakage from the antenna mismatch, may be sensitive to environmental changes and is considered to be time varying. Most leakage suppression methods use the direct coupling compensation architecture [5–7]. The amplitude and phase of the sampled transmitter signal are tuned by variable attenuator and phase shifter to cancel the carrier leakage. The amplitude and phase of the sampled signal are ultimately tuned to cancel the carrier leakage. The block diagram of our RF front-end for carrier leakage suppression is shown in Fig. 1. Input port in Fig. 1 is connected to the receiver input of the antenna. The output port is connected to the receiver input of the reader. The combined signal of the leakage 1 and 2 appear at the input of our carrier leakage suppression front-end which makes the input signal slowly time-varying. In order to extract the carrier, the received signal is split and the carrier is obtained from the limiter and filter. The filter is used for harmonic elimination. The amplifier before the limiter is used to

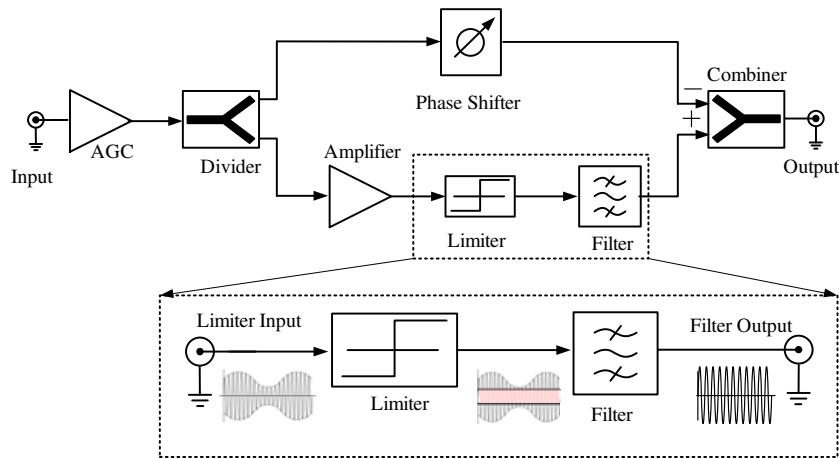


Figure 1: Proposed carrier leakage suppression front-end.

provide sufficient power to the latter. Note that the output power of the limiter is almost fixed for changes in the input power. Thus, the output power level of AGC is selected for the output power of the phase shifter to yield exactly the same as the filtered output power of the limiter. This can be realized using the AGC control input. Once set, the output powers of the two paths can be made equal without regard to the input power change. Thus, amplitude tuning for matching is not necessary, and leakage cancellation is achieved by phase tuning alone. The phase shifter is used for the cancellation of carriers by providing a phase difference of 180° . The carrier component is eliminated when the phase difference of the two paths becomes 180° .

The operation of the proposed carrier leakage suppression RF front-end was simulated using Agilent ADS. First, the backscattered signal was generated using the tag emulator when it was exposed to a 910 MHz CW signal of 30 dBm. The tag emulator generates the standard Class-1 FM0 40 kbps signal. The backscattered power was set to -80 dBm considering a 20 m distance. The CW leakage carrier signal is then superimposed. The power level of the carrier leakage was set to -10 dBm, assuming the transmitting power of the reader to be 30 dBm and the isolation between transmitter and receiver to be 40 dB. Then, the carrier to data ratio is found to be about 70 dBc. The synthesized signal was then converted to data using VSA 89600, which can be used as an envelope simulation source in ADS. From Fig. 2(a), the carrier to data ratio is found to be about 34 dBc, and the improvement in carrier to data ratio is estimated to be about 36 dB. Thus the improvement in reader sensitivity can be expected from the improvement in carrier to data ratio. Fig. 2(b) shows the output power of the proposed carrier leakage suppression front-end for changes in the return loss of antenna. The return loss of the antenna is varied from 5 to 45 dB and the circulator isolation of 40 dB. The output power is observed to be almost constant.

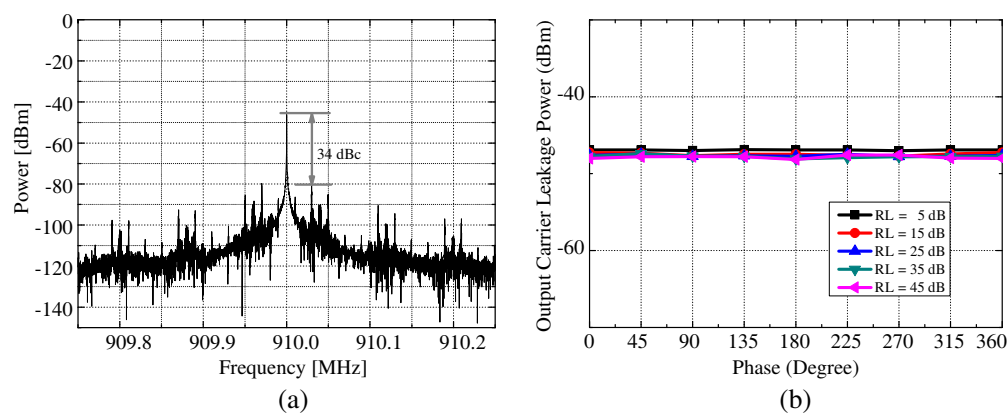


Figure 2: Simulation results at the output of the proposed carrier leakage suppression front-end: (a) simulated spectrum and (b) carrier leakage power.

3. EXPERIMENTAL RESULTS

Figure 3 shows the measurement setup for the proposed carrier leakage suppression front-end. The Tx module was set to transmit a CW power level of 1W. The tag emulator can then generate the backscattered signal compliant UHF RFID standard. The signal generated by the tag emulator is attenuated to reflect the path loss.

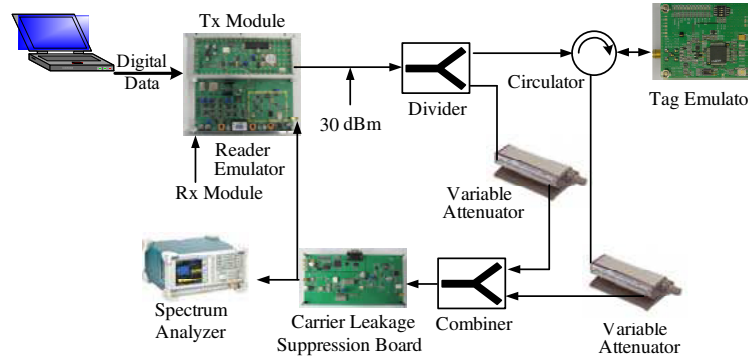


Figure 3: Measurement setup of the carrier leakage suppression front-end.

The power level from the tag emulator was set to -80 dBm which corresponds to 20 m distance. The attenuated tag signal is combined with the TX signal, which represents the leakage power of the transmitter. This signal is sent to the read emulator with the proposed carrier leakage suppression front-end. The measured spectrum of the carrier leakage to the receiver input of the reader emulator is shown in Fig. 4, with and without the proposed carrier leakage suppression front-end. The carrier leakage power was set to -10 dBm and the tag power of the 40 kbps signal was set to -80 dBm to reflect the path loss of 20 m distance using the variable attenuators in Fig. 3. The carrier to data ratio, without the carrier leakage suppression front-end, is found to be about 73 dBc as shown in Fig. 4(a). With the carrier leakage suppression front-end, the carrier to data ratio is observed to be about 37 dBc from Fig. 4(b), which is about a 36 dB improvement.

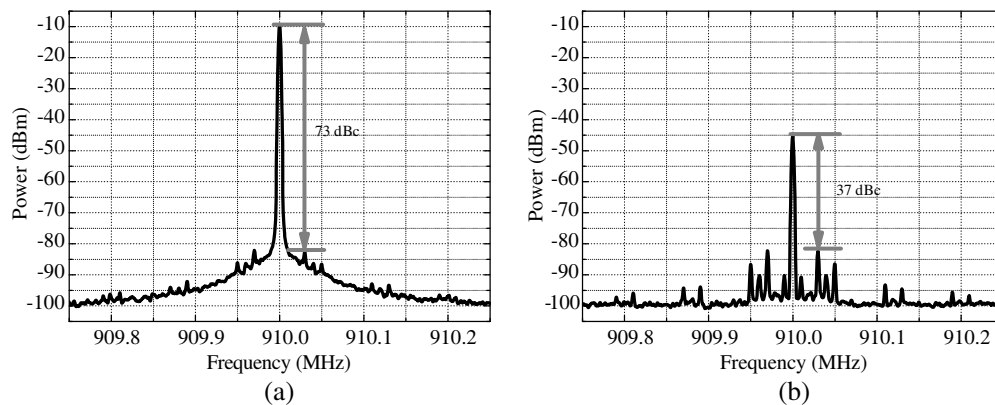


Figure 4: Measured spectra at the receiver input of the reader: (a) without and (b) with the proposed carrier leakage suppression front-end.

4. CONCLUSIONS

We proposed the novel carrier leakage suppression front-end based on received signal alone. The leakage suppression is thus less dependent on time and environmental changes. We demonstrated experimentally that the proposed carrier leakage suppression front-end shows a significant improvement in the carrier to data ratio by up to more than 36 dBc, which is consistent with our simulation results. Also, we demonstrated through indoor experiment that the reader with the proposed carrier leakage suppression front-end can detect a commercial Class-1 battery-assisted tag at a distance of above 20 m.

ACKNOWLEDGMENT

This work was supported by ETRI R & D Program [13ZC1100, The Development of an Intelligent Situation Cognition and an IoT Basic Technology] funded By the Government of Korea.

REFERENCES

1. Finkenzeller, K., *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*, 2nd Edition, 1–59, John Wiley & Sons, 2003.
2. Want, R., “An introduction to RFID technology,” *IEEE Pervasive Comput.*, Vol. 5, No. 1, 25–33, Jan. 2006.
3. “Electronic product code (EPC) standard, radio-frequency identity protocols Class-1 Generation-2 UHF RfId protocol for communications at 860 MHz–960 MHz for communications Version 1.1.0,” 2006.
4. Paret, D., *RFID at Ultra and Super High Frequencies: Theory and Application*, 281–287, John Wiley & Sons, 2009.
5. Jung, J.-W., H.-H. Roh, J.-C. Kim, H.-G. Kwak, M. S. Jeong, and J.-S. Park, “Adaptive TRX isolation scheme by using Tx leakage canceller at variable frequency,” *Microw. Optical Tech. Lett.*, Vol. 50, No. 8, 2043–2045, Aug. 2008.
6. Xiong, T., X. Tan, J. Xi, and H. Min, “High TX-to-RX isolation in UHF RFID using narrow-band leaking carrier canceller,” *IEEE Microw. Wireless Comp. Lett.*, Vol. 20, No. 2, 124–127, Feb. 2010.
7. Villame, D. P. and J. S. Marciano, “Carrier suppression locked loop mechanism for UHF RFID readers,” *2010 IEEE International Conference on RFID*, 141–145, Apr. 2010.

Simulation and Implementation of TDOA Monitoring System for Broadcasting Interferences

Yao-Tang Chang

Department of Information Technology, Kao Yuan University, Kaohsiung 82151, Taiwan

Abstract— Since the rapid development of the wireless broadband communication technology the location accuracy performance of the radio monitoring station configured by signal angle of arrival (AOA) location technology has been seriously degraded in Taiwan. In this study, by applying F(50, 50) transmission channel model of applied in Federal Communications Commission (FCC) and following specifications and requirement of TDOA-based receiver, the required number of TDOA-based in metropolitan area of Tainan is simulated and evaluated. The results show that the number of 3 to 5 TDOA-based location stations are needed to install for 20 km ~ 30 km coverage area in metropolitan of Tainan. Furthermore, three TDOA-based radio monitoring station called Luzhu, Tainan Gaote and Jinkang are first installed to locate the radio transmitter such as interference sources in Taiwan. By monitoring the frequency modulation (FM), the broadcasting station of 88.3, 91.5, 89.1 and 91.9 MHz are located on metropolitan, rural and urban area, respectively. The TDOA-based locating technology is first implemented and be successful. Here, the results show that the locating accuracy at circle error probability (CEP) 50% is less than 950 m distance under multi-path effect in metropolitan area.

1. INTRODUCTION

In order to execute telecommunication act in Taiwan, the radio frequency monitoring system is applied to protect and maintain a great radio frequency environment for easy communication at anywhere, anytime and any-devices. Since the rapid developing of economy in Taiwan in recently years, a large numbers of constructions have made the changes on the terrain and environmental surrounding of the radio monitoring stations rapidly. Hence, the accuracy performance of the radio monitoring station configured by signal angle of arrival (AOA) location technology has been seriously degraded.

Furthermore, more complex-communications technology and equipment has been invented. Hence, the monitoring station that configured originally and based on signal angle of arrival (angle of arrival, AOA) location technology is unable to meet the changing of geography and radio-frequency environment [1]. That is, the accuracy performance of the radio monitoring station has been seriously degraded.

The special interference case increased rapidly in advanced communicational technologies such as low transmitted power, burst and weak signal and metropolitan interference [2, 3]. As new monitoring and direction finder technologies are invented to prevent the radio interference, the TDOA-based location system is proposed to integrate the previous AOA-based location system [4]. Here the implementation and experiments in Taiwan are supported by the radio frequency interference. Hence the interference transmitter for low transmitted power, burst and weak signal and metropolitan interference is easy identified, analyzed (located) and prevented.

2. THE DESIGN OF TDOA-BASED MONITORING SYSTEM

In order to implement TDOA monitoring system, the design configuration consists of hardware selection, the cross-correlation algorithm and application user interface and estimation of accuracy location. Time difference of arrival (TDOA) is the use of a plurality on receivers to measure the arrival time from the location of pending signal to each receiver.

To obtain an accuracy solution, the TDOA technology hardware and software design of the experiment have to be based on the performance of the receiver, time synchronization and network connectivity. The installation of system architecture and the location requirements is shown in Fig. 1. The important TDOA-based monitoring is that the implementation of signal synchronization. Here, the TDOA location systems typically use a GPS receiver to establish a common time reference. In this study, the experimental monitoring station is configured as three layers to install and analyze the location accuracy including hardware selection, software design (cross-correlation algorithm) and estimation of location error.

Under the cost consideration, the proposed integrated AOA/TDOA receiver configuration is proposed and shown in Fig. 2. The AOA-based technology needs a pair of modular of local oscillator,

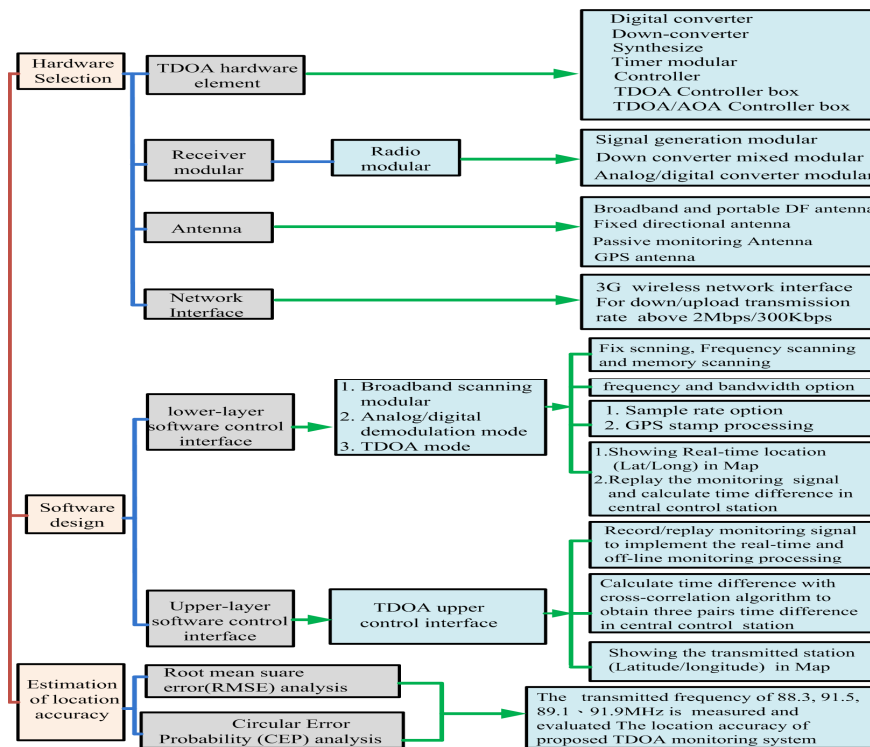


Figure 1: The installation of proposed system architecture and the location requirements.

down converter and digitizer. The only one modular is need for TDOA-based technology. Here, the integrated AOA/TDOA is easy implemented while the AOA or TDOA software algorithm is applied.

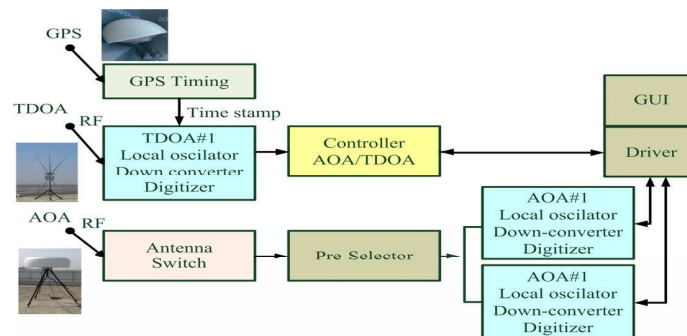


Figure 2: The proposed receiver configuration of integrated AOA/TDOA location system.

3. SIMULATION AND IMPLEMENTATION OF TDOA-BASED MONITORING SYSTEM

By applying F(50, 50) transmission channel model that applied in Federal Communications Commission (FCC) and followed with specifications and requirement of TDOA-based receiver, the required number of TDOA-based in metropolitan area of Tainan is simulated and evaluated. The simulation parameter and result is summed in Table 1 and the required number of TDOA-based in metropolitan area of Tainan is simulated and evaluated.

The results show that the number of 3 to 5 TDOA-based location stations are needed to install for 20 km ~ 30 km coverage area in metropolitan of Tainan. Hence, the proper position TDOA-based monitoring station is selected and shown in triangle of Fig. 3. Moreover, as shown the broadcasting station of 88.3 MHz, 91.5 MHz, 89.1 MHz and 91.9 MHz are specified to verify the location performance by moving from the center to outside of TDOA coverage area.

In this trial experiment, the three monitoring station including Jinkang station in Tainan (Tainan metropolitan area) and Tainan Gaote is selected to create the proper triangular arrangement. The

Table 1: The required number of TDOA-based foe applied simulation parameter.

Simulated monitoring station	frequency	Antenna factor (AF)	Coverage area (km)	Receiver sensitivity	Antenna height (transmitter)	Antenna height (receiver)	The number for TDOA
Kaoshiung Science Park (central control station (Lujhu city))	30 MHz	-4.2 dB	Simulated area from north Kaoshiung to south Tainan	≤ -100 dBm	2m	20m	Simulated three monitoring sites
	300 MHz	15.8 dB	15				
	1GHz	26.3 dB	15				
	3GHz	35.8 dB	8				

central station is installed in Luzhu (Kaohsiung Science Park). At the three sites, a distance of about 9.94 km is calculated between Jinkang Station and Luzhu (Kaohsiung Science Park).

As shown in Fig. 3, the proposed experiment configuration is easy to verify the TDOA limitation of location performance in multi-path effect of metropolitan area.

In this study, a central station of AOA/TDOA #1 is installed in Luzhu (Kaohsiung Science Park). The TDOA #2 and TDOA #3 (Tainan metropolitan area) is act as remote monitoring station. This monitoring station is selected under considering the barrier impact of the aforementioned buildings, ground reflection and the reflected wave. Also, the up-load transmission rate of 3G mobile network design is needed. The related position (longitude and latitude) of TDOA monitoring and transmitted station is shown in Table 2. For an example of measured FM 88.3 MHz, the location result is shown in Fig. 4.



Figure 3: The proposed experiment configuration for investigating the TDOA limitation of location performance in multi-path effect of metropolitan area.

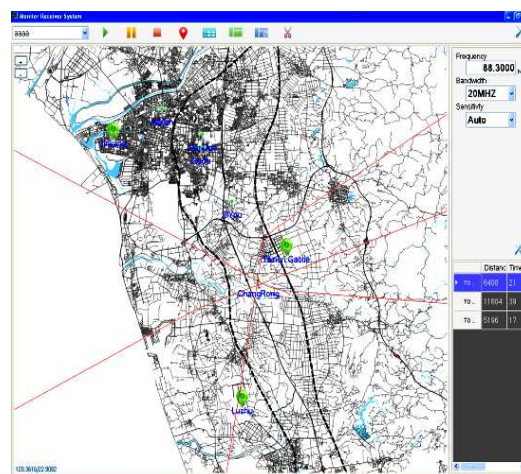


Figure 4: The location result of an example of FM 88.3 MHz in experiment.

The experiment results are summarized in Table 3. The various location error depend on the various broadcasting station of 88.3, 91.5, 89.1 and 91.9 MHz. The result show the least locating accuracy for 91.9 MHz at circle error probability (CEP) 50% is less than 950 m distance under multi-path effect. Hence, the average and deviation error of distance is proved to apply in metropolitan interference.

4. THE PROPOSED INTEGRATED AOA/TDOA-BASED MONITORING SYSTEM

By integrating advantage of the AOA and TDOA-based technologies, the AOA/TDOA monitoring system is proposed to improve and the capacity of monitoring area and strengthen the ability of anti-interference. As shown in Fig. 5, the fixed AOA/TDOA monitoring station is installed in higher mountain surrounding the metropolitan area. However, mobile TDOA monitoring station is assigned on vehicle to move randomly in metropolitan area. Here, the AOA-based system can

Table 2: The proposed monitoring and transmitted position of proposed experiment.

Proposed monitoring and transmitted station	Longitude	Latitude
TDOA #1 monitoring	120° 15' 42.41"	22° 50' 21.70"
TDOA #2 monitoring	120° 17' 9.00"	22° 55' 29.00"
TDOA #3 monitoring	120° 10' 14.60"	22° 59' 49.70"
88.3 MHz broadcasting station	120° 16' 26.36"	22° 54' 15.15"
91.5 MHz broadcasting station	120° 13' 45"	22° 59' 57"
89.1 MHz broadcasting station	120° 13' 45"	22° 59' 57"
91.9 MHz broadcasting station	120° 12' 33"	23° 0' 45"

Table 3: The summarized result of experiment for measured location.

Center frequency of Transmitted Station (MHz)	Precise performance of location @CEP50% (m)	Average distance error of measured location (m)	Deviation distance error of measured location (m)
88.3	235	230.31	22.24
91.5	230	260.50	76.97
89.1	550	515.22	137.25
91.9	950	929.26	275.76

calculate the straight line of bearing and achievement of TDOA-based hyperbola curve. Hence, the proposed efficiency configuration of integrated AOA/TDAO monitoring system is achieved to prevent broadcasting interferences occurred in metropolitan area.

Following the integrated AOA/TDOA configuration in Taiwan, the interference transmitter for flight, low transmitted power, burst and weak signal and metropolitan interference is easy identified, analyzed (located) and prevented. Moreover, as shown in Fig. 6, the major advantage of integrated AOA/TDOA monitoring system is to extend the coverage area and improve the location performance of previous AOA-based monitoring system installed in Taiwan. For location accuracy of TDOA consideration, the surrounding of three monition stations are characterized with better location-performance. On the contrary, AOA monitoring system is suitable for the inside as well as outside of monitoring system. Hence, the possible architecture, equipment specification of proposed TDOA location system is proposed to solve the special interference and improve the monitoring network.

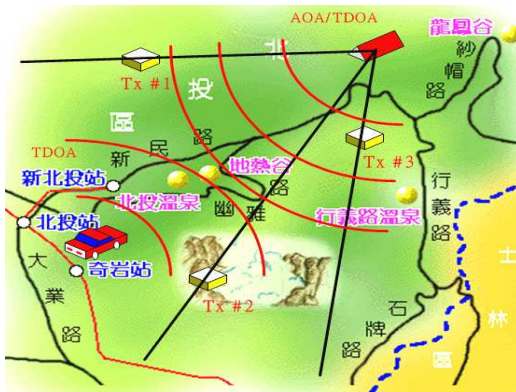


Figure 5: The proposed efficiency configuration of integrated AOA/TDAO monitoring system is used to prevent broadcasting interferences occurred in metropolitan area.

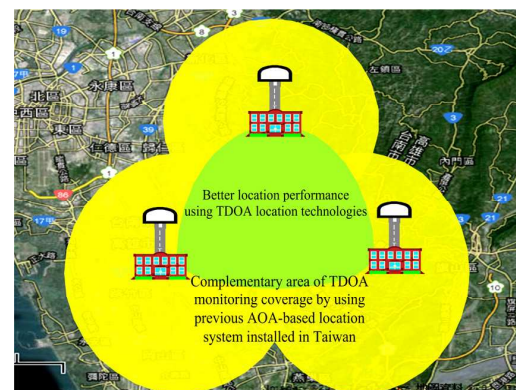


Figure 6: Integrated AOA/TDOA monitoring system for extending the coverage area and improve location accuracy.

5. CONCLUSION

By applying F(50, 50) transmission channel model of applied in Federal Communications Commission (FCC) and following specifications/requirement of TDOA-based receiver, the required number of TDOA-based in metropolitan area of Taiwan is simulated and evaluated. Since the number of 3 to 5 TDOA-based location stations are needed to install for 20 km ~ 30 km coverage area in metropolitan of Tainan. The results show that the locating accuracy at circle error probability (CEP) 50% is less than 950m distance under multi-path effect in metropolitan area. Hence, the solution of integrated the AOA/TDOA technologies can be applied and the interference transmitter characterized with low transmitted power, burst and weak signal and metropolitan interference is easy identified, analyzed (located) and prevented.

ACKNOWLEDGMENT

This study was supported under grant No. NSC 102-2221-E-244-001 from the National Science Council and the special-interference project of National Communications Commission (NCC) in Taiwan.

REFERENCES

1. ITU, *Spectrum Monitoring Handbook*, 2011.
2. Liang, Y. C., A. R. Leyman, and B. H. Song, "Multipath time delay estimation using higher order statistics," *Higher-order Statistics, Proceedings of the IEEE Signal Processing*, 1997.
3. Nikias, C. L. and R. Pan, "Time delay estimation in unknown Gaussian spatially estimation for large BT signals," *IEEE Transactions on Signal Processing*, Vol. 39, No. 4, Apr. 1991.
4. Li, W., "Hybrid TDOA/AOA mobile user location for wideband CDMA cellular systems," *IEEE Transactions on Wireless Communications*, Vol. 1, No. 3, Jul. 2002.

Design Guidelines of a Microwave Active Front-end Receptor: A LNA Approach

J. C. Bonoto¹, C. E. Capovilla², and H. X. Araujo¹

¹Universidade Federal de São João Del Rei — UFSJ, Ouro Branco, Brazil

²Universidade Federal do ABC — UFABC, Santo André, Brazil

Abstract— In this work, design guidelines of the first stage of a CMOS 0.35 μm front-end receptor, to attend the Brazilian 4G mobile technology is shown. This paper deals with different topologies of LNA — Low Noise Amplifier and guidelines of design, layout, and simulations. Thereby, the development of the proposed device is shown since the hand calculations to the final layout which is ready to be run by the Foundry. In this context, this paper aims the improvement and development of new technologies to meet the growing demands on the communication and mobile field.

1. INTRODUCTION

The current economic scenario, coupled with the lifestyle that people have in their homes, always seeking comfort and connectivity, makes increase the need for faster electronic devices with high efficiency and low cost. These devices have features to make life easier and give comfort to people. Among these features, can be cited the use of the internet as a vehicle of communication by video conference through mobiles, something unimaginable in the early twenty-first century.

In order for such expectations becomes fulfilled, it took many technological advances in the electronic field, including the stage of RF (radio frequency) devices for communications, which is considered the basis for many advances, due to the growing demand for data traffic, voice and image, which increases the need of high speeds transfers that support these services, making it essential that there be a monitoring of its technological development.

The progress related to the RF microelectronic happened in a faster way [1, 2], thus requiring constant updating and training of researchers and technicians involved in this area. This search for knowledge is due to the fact that the market always needs new devices.

The LNA is the first stage in the signal reception, as can be seen in Fig. 1, so it must be well designed, because if the amplified signal provides high levels of noise all other information will be compromised [3, 4]. Following the amplification stage the signal will pass through several intermediate stages until it can be extracted the original information. On the other hand, the mixer translates an incoming RF signal to a lower frequency, known as the intermediate frequency — IF [5, 6], while the local oscillator is applied to generate signal on a desire frequency. Several topologies are deal on the literature, like the Voltage Controlled Oscillator — VCO, *Colpitts*, ring oscillator and others [7, 8]. The power amplifiers may be divided into several categories, depending whether it is a broadband or narrowband. Relative to its linearity, it can be linear or constant-envelope operation, which generally produces an output whose amplitude is ideally independent of the input, unlike the linear ones, known for produce an output identical to the original input [9].

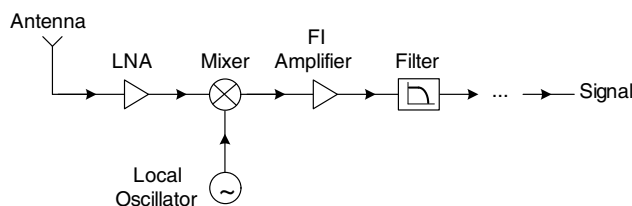


Figure 1: A generic front-end receptor [10].

The 4G technology is still in its early stages of use in Brazil, since the required parameters for the existence of a difference between 3G and 4G technologies have not been fully achieved. From these parameters can be mentioned the download rate, which should be between 100 Mbps to 1 Gbps but is still operating in the range of 5 Mbps. With these divergences, the technology specifications were revised, so that one could use the 4G nomenclature even though this is not within the preset

parameters. In the United States the 4G operational frequency is in the range of 700 MHz, while in Brazil the 4G devices work at 2.5 GHz.

Therefore, in this work it will be shown design guidelines of a 0.35 μm CMOS first stage front-end, LNA for the Brazilian 4G technology. It will be shown since the hand calculation to the final layout which will be manufactured by the AMS — Austria Micro System Foundry.

2. FRONT-END FIRST STAGE

The low noise amplifier is usually used as the first stage of signal reception. For being a device with good gain and low noise, it plays a fundamental role in the circuit and has a high impact on the continuity of the RF signal. The LNA can normally be configured in four different ways to realize the impedance matching. However, each configuration present disadvantages on the system as a whole.

The first architecture employs a resistive termination to match the impedance at 50 Ω . The disadvantage of this technique is the use of a resistor for the impedance matching which has a very large contribution in the signal to noise ratio, greatly increasing the level of noise in the received signal.

The second topology uses a common gate amplifier for the LNA. For this configuration we can mention two distinct problems: noise figure, which in theory has a minimum value of 2.2 dB for the CMOS technology, besides the problem of signal linearity which for this setting is too low. Therefore, a commitment between power and gain must be assumed.

In the third configuration is shown a gain stage with resistive feedback. It is insensitive to the parasitic elements of the input match network. However, this topology requires a large transconductance and thus a large current, making this project with high power consumption.

The fourth architecture employs an inductive degeneration, to generate a real term on the input impedance. Such topology uses an inductor connected to the source to provide the impedance match at the input without a lumped resistor. In general, this type of configuration has the characteristic of being narrow-band which does not compromise the application of this the configuration for this work.

3. LNA DESIGN GUIDELINES

The cascode topology increases the LNA bandwidth due to the Miller effect which modifies the effective capacitance of the circuit, so that this condition contributes to the linearity of the circuit. The common gate amplifier has two important roles in the reverse isolation of the LNA which decreases the leakage current produced by the oscillator and the mixer and minimizes the feedback of the output to the input. One of the main problems in designing a cascode LNA is the calculation of the channel width of the transistors.

The elimination of the noise in a complete way is still the major goal of any project within the area of transmission and reception of signals. However, as the current technology does not provide this capability, that there are means to reduce the noise by more efficient circuitries of transmitting and receiving data. The fact that there is noise at all stages of amplification leads to the conception that the gain stage at the input must find a relationship where noise must be mitigated as much as possible without losing the minimum of amplification. This relationship is necessary so that the signal can pass through all the stages of amplification without losing its quality with noise that were not added yet.

To design an active microwave device, some information must be obtained with the Foundry. Therefore, the hand calculations were done taking in account the typical configuration of the parameters provided by the AMS Foundry. The first expression is concerning to the calculation of the channel width — W of the CMOS transistor. Thus, the W can be found by:

$$W = \frac{1}{3\omega LC_{ox}R_s} \quad (1)$$

where L is the length of the transistor's channel, R_s the input resistance, 50 Ω , and ω the angular frequency

To attend the 4G Brazilian mobile specification [11], with central frequency around 2.6 GHz and from the process parameters, the transistors channel width $W \approx 256 \mu\text{m}$ was found. As the value of the NMOS RF type transistors is limited in terms of channel width, the use of 12.5 fingers was needed, making a width around 125 μm . To achieve $W \approx 256 \mu\text{m}$, it was placed another 125 μm

transistor in parallel. The capacitance between the gate and source can be found by the following expression:

$$C_{gs} = \frac{2}{3}C_{ox}WL + C_{gs0}W \quad (2)$$

And the existing conductance inside the transistor is expressed by

$$G_{m1} = K_{pn} \frac{W}{L} (V_{gs} - V_{th}) \quad (3)$$

As the inductors are pre-formed cells provide by the Foundry, with established values, the transistor width should be design in function of these existing cells. Therefore, the impedance match can be achieved by the following expressions:

$$Z_{in}(s) = s(L_s + L_g) + \frac{1}{sC_{gs}} + \left(\frac{G_{m1}}{C_{gs}} \right) L_s \quad (4)$$

where L_s and L_g are inductances connected to the source and gate, respectively.

On the other hand, the drain current can be obtained by the following expressions:

$$I_d = \frac{K_{PN} W}{2} \frac{W}{L} (V_{gs} - V_{th})^2 \quad (5)$$

From the design equations showed above, with a $V_{DD} = 3.3\text{V}$ and $V_{gs} = 1\text{V}$, a drain current around 15 mA is found.

In a second analysis, using the inductive values from (4), it is possible define (6):

$$\omega_0(L_g + L_s) - \frac{1}{\omega_0 C_{gs}} = 0 \quad (6)$$

Replacing the obtained process parameters, and using the pre-existing cells, the relationship between L_g and L_s is: $L_g = 12.08\text{ nH}$ and $L_s = 1.04\text{ nH}$. Note that the L_g inductor must assume values around 1 nH. In Fig. 2(a) is shown the gain S_{21} and reverse isolation S_{12} of the designed LNA. From the obtained results, it can be observed a gain around 15 dB, which validate the good behavior of the device. Moreover, in Fig. 2(b) is shown the return loss S_{11} , where it can be observed

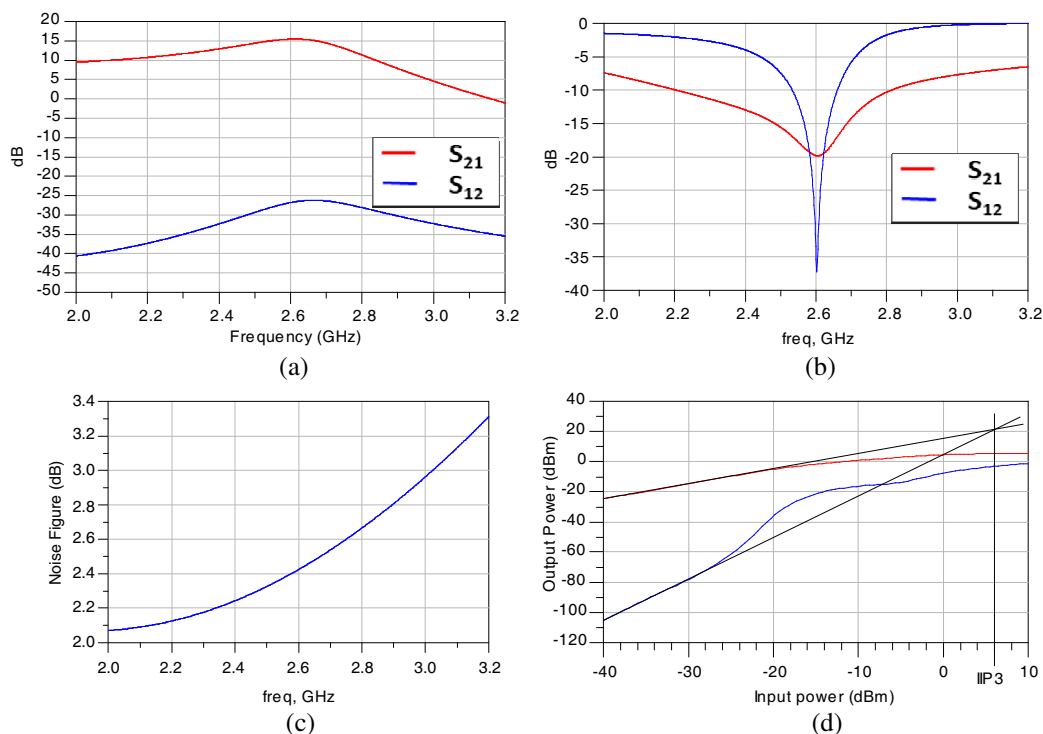


Figure 2: (a) Gain and reverse isolation; (b) Return loss; (c) Noise figure; (d) The linearity response — IP3.

the agreement of the resonance peak with the frequency of design. It is still notice in Fig. 2(c) the noise figure around 2.4 dB at the interest frequency. This is a good value due to the limitation of the used technology.

In addition to the gain, input match and noise figure, the linearity is an important consideration because the LNA usually performs more than only amplify signals without adding noise, remaining linear besides the signals that it receives. Therefore, in Fig. 2(d) is shown the IP3 analysis, which represents the linearity performance of the designed device.

From the obtained results, it can be clearly observed that that designed LNA is working correctly, with good gain, resonance and linearity. In Fig. 3(a) is shown the schematic of the designed device, while in Fig. 3(b) is shown its layout. In this schematic the bias circuit, RF PADS, wirebonds, and output matching network was intentionally neglected to facilitate the visualization of the proposed circuit. The presence of four layers of metal is highlighted, and is important to observe that the fourth metal layer is thick to guarantee a moderate Q of the integrated inductors. Countless tests are required to check the layout before it can be submitted to Foundry.

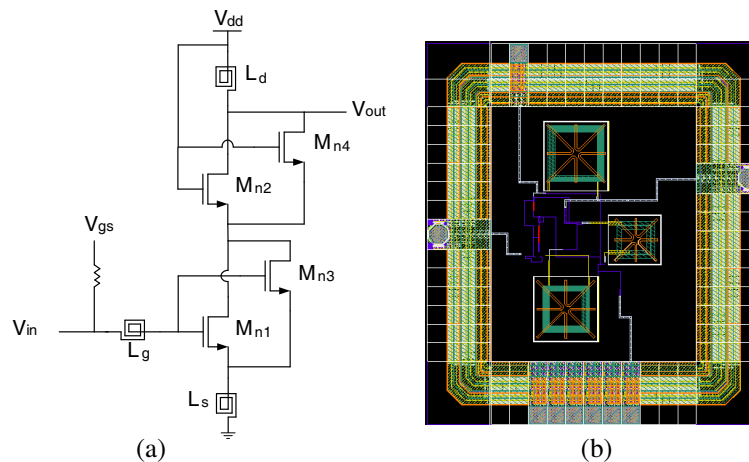


Figure 3: (a) Schematic; (b) Layout of the Front-end first stage: Low noise amplifier.

4. CONCLUSIONS

In this work, design guidelines of a LNA to attend the Brazilian 4G mobile technology were presented. It was shown the development of the proposed device since the design by hand calculations to the final layout, which is ready to be run by the Foundry. From the obtained results, it can be clearly observed that the designed LNA is working correctly, with good gain, tuning, noise figure and linearity.

ACKNOWLEDGMENT

The authors would like to thank FAPEMIG — Minas Gerais Research Foundation, for the partial financial support.

REFERENCES

1. Li, R., C. Jin, S. C. Ong, T. G. Lim, K. F. Chang, and S. W. Ho, "Embedded wafer level packaging for 77-GHz automotive radar front-end with through silicon via and its 3-D integration," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2013.
2. Thabet, H., S. Meillere, M. Masmoudi, J. Seguin, H. Barthelemy, and K. Aguir, "Combiners based on CMOS inverters and application in RF transmitter for wireless sensors," *7th International Conference on Design & Technology of Integrated Systems in Nanoscale Era*, 2012.
3. Capovilla, C. E., A. S. A. Tavora, and L. C. Kretly, "A fully integrated 2.5 GHz band CMOS Low Noise Amplifier with multiple switched inputs for diversity wireless communications," *SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference — IMOC 2007*, 2007.
4. Kretly, L. C., C. E. Capovilla, and A. A. S. Tavora, "A 1.9-GHz CMOS low noise amplifier with partial source degeneration," *SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference (IMOC)*, 2009.

5. Na, D. and T. W. Kim, "A 1.2 V, 0.87–3.7 GHz wideband low-noise mixer using a current mirror for multiband application," *IEEE Microwave and Wireless Components Letters*, 2012.
6. Pozar, D. M., *Microwave Engineering*, 3rd Edition, Wiley, 2004.
7. Wang, J. and Z. Liu, "A merged up-conversion mixer-VCO based on current reuse," *IEEE Subthreshold Microelectronics Conference*, 2012.
8. Razavi, B., "CMOS technology characterization for analog and RF design," *IEEE Journal of Solid-State Circuits*, 268–276, 2002.
9. Werker, S., C. Stedler, and R. Kronberger, "Highly linear low noise amplifier for 2.45 GHz," *IEEE Microwave Magazine*, 146–151, 2012.
10. Lee, T. H., *Design of CMOS Radio-frequency Integrated Circuits*, 2nd Edition, Cambridge University Press, 2004.
11. <http://www.anatel.gov.br/Portal/exibirPortalPaginaEspecialPesquisa.do?acao=&tipoConteudoHtml=1&codNoticia=25255>.

Statistical Analysis of On-body Radio Propagation Channel for Body-centric Wireless Communications

H. A. Rahim¹, F. Malek¹, N. Hisham¹, and M. F. A. Malek²

¹Embedded, Networks and Advanced Computing Research Cluster (ENAC)

School of Computer and Communication Engineering

Universiti Malaysia Perlis, P. O. Box 77, d/a Pejabat Pos Besar, Kangar 01000, Malaysia

²MISC LNG Japan MISC Berhad, Yokohama, Japan

Abstract— This paper presents the experimental investigation of on-body radio propagation channel utilizing textile monopole antenna at 2.45 GHz. The measurement campaign was carried out in the anechoic chamber, considering stationary movement of the human body. The position of the transmitted antenna was fixed on the right of the upper arm of the human body and the received antenna was varied on several different potential on-body locations for body-centric wireless communication (BCWC) applications. The investigation was aimed to characterize the reflection coefficient and path loss of on-body radio channel when the antenna was placed in the vicinity of human body. A statistical analysis of path loss was also performed. The results showed that the measured reflection coefficient of four on-body positions experienced an upward frequency shift at a minimum of 0.2% compared to the simulated results due to the body coupling effect. In the anechoic chamber, the highest path loss was found for right upper arm-left ankle link while the lowest path loss was observed on the right upper arm-right chest link, proving that the closer a received antenna to the transmitted antenna, the better signal reception will be obtained. Based on the measurement results, it could be seen that the lognormal distribution fits very well to the on-body radio channel for narrowband frequency.

1. INTRODUCTION

BCWC has received a lot of attention recently [1–4]. Since BCWC is intended to be implemented on the user's body, it is more practical to utilize a textile antenna with an omni-directional pattern in BCWC applications as it can be integrated into clothing. On-body radio propagation channel has been extensively published in the open literature [1–4]. However, very few work reported the utilization of textile antenna for on-body radio channel [3]. Thus, this paper presents the experimental investigation of reflection coefficient and path loss characteristics using planar textile monopole antenna at 2.45 GHz. The stationary on-body propagation channel for different body positions is derived and statistically analyzed.

2. MEASUREMENT SETUP

The experiment was carried out in an anechoic chamber at Electromagnetic Hyper Sensitivity (EHS) Laboratory to eliminate multipath reflections from surrounding environment. The antenna utilizes a portable Agilent Field Fox model number N9923A 2-port Vector Network Analyzer (VNA) to generate (transmit) and measure (receive) the signal. A total number of sampled points per acquisition $N = 1001$ is set. Measurement was performed on a female subject of weight 51 kg with a height of 1.49 m. Two planar textile monopole antennas were used in this measurement campaign [5, 6]. The transmitter antenna (Tx) was placed fixed at the right side of the upper arm (RU). The receiver (Rx) was placed on the 11 other positions: right chest (RC), left chest (LC), right waist (RW), left waist (LW), right thigh (RT), left thigh (LT), right ankle (RA), left ankle (LA), center of back (B), right back (RB), left back (LB) and left upper arm (LU). Fig. 1(a) shows the location of the Tx and positions of Rx antennas. Two 5 m low loss semi-rigid coaxial cables were used in the measurement campaign. The cables were wrapped with Eccosorb Flexible Broadband Urethane Absorber model: FGM-U-20-SA microwave absorbing foams to minimize the spurious radiation from, and coupling between, the coaxial cables. The measurement setup for on body shows in Fig. 1(b). A 10 mm separation was set between the antenna and the body. Table 1 shows the distance between the Tx-Rx for on-body measurement. Five sweep durations were performed for each location. The simulation was performed using CST HUGO body model software where the HUGO model was defined at $8 \times 8 \times 8 \times 8 \text{ mm}^3$ voxel resolution.

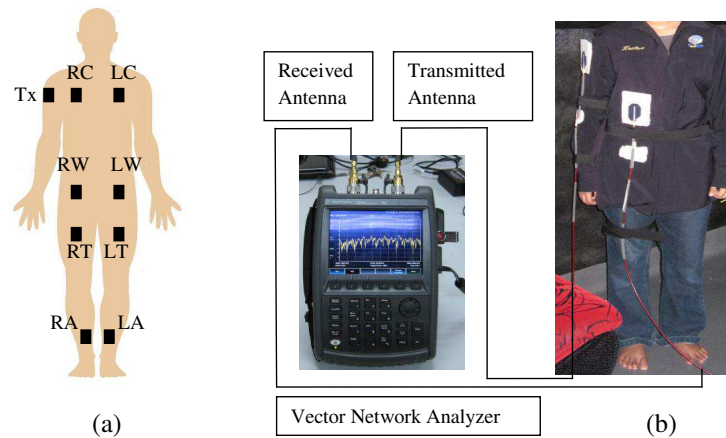


Figure 1: On-body measurement setup. (a) Placements of transmitted and received antennas. (b) Measurement in anechoic chamber.

Table 1: Distance between Tx-Rx for on-body measurement.

Positions	Right Chest (RC)	Left Chest (LC)	Right Waist (RW)	Left Waist (LW)	Right Thigh (RT)	Left Thigh (LT)	Right Ankle (RA)	Left Ankle (LA)	Center Back (B)	Right Back (RB)	Left Back (LB)	Left Upper Arm (LU)
Distance (cm)	19	35	22	38	27	40	101	107	52	27	40	52

3. RESULTS AND ANALYSIS

The measured reflection coefficient for on-body static at all positions is illustrated in Fig. 2. It is observed that the measured reflection coefficient of all on-body Rx placements shifted to the right up to 7.2% due to the body coupling effect. This result also shows that the textile monopole demonstrated reflection coefficient, $S_{11} < -10$ dB for all on-body locations. Voltage Standing Wave Ratio (VSWR) is a function of the reflection coefficient, which describes the power reflected from the antenna. Fig. 3 shows the VSWR for on-body static for free space and nine positions of textile monopole. From the graph, it is seen that VSWR for all positions and free space is less than 2. In general, if the VSWR is less than 2, the antenna matching is considered excellent. Fig. 4 shows the comparison between simulated and measured reflection coefficient for four on-body positions, i.e., LU, LB, RC and RT of textile monopole antenna. The result clearly showed that the minimum frequency detuning occurred when the antenna was placed on the left back by 0.2% as compared to the simulated reflection coefficient result. It is evident that the left back is the least affected

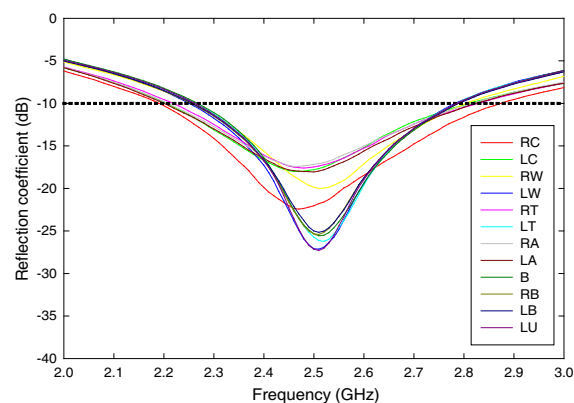


Figure 2: Measured reflection coefficient for several positions of Rx.

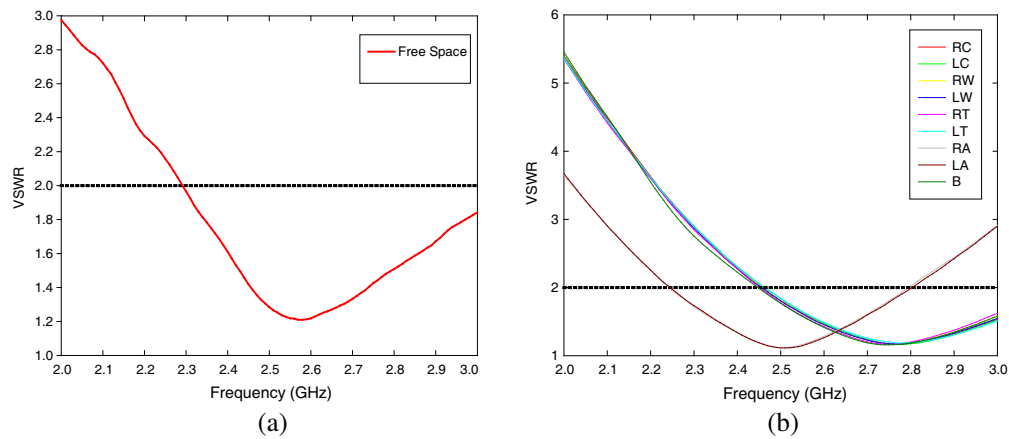


Figure 3: Measured VSWR for stationary on-body in (a) free spaces and (b) positions of Rx.

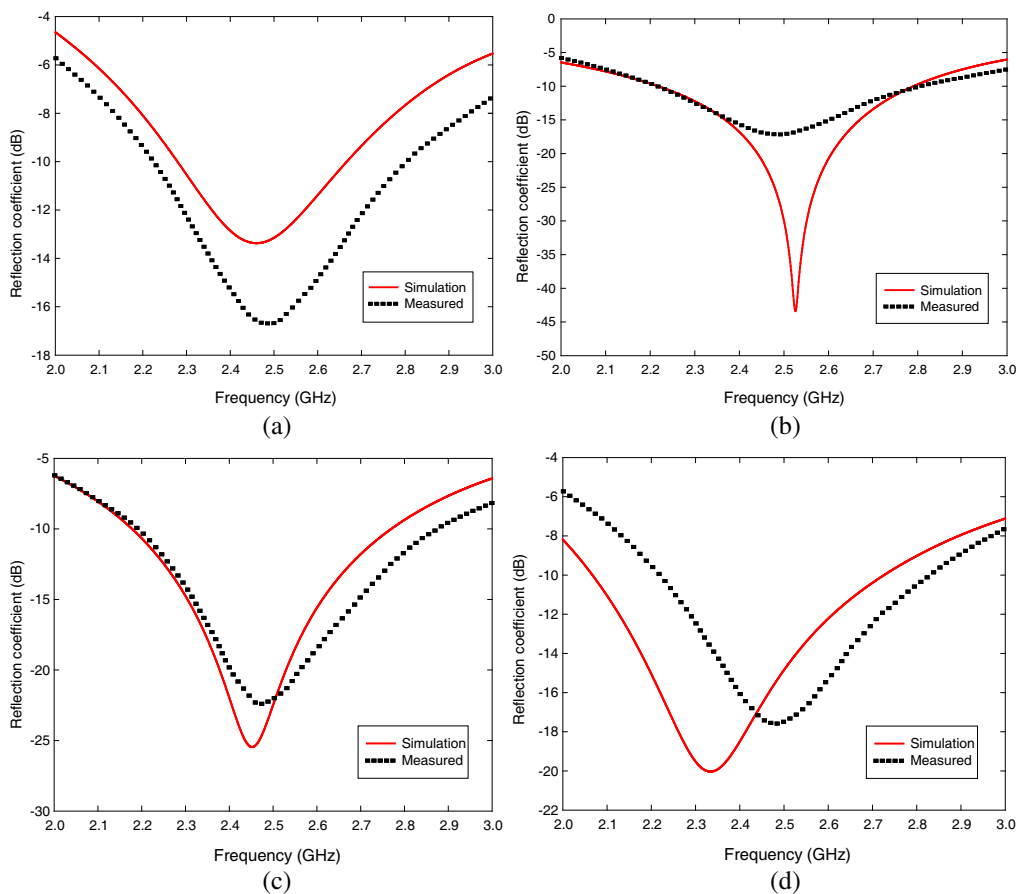


Figure 4: Comparison between (a) left upper arm, (b) left back, (c) right chest, (d) right thigh for simulated and measured reflection coefficient.

location by the body coupling when textile monopole was used as Tx. However, this on-body location is impractical to be applied in BCWC as it will make the user feel uncomfortable if the antenna is to be attached to the clothing. Hence, left upper arm is chosen as Tx position since the frequency detuning is less than 2% compared to the simulation result.

The path loss is defined as the ratio of received to transmitted power computed from the measured data, averaging over the measured frequency transfers at each frequency point [7]. Fig. 5 shows the path loss for 5 locations of stationary on-body. The result shows that highest path loss at 2.45 GHz was obtained for RU-LA link with a maximum value of -53 dB due to the longest distance between Tx and Rx. Meanwhile the lowest path loss was observed on the RU-RC link with a maximum value of -43 dB. Since the propagation distance is shorter between RU and RC

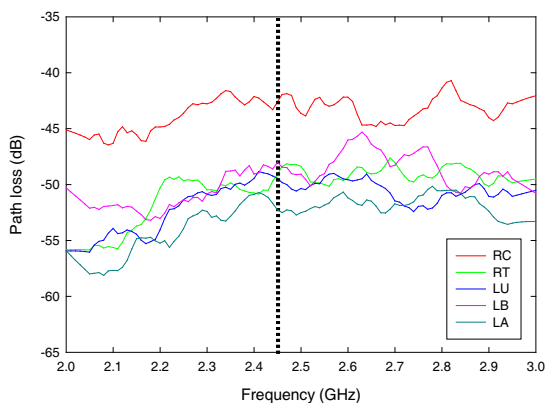


Figure 5: Path loss of five Rx position for stationary on-body propagation.

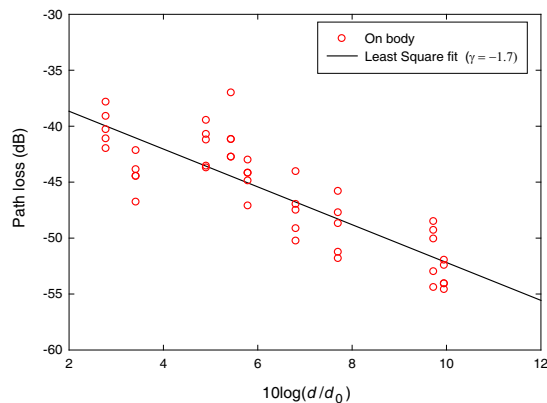


Figure 6: Measured and modeled path loss for stationary on-body channel.

locations, the electromagnetic wave can propagate in a direct path from transmitter to the receiver. The on-body radio channel can be modelled as a linear function of the logarithmic distance d between transmitter and receiver, expressed as [2]

$$PL_{dB}(d) = PL_{dB}(d_0) + 10\gamma \log\left(\frac{d}{d_0}\right) + X_\sigma \quad (1)$$

where $PL_{dB}(d_0)$ is the average path loss at 0.1 m and γ is the path loss exponent. X_σ represents a shadowing (large-scale) fading defined as variation of the local mean around the path loss, Gaussian distributed random variable with standard deviation σ in dB. In order to obtain the average path loss at d_0 and the path loss exponent γ , a least square fit technique is applied. Fig. 6 shows the measured and modeled path loss value for stationary on-body radio propagation channel, involving nine on-body positions, i.e., RC, LC, RW, LW, RT, LT, RA, LA and B. The path loss exponent for this case is $\gamma = -1.7$ and the mean path loss is -32.3 dB. A shadowing factor is determined by computing the deviation between measured and the calculated average path losses. Fig. 7(a) presents the measured CDF of path loss for stationary on-body radio channel in the chamber fitted to normal distribution ($\sigma = 2.6$). This explains that human body shadowing plays insignificant role to the stationary on-body path loss variation when utilizing an omni-directional antenna. A statistical analysis is also performed to the measured path loss by fitting the data to an empirical distribution, lognormal distribution. The measured CDF of stationary on-body path loss is shown in Fig. 7(b). The result exhibits that the measured on-body path loss in the chamber is very well fit to lognormal distribution ($\mu = 3.82$, $\sigma = 0.10$). A smaller spread of data, indicating by $\sigma = 0.10$, shows that there is a direct path of propagation occurred along the body surface.

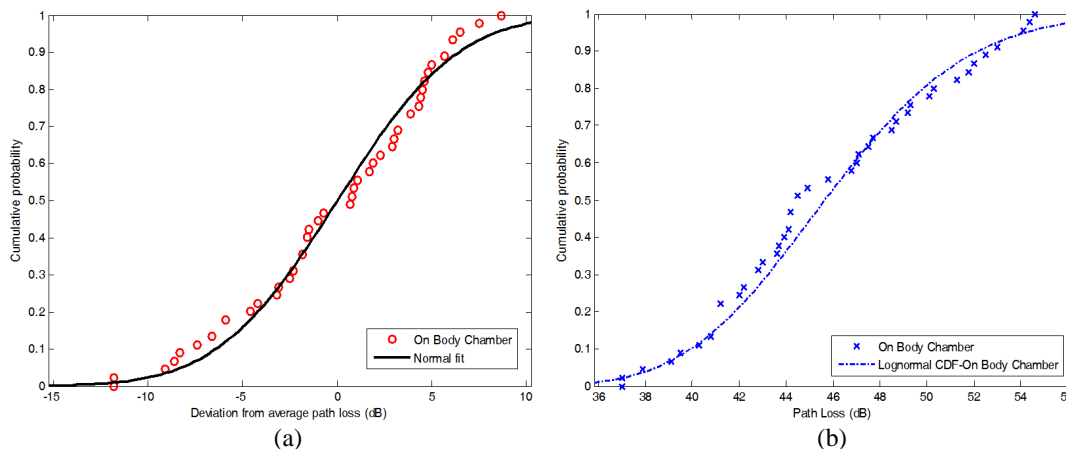


Figure 7: Measured CDF of (a) deviation from average path loss fitted to normal distribution, (b) path loss in the chamber.

4. CONCLUSIONS

The stationary on-body radio channel involving different body positions was carried out in an anechoic chamber. The characteristics of the reflection coefficient and path loss of on-body radio propagation channel were studied. The channel model derivation and statistical analysis of stationary on-body radio channel were also performed. The measured path loss was fitted to an empirical distribution function. Results exhibit that the textile monopole obtained $S_{11} < -10$ dB for all on-body positions and demonstrated an upward frequency shift of four selected on-body locations at a minimum of 0.2% compared to the simulated results. The results also confirm the distance dependency between Tx-Rx of stationary on-body radio channel in non-reflecting environment. The measured path loss was very well fitted to the lognormal distribution.

REFERENCES

1. Cotton, S. L., G. A. Conway, and W. G. Scanlon, "A time-domain approach to the analysis and modeling of on-body propagation characteristics using synchronized measurements at 2.45 GHz," *IEEE Trans. on Antennas and Propagation*, Vol. 57, No. 4, 943–955, Apr. 2009.
2. Sani, A., Y. Zhao, Y. Hao, S.-L. Lee, and G.-Z. Yang, "A subject-specific radio propagation study in wireless body area networks," *2009 Loughborough Antennas and Propagation Conference (LAPC)*, 80–83, Loughborough, UK, Nov. 16–17, 2009.
3. Michalopoulou, A., A. A. Alexandridis, K. Peppas, T. Zervos, F. Lazarakis, K. Dangakis, and D. I. Kaklamani, "On-body channel modelling: Measurement and statistical analysis," *2010 Loughborough Antennas and Propagation Conference (LAPC)*, 201–204, Loughborough, UK, Nov. 8–9, 2010.
4. Abbasi, Q. H., M. M. Khan, S. Liaqat, A. Alomainy, and Y. Hao, "Experimental investigation of ultra wideband diversity techniques for on-body radio communications," *Progress In Electromagnetics Research C*, Vol. 34, 165–181, 2013.
5. Rahim, H. A., F. Malek, I. Adam, S. Ahmad, N. B. Hashim, and P. S. Hall, "Design and simulation of a wearable textile monopole antenna for body centric wireless communications," *PIERS Proceedings*, 1381–1384, Moscow, Russia, Aug. 19–23, 2012.
6. Rahim, H. A., F. Malek, I. Adam, S. Ahmad, N. B. Hashim, and P. S. Hall, "On-body textile monopole antenna characterisation for body-centric wireless communications," *PIERS Proceedings*, 1377–1380, Moscow, Russia, Aug. 19–23, 2012.
7. Dabin, J. A., N. Ni, M. Haimovich, E. Niver, and H. Grebel, "The effects of antenna directivity on path loss and multipath propagation in UWB indoor wireless channels," *Proc. of IEEE Conf. Ultra Wideband Syst. Technol.*, 305–309, Newark, New Jersey, 2003.

Path Loss Characterization of Body-to-body Radio Propagation Channel on the Angular Variations

H. A. Rahim¹, F. Malek¹, A. I. Ahmad¹, and M. F. A. Malek²

¹Embedded, Networks and Advanced Computing Research Cluster (ENAC)

School of Computer and Communication Engineering

Universiti Malaysia Perlis, P. O. Box 77 d/a Pejabat Pos Besar, Kangar 01000, Malaysia

²MISC LNG Japan MISC Berhad, Yokohama, Japan

Abstract— This paper investigated the path loss characteristics of body-to-body radio propagation channel in a controlled lab environment using a pair of simple structure textile patch antennas operating at 2.45 GHz. A line-of-sight (LOS) case was considered in this measurement campaign. The transmitted antenna was mounted parallel to a right upper arm of the first human body, separated by 2 mm air gap and the received antenna was placed on a left upper arm of second human body. The orientation between transmitted antenna on the first body and received antenna on the second body was carried out for 0, 90 and 180 degrees. The distance between transmitted antenna on the first body and received antenna on the second body was varied from 40 to 180 cm for each angle orientation. Both of the subjects were in a static movement throughout the measurement campaign. This investigation was carried out to observe the path loss variation in LOS scenario in terms of angle orientation and distance variation between two human bodies. The results showed that the measured path loss varied up to 10 dB higher at the orientation 180° compared to other angular orientations as the antenna separation between the first body and the second body increased from 40 to 180 cm. The results also exhibited that the measured path loss varied up to 6.07 dB when the angle between on-body transmitter and on-body receiver was oriented from 90° to 180°. The mean path loss varied on the average of 1.82 dB for all angle orientations.

1. INTRODUCTION

Body-centric wireless communications (BCWC) appears to be different than typical radio communication as it involves the presence of human body which gives significant effects on the radio propagation channel. Due to its promising future technology in numerous fields such as medical, military, emergency services, entertainment and sports, many literatures have focused on investigating and analyzing its characteristics around the body [1–9].

Body-to-body channel falls under one of BCWC categories that involves communication link between radio devices worn on two bodies. This type of communication happens when a radio device mounted on one user's body is transmitting/receiving with a transceiver located on other user's body [4]. To date, very few studies have been reported in the open literature that investigate the channel characteristics of body-to-body radio channel [3–5]. The body-to-body radio propagation channel confronts several issues that need to be addressed, such as on how the position orientations and distances between on-body transmitter and on-body receiver affect the characterization of path loss. Hence, this paper investigates and characterizes body-to-body radio channel propagation in terms of angular orientations and distances variations at 2.45 GHz.

2. EXPERIMENTAL SETTING

This measurement campaign was performed in a controlled indoor environment at Embedded Computing Research Cluster, Universiti Malaysia Perlis, where no other people were presented when the measurement was performed, except two male subjects and data collector. Two simple structure of textile patch antennas with excellent impedance matching were used in this investigation. A portable Agilent Field Fox model number N9923A 2-port Vector Network Analyzer (VNA) was utilized in the measurement setup to capture the frequency response, S_{21} . A total number of sampled points per acquisition $N = 1001$ was set. The average height and weight of two male subjects is 170 cm and 65 kg. The placement of transmitter (Tx) was fixed at the right side of upper arm on the Body 1 (B1) and the placement of receiver (Rx) was fixed at the left side of upper arm on the Body 2 (B2). The antenna to body separation for both subjects was 2 mm air gap. Two 5 meter low loss semi-rigid coaxial cables were used in the measurement campaign, manufactured by Huber-Suhner. The VNA was calibrated during the measurements to ensure the measured data

consists of the reflected signal measured at the antenna ports and to exclude losses from the cables. Accurate antenna placements were ensured during the measurement campaign for each on-body location. Sampling time is 4.25 msec and observation time/sweep duration, T_{obs} is 4.25 sec. The orientation between transmitted antenna on the first body and received antenna on the second body was varied from 0° to 180° with an interval of 90° shown in Fig. 1. The distance between transmitted antenna on the first body and received antenna on the second body was varied from 40 to 180 cm for each angle orientation. Both of the subjects were in a static movement throughout the measurement campaign. Fig. 2 shows the measurement setup utilizing textile patch antennas worn on two subjects' bodies in the indoor environment.

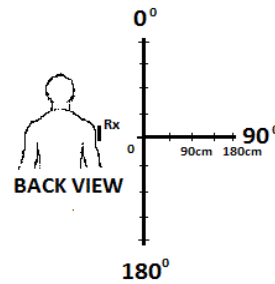


Figure 1: Orientations of the antenna in the measurement campaign for body-to-body communication.

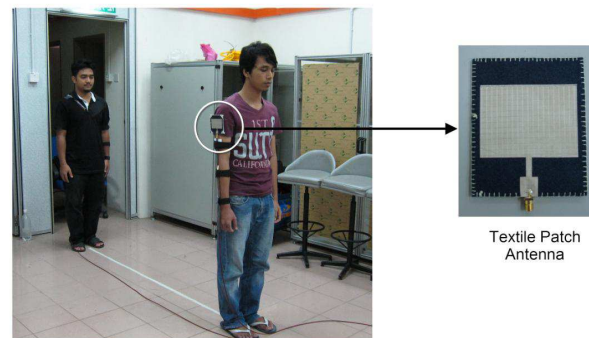


Figure 2: Measurement setup utilizing textile patch antennas worn on two subjects' bodies in indoor environment.

3. BODY-TO-BODY PATH LOSS CHARACTERIZATION

In realistic BCWC operation scenarios, the human will be immersed in a realistic (scattering) propagation environment, and this can have significant influence on the signal propagation around the user body. For the experimental investigation presented in this work, the chosen indoor environment was a lab room and two male subjects were located in the center of the lab room. In order to identify the influence of the environment on the measured radio channel, variation in path loss due to different angular orientations and distance has been investigated.

Figure 3 shows the measured path loss of body-to-body radio channel for three distinct orientations in LOS scenario. Path loss is averaged from five sweep durations for different angular orientation, 0° – 180° with interval of 90° with respect to distance, ranging from 40 cm to 180 cm. The result clearly shows that when the orientation between Tx and Rx increased to 180° the path loss varied up to 10 dB higher than other angular orientations as the distance increased from 40 to 180 cm. This result indicates that a larger path loss variation occurs when the on-body receiver was oriented in maximum angular orientation from the on-body transmitter as the antenna separation between Tx and Rx increased.

Table 1 summarizes the measured mean path loss that is averaged over distances (40–180 cm) in LOS scenario for different angular orientations. The difference values shown in Table 1 represent the difference between the present mean path loss and previous mean path loss. Results show that the mean path loss varies by a maximum of 6.07 dB when the orientation changes from 90° to 180° and 1.82 dB on the average.

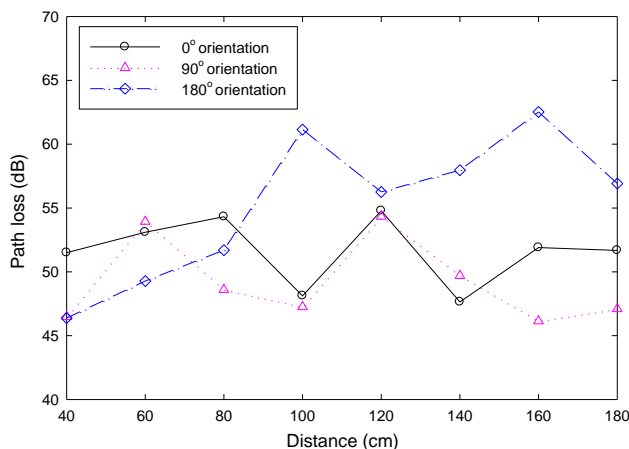


Figure 3: Measured path loss for different angular orientations.

Table 1: Mean path loss for different orientations in LOS scenario.

Mean Path Loss (dB)		
Angle	LOS	Difference
0°	51.63	
90°	49.20	-2.43
180°	55.27	6.07
μ		1.82

Table 2: The path loss exponent and average path loss of body-to-body radio channel.

	Body-to-body	
	LOS	
Angle	γ	$PL_{dB}(d_0)$
0°	0.22	51.62
90°	0.16	49.19
180°	2.16	55.36

Table 2 presents the path loss exponent, γ and the average path loss for each angular orientation. The path loss (PL) of body-to-body radio channel is derived from a function of the logarithmic distance between Tx and Rx, represented as

$$PL_{dB}(d) = PL_{dB}(d_0) + 10\gamma \log\left(\frac{d}{d_0}\right) \quad (1)$$

where $PL_{dB}(d_0)$ is the average path loss at reference distance, $d_0 = 1$ m. d is defined as a distance between Tx and Rx. The PL is calculated for different angular orientations over a range of distance, 40–180 cm. It can be seen that the path loss exponent at the orientation of 180° suffers the least multipath propagation in the scattering environment for body-to-body radio channel communication, i.e., lab environment, compared to other orientations since γ is closer to the path loss exponent of free space ($\gamma_0 = 2$). Results show that smaller values of path loss exponents were obtained at the orientation of 0° and 90° due to the multipath contributions in the indoor environment. An increase of $PL_{dB}(d_0)$ was noticeable as the orientation changed from 0° to 180°, 3.74 dB higher than 0° orientation. Results presented in Table 2 prove that the specific angular orientation plays an important role in the characterization of radio channel involving the presence of two human bodies.

4. CONCLUSIONS

Body-to-body radio propagation channel experimental investigation has been conducted in an indoor environment with respect to different distances and angular orientation between two human bodies for LOS case. The transmitter was placed on the B1, while the receiver was mounted on the B2. The results showed that the measured path loss varied up to 10 dB higher at the orientation 180° compared to other angular orientations as the antenna separation between the first body and the second body increased from 40 to 180 cm. The results indicate that the measured path loss experienced a maximum variation up to 6.07 dB when the angle between on-body transmitter and on-body receiver was oriented from 90° to 180° . It was also found that the mean path loss varied on the average of 1.82 dB for all angle orientations. The measurement shows that the specific angular orientation plays an important part in the body-to-body radio propagation channel characteristics.

ACKNOWLEDGMENT

This work is partially supported by the Short Term Grant (STG) scheme Universiti Malaysia Perlis, under grant 9001-00436. The authors would like to express their highest gratitude to Universiti Malaysia Perlis.

REFERENCES

1. Alomainy, A., et al., “Statistical analysis and performance evaluation for on body radio propagation with microstrip patch antenna,” *IEEE Trans. on Antennas and Propagation*, Vol. 55, No. 1, 245–248, 2007.
2. Alomainy, A., A. Sani, A. Rahman, J. G. Santas, and Y. Hao, “Transient characteristics of wearable antennas and radio propagation channels for ultrawideband body-centric wireless communications,” *IEEE Trans. on Antennas and Propagation*, Vol. 57, No. 4, 875–884, Apr. 2009.
3. See, T. S. P., J. Y. Hee, C. T. Ong, L. C. Ong, and Z. N. Chen, “Inter-body channel model for UWB communications,” *European Conference on Antennas and Propagation (EuCAP)*, 3519–3522, 2009.
4. Cotton, S. L. and W. G. Scanlon, “Channel characterization for single and multiple antenna wearable systems used for indoor body-to-body communications,” *IEEE Trans. on Antennas and Propagation*, Vol. 57, No. 4, 980–990, 2009.
5. Wang, Y., I. B. Borev, J. Ø. Nielsen, I. Z. Kovacs, and G. F. Pedersen, “Characterization of the indoor multi-antenna body-to-body radio channel,” *IEEE Trans. on Antennas and Propagation*, Vol. 57, No. 4, 972–979, 2009.
6. Liu, L., R. D’Errico, L. Ouvry, P. D. Doncker, and C. Oestges, “Dynamic channel modeling at 2.45 GHz for on-body area networks,” *Advances in Electronics and Telecommunications*, Vol. 2, No. 4, 19–27, 2011.
7. Michalopoulou, A., A. A. Alexandridis, K. Peppas, T. Zervos, F. Lazarakis, K. Dangakis, and D. I. Kaklamani, “Statistical analysis for on-body spatial diversity communications at 2.45 GHz,” *IEEE Trans. on Antennas and Propagation*, Vol. 60, No. 8, 4014–4019, Aug. 2012.
8. Abbasi, Q. H., M. M. Khan, S. Liaqat, A. Alomainy, and Y. Hao, “Experimental investigation of ultra wideband diversity techniques for on-body radio communications,” *Progress In Electromagnetics Research C*, Vol. 34, 165–181, 2013.
9. Bari, R. D., Q. H. Abassi, A. Alomainy, and Y. Hao, “An advanced UWB channel model for body-centric wireless networks,” *Progress In Electromagnetics Research*, Vol. 136, 79–99, 2013.

The Influence of Textile Dielectric Properties in On-body Radio Communication Channel Performance at 2.45 GHz

H. A. Rahim¹, F. Malek¹, and M. F. A. Malek²

¹Embedded, Networks and Advanced Computing Research Cluster (ENAC)
School of Computer and Communication Engineering, Universiti Malaysia Perlis
P. O. Box 77, d/a Pejabat Pos Besar, Kangar 01000, Malaysia

²MISC LNG Japan MISC Berhad, Yokohama, Japan

Abstract— This paper presents the characterization of on-body radio communication channel considering the effect of textile dielectric properties at 2.45 GHz. The experiment was performed in the anechoic chamber, involving six subjects that utilized a pair of omni-directional antenna, textile monopole. Four types of textiles clothing with different dielectric properties were considered in this investigation; cotton, nylon, satin and polyester. The transmitted antenna was positioned on a fixed location, right upper arm of the human body and the received antenna position was varied on seven realistic on-body locations for both line-of-sight (LOS) and non-line-of-sight (NLOS) scenarios. This investigation was carried out to determine whether the dielectric properties of textile clothing give significant effect on the characterization of on-body radio communication channel. The path loss model was determined for on-body radio communication channel for LOS and NLOS scenarios. The comparison of path loss values between with and without textile cases was also presented in this work. The result showed that the path loss variation for NLOS was higher than LOS scenario, up to 13.9% for all materials. The results also exhibits that path loss variation was more dependent on the dielectric properties as well as thickness of textile in LOS case compared to NLOS.

1. INTRODUCTION

In recent years, body-centric wireless communications (BCWC) has received great interest because of its huge potential in the near-future for personalized and integrated wearable communications systems [1–4]. As this application will be integrated in the vicinity of human body with clothing, it is crucial to investigate whether textile clothing can give significant effects to the path loss variation of on-body radio communication channel, operating at 2.45 GHz. Very few work has been published in investigating the effect of textile clothing in higher frequency range of 60 GHz [5, 6]. However, none of the literature studies the effect of textile clothing for lower frequency, such as 2.45 GHz. Therefore, this paper is aimed to investigate and analyze the effect of textile properties, such as thickness and dielectric constant, on the path loss characteristics of on-body radio communication channel, considering LOS and NLOS scenarios.

2. MEASUREMENT SETTING

The investigation of the influence of textile in on-body radio communication channel has been performed in an anechoic chamber, located at Electromagnetic Hypersensitivity (EHS) Laboratory, Politeknik Tuanku Syed Sirajuddin, Perlis, Malaysia, by utilizing a portable vector network analyzer (VNA) and a pair of omni-directional planar textile monopole antennas, as analyzed in [1, 2]. Two low loss Huber Suhner coaxial cables of 5 meters were used to connect two antennas to VNA in order to measure the frequency response, S_{21} parameter. A sweep time of 4.25 s with a sampling time of 4.25 ms was configured. The transmitting antenna (Tx) was mounted on a fixed on-body location at right upper arm (RU) of the subject body. The receiving antenna (Rx) was placed at four different locations on the upper torso for LOS scenario, i.e., right chest (RC), left chest (LC), right waist (RW) and left waist (LW). Three other locations were on the back of the body for NLOS case, i.e., right back (RB), left back (LB) and left upper arm (LU), as shown in Fig. 1. 10 mm of separation between antenna and the body is set. The average height and weight of all subjects are 161.1 cm and 62.7 kg, respectively. The measurement was repeated for five times to obtain average value of S_{21} parameter. Table 1 listed the average distance of Tx-Rx separations of on-body channel of 6 subjects. Only static movement was considered in this measurement campaign where the subject was in a standing position throughout the campaign.

Random of shirt samples were chosen for this experimental investigation. The dielectric constant values, ϵ_r , of four textile types such as satin, cotton, nylon and polyester, at 2.58 GHz, have been

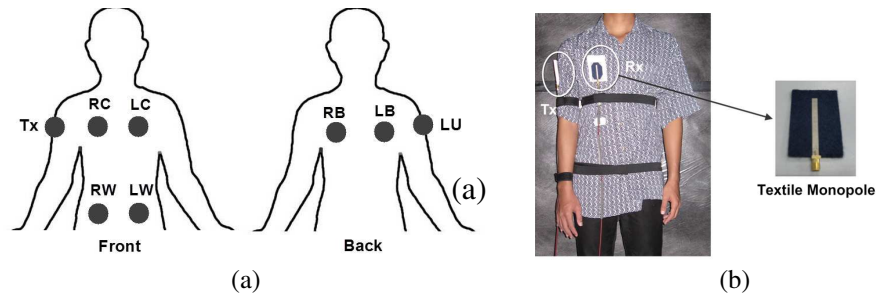


Figure 1: Measurement setup for on-body radio communication channel (a) location of transmitter and receivers and (b) subject wearing nylon shirt with textile monopole antennas.

Table 1: Average distance of Tx-Rx separations of on-body channel of 6 subjects.

On-body Positions	Right Chest	Left Chest	Right Waist	Left Waist	Right Back	Left Back	Left Upper Arm
Distance (cm)	23.3	37.2	29.3	43.8	28.3	42.7	58.3

measured using open-ended coaxial probe technique. The thickness and dielectric constant values of textile materials are summarized in Table 2.

Table 2: Thickness and dielectric constant of four types of textiles.

Material	Thickness, t (mm)	Dielectric constant, ϵ_r
Satin	0.40	1.83
Cotton	0.64	1.76
Nylon	0.47	1.88
Polyester	1.94	1.61

3. RESULTS AND DISCUSSION

The comparison of measured path loss, averaging over 6 subjects, between without any layer of textile and with single layer of four different textiles are illustrated in Fig. 2. It is seen that the measured mean path loss of on-body radio communication channel was generally decreased for all textile materials compared to without textile. The mean path loss decreased from -50 dB at 2.45 GHz when the wave propagated along the body's skin without any layer of textile to a minimum of -48.2 dB when the body torso was covered with nylon. The highest path loss variation

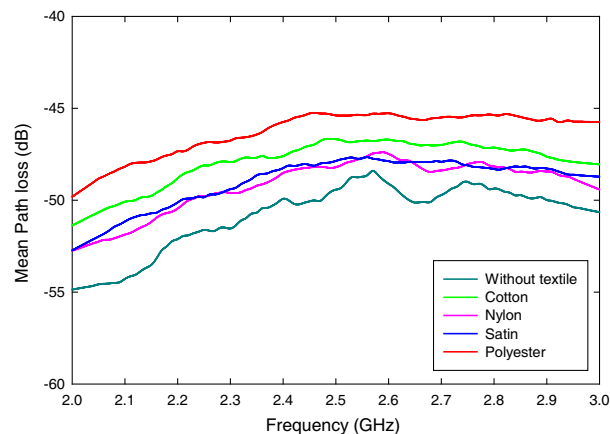


Figure 2: Measured mean path loss of on-body radio channel without and with layer of textiles.

of 10% was exhibited by polyester that has the lowest dielectric constant. This explains that more electromagnetic (EM) wave was reflected in the presence of single layer textile than without textile. This scenario clearly demonstrated that the air-body interface increased the reflection loss of EM wave as the EM wave propagated along the body skin's surface.

Moreover, the statistical analysis has been performed to further investigate the effect of textile on the path loss variation of on-body radio communication channel for different body parts, both for LOS and NLOS scenarios. The mean, μ and standard deviation, σ are determined by fitting to well-known distribution function, log-normal for different on-body radio links. The results are summarized in Table 3. The highest path loss was shown by cotton, occurring on the RU-upper arm link in NLOS case. Meanwhile, nylon obtained the highest path loss for RU-waist link in LOS case, in which has the highest value of dielectric properties. This indicates that for on-body LOS communication, the higher dielectric properties value of textile resulted in the higher losses. The results also show that the path loss variation for on-body NLOS communication was less affected by the dielectric properties value of textile. It is noticed that the path loss values for NLOS case is up to 13.9% higher than LOS case for all materials as there is no direct path of propagation. Satin yielded the highest standard deviation values, both in LOS and NLOS scenarios. Due to satin's smallest thickness, the interface between body and cloth become smaller, resulting an increased in reflection losses.

Table 3: Mean and standard deviation of path loss using log-normal distribution for different textile materials and body-link in LOS and NLOS scenarios.

Material	Path Loss (dB)							
	LOS				NLOS			
	Chest		Waist		Back		Upper Arm	
	μ	σ	μ	σ	μ	σ	μ	σ
Cotton	43.80	3.31	44.29	3.49	46.76	4.78	49.50	2.25
Nylon	45.95	1.0	47.18	3.03	45.24	3.96	47.89	0.55
Satin	46.01	6.75	45.87	8.21	48.35	5.21	47.10	0.09
Polyester	43.23	2.52	42.63	0.76	45.87	4.50	48.75	0.76

Table 4 presents the path loss variation model of on-body radio communication channel for different textile materials in LOS and NLOS scenarios. The mean PL is being calculated. The results demonstrate that satin showed the highest path loss exponent value in LOS case, 93% higher than in NLOS case. For NLOS case, cotton experienced the highest γ value of 16% more than in LOS case. This result also confirmed that the path loss variation in NLOS is less susceptible to the dielectric properties of textile. Larger variation in γ was observed in LOS case compared to NLOS case for all materials.

Table 4: Path loss exponent and mean path loss of on-body radio channel for different textile materials both in LOS and NLOS scenarios.

Material	On-Body			
	LOS		NLOS	
	γ	Mean PL (dB)	γ	Mean PL (dB)
Satin	2.04	35.50	0.14	47.06
Cotton	1.23	37.75	1.46	38.66
Nylon	1.19	40.47	1.38	37.63
Polyester	0.13	42.28	1.39	38.28

Figures 3 and 4 show the dependence of measured mean path loss on the thickness and dielectric properties of textiles for LOS and NLOS scenarios, respectively. As the thickness increased, the mean path loss decreased for both cases, LOS (Fig. 3(a)) and NLOS (Fig. 3(b)). However, the mean path loss decreased, up to 7 dB for LOS, in comparison to NLOS case when the cloth thickness increased to a value of 1.94 mm. The highest rate of change of path loss occurred on RU-LW link

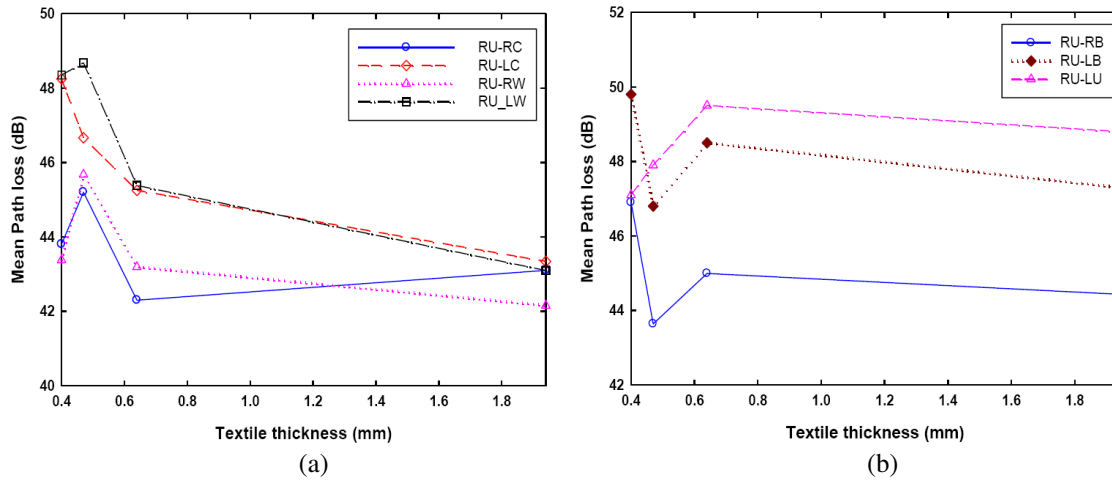


Figure 3: Mean path loss versus different thickness of textiles for different body positions in (a) LOS and (b) NLOS scenarios.

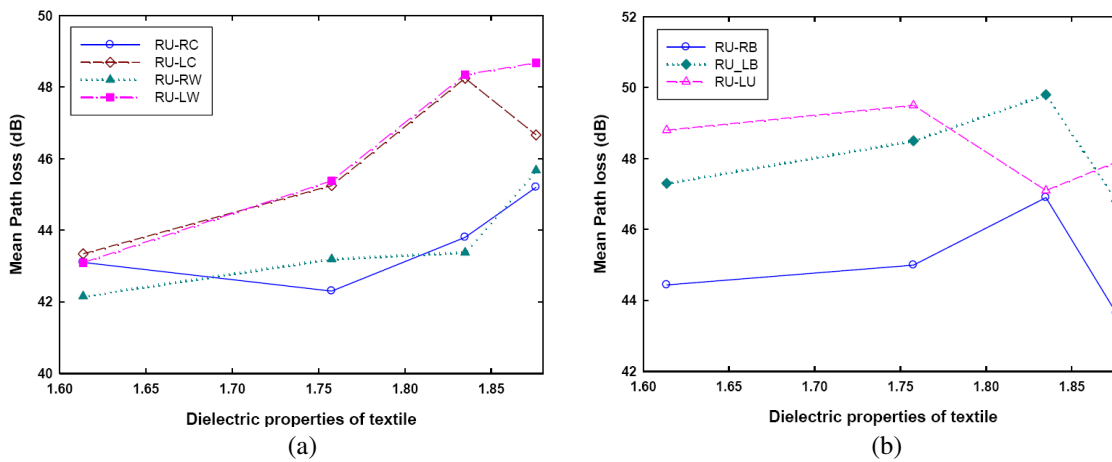


Figure 4: Mean path loss versus different dielectric properties of textiles for different body positions in (a) LOS and (b) NLOS scenarios.

for the case of LOS. The path loss improved significantly as the relative distance between Tx and Rx is larger with an increase of cloth thickness. This is mainly due to the increased of reflected signal in the direct path of propagation (LOS), resulting to the increased of received power. As for NLOS scenario, the textile thickness gives less influence to the path loss variation for all on-body radio links.

The results in Fig. 4 show that the mean path loss increased significantly by a maximum of 5 dB, as the dielectric properties of textile increased, except for RU-LC link in LOS whereas the mean path loss decreased at dielectric properties of 1.83 to 1.88 for the case of RU-RB and RU-LB links. The path loss of LOS increased more than in NLOS scenario by 60% when the dielectric constant of textile rises from 1.61 to 1.88. The results presented in Fig. 3(a) confirms the dependency of path loss variation to dielectric properties of textiles in a main direct propagation path scenario.

4. CONCLUSIONS

Characterization of on-body radio communication channel considering the effect of textile dielectric properties at 2.45 GHz is presented in this work. The experimental investigation was performed in the anechoic chamber, involving six subjects. The aim of this investigation is to determine whether the dielectric properties of textile clothing give significant effect on the characterization of on-body radio communication channel for LOS and NLOS scenarios. The comparison of path loss values between with and without textile cases was also presented in this work. The result showed that the path loss variation for NLOS was higher than LOS scenario, up to 13.9% for all materials. The

results also showed that path loss variation was more dependent on the dielectric properties as well as thickness of textile in LOS case compared to NLOS.

ACKNOWLEDGMENT

This work is partially supported by the Short Term Grant (STG) scheme Universiti Malaysia Perlis, under grant 9001-00436. The authors would like to express their highest gratitude to Universiti Malaysia Perlis.

REFERENCES

1. Alomainy, A., et al., "Statistical analysis and performance evaluation for on body radio propagation with microstrip patch antenna," *IEEE Trans. on Antennas and Propagation*, Vol. 55, No. 1, 245–248, 2007.
2. Liu, L., R. D'Errico, L. Ouvry, P. D. Doncker, and C. Oestges, "Dynamic channel modeling at 2.45 GHz for on-body area networks," *Advances in Electronics and Telecommunications*, Vol. 2, No. 4, 19–27, 2011.
3. Michalopoulou, A., A. A. Alexandridis, K. Peppas, T. Zervos, F. Lazarakis, K. Dangakis, and D. I. Kaklamani, "Statistical analysis for on-body spatial diversit communications at 2.45 GHz," *IEEE Trans. on Antennas and Propagation*, Vol. 60, No. 8, 4014–4019, Aug. 2012.
4. Abbasi, Q. H., M. M. Khan, S. Liaqat, A. Alomainy, and Y. Hao, "Experimental investigation of ultra wideband diversity techniques for on-body radio communications," *Progress In Electromagnetics Research C*, Vol. 34, 165–181, 2013.
5. Xiao, Z., J. Xu, and T. Hu, "Research on the transmissivity of some clothing materials at millimeter-wave band," *2008 International Conference on Microwave and Millimeter Wave Technology Proceedings*, 1750–1753, Nanjing, China, Apr. 2008.
6. Khaleda, A., A. Brizzi, A. Pellegrini, Y. Hao, and A. Alomainy, "Investigation of the effect of fabric in on-body communication using finite difference time domain technique at 60 GHz," *Loughborough Antennas and Propagation Conference (LAPC)*, 1–4, Loughborough, UK, Nov. 12–13, 2012.
7. Rahim, H. A., F. Malek, I. Adam, S. Ahmad, N. B. Hashim, and P. S. Hall, "Design and simulation of a wearable textile monopole antenna for body centric wireless communications," *PIERS Proceedings*, 1381–1384, Moscow, Russia, Aug. 19–23, 2012.

Mathematical and Numerical Analysis of Dielectric Waveguides by the Integral Equation Method

E. Karchevskiy¹ and Y. Shestopalov²

¹Kazan Federal University, Russia

²Karlstad University, Sweden

Abstract— The eigenvalue problems for generalized natural modes of an inhomogeneous dielectric waveguide without a sharp boundary and a step-index dielectric waveguide with a smooth boundary of cross-section are formulated as problems for the set of time-harmonic Maxwell equations with partial radiation conditions at infinity in the cross-sectional plane. The original problems are reduced by the integral equation method to nonlinear spectral problems with Fredholm integral operators. Properties of the spectrum are investigated. The Galerkin and collocation methods for the calculations of generalized natural modes are proposed and convergence of the methods is proved. Some results of numerical experiments are discussed.

1. INTRODUCTION

Many different numerical techniques are applied for computing eigenmodes of dielectric waveguides [1, 2]; namely, Finite-element, Finite-difference, beam propagation, and spline collocation methods, as well as multidomain spectral approach. Often the authors concentrate on the algorithm's features and physical interpretation of the numerical results rather than on fundamental mathematical aspects including the existence, properties, and distribution of the spectra on the complex plane of the spectral parameter. In this study, we propose a new approach to mathematical and numerical analysis of dielectric waveguides based on the methods of spectral theory of operator-valued functions [3, 4] and integral equations (IEs) [4–6]. The eigenvalue problems for the determination of natural modes (surface, leaky, and complex eigenmodes) of inhomogeneous optical waveguides and step-index optical waveguides with the smooth cross-sectional boundary are formulated [3–5] for the time-harmonic Maxwell equations with partial radiation conditions at infinity in the cross-sectional plane. The initial problems are reduced with the aid of the integral equation (IE) method (using appropriate Green functions) to nonlinear spectral problems with Fredholm integral operators. Theorems on the spectrum localization are proved. It is shown that the sets of all eigenvalues of the initial problems may consist of isolated points on the Riemann surface of the spectral parameter (longitudinal wavenumber) and each eigenvalue depends continuously on the frequency and permittivity and can appear or disappear only at the boundary of the Riemann surface. The initial problems for surface waves are reduced to linear eigenvalue problems for integral operators with real-valued symmetric weakly singular kernels. The existence, localization, and dependence of the spectrum on parameters are investigated. The collocation and Galerkin methods for the calculation of natural modes are proposed, the convergence of the methods is proved, and some results of numerical experiments are discussed.

2. GENERALIZED NATURAL MODES OF A STEP-INDEX DIELECTRIC WAVEGUIDE

Let the three-dimensional space be occupied by an isotropic source-free medium, and let the permittivity be prescribed as a positive real-valued function $\varepsilon = \varepsilon(x)$ independent of the longitudinal coordinate and equal to a constant $\varepsilon_\infty > 0$ outside a cylinder. In this section, we consider the generalized natural modes of a step-index optical fiber and suppose that the permittivity is equal to a constant $\varepsilon_+ > \varepsilon_\infty$ inside the cylinder. The axis of the cylinder is parallel to the longitudinal coordinate and its cross section is a bounded domain Ω_i with a twice continuously differentiable boundary γ (see Fig. 1). The domain Ω_i is a subset of a circle with radius R_0 . Denote by Ω_e the unbounded domain $\Omega_e = \mathbb{R}^2 \setminus \bar{\Omega}_i$, by U the space of complex-valued continuous and continuously differentiable in $\bar{\Omega}_i$ and $\bar{\Omega}_e$, twice continuously differentiable in Ω_i and Ω_e functions, and by Λ the Riemann surface of the function $\ln \chi_\infty(\beta)$, where $\chi_\infty = \sqrt{k^2 \varepsilon_\infty - \beta^2}$. Here $k^2 = \omega^2 \varepsilon_0 \mu_0$, ω is a given radian frequency and ε_0, μ_0 are the free-space dielectric and magnetic constants, respectively. Denote by Λ_0 the principal (“proper”) sheet of this Riemann surface specified by the condition $\text{Im} \chi_\infty(\beta) \geq 0$.

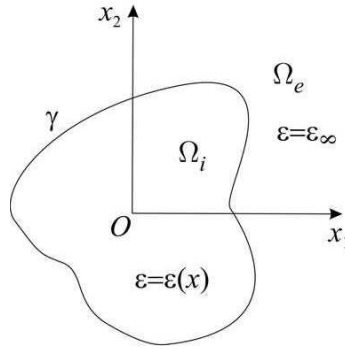


Figure 1: A schematic waveguide's cross-section.

A nonzero vector $\{\mathbf{E}, \mathbf{H}\} \in U^6$ is referred to as a generalized eigenvector (or eigenmode) of the problem corresponding to an eigenvalue $\beta \in \Lambda$ if the following relations are valid [7]:

$$\text{rot}_\beta \mathbf{E} = i\omega\mu_0 \mathbf{H}, \quad \text{rot}_\beta \mathbf{H} = -i\omega\varepsilon_0 \varepsilon \mathbf{E}, \quad x \in \mathbb{R}^2 \setminus \gamma, \quad (1)$$

$$\nu \times \mathbf{E}^+ = \nu \times \mathbf{E}^-, \quad x \in \gamma, \quad (2)$$

$$\nu \times \mathbf{H}^+ = \nu \times \mathbf{H}^-, \quad x \in \gamma, \quad (3)$$

$$\begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix} = \sum_{l=-\infty}^{\infty} \begin{bmatrix} A_l \\ B_l \end{bmatrix} H_l^{(1)}(\chi_\infty r) \exp(il\varphi), \quad r \geq R_0. \quad (4)$$

Here differential operator rot_β is obtained from the standard operator by replacing the generating waveguide line derivative with $i\beta$ multiplication and $H_l^{(1)}(z)$ is the Hankel function of the first kind and index l . The conditions (4) are the the partial radiation conditions.

Theorem 1 (see [7]). *The imaginary axis \mathbb{I} and the real axis \mathbb{R} of the sheet Λ_0 except the set $G = \{\beta \in \mathbb{R}: k^2\varepsilon_\infty < \beta^2 < k^2\varepsilon_+\}$ are free of the eigenvalues of the problem (1)–(4). Surface and complex eigenmodes correspond to real eigenvalues $\beta \in G$ and complex eigenvalues $\beta \in \Lambda_0$, respectively. Leaky eigenmodes correspond to complex eigenvalues β belonging to an “improper” sheet of Λ for which $\text{Im}\chi_\infty(\beta) < 0$.*

Theorem 1 generalizes the well-known results on the spectrum localization of a step-index circular dielectric waveguide which were obtained by the separation of variables method (see, for example [8]).

We use representation of the eigenvectors of problem (1)–(4) in the form of single-layer potentials u and v :

$$\mathbf{E}_1 = \frac{i}{k^2\varepsilon - \beta^2} \left(\mu_0\omega \frac{\partial v}{\partial x_2} + \beta \frac{\partial u}{\partial x_1} \right), \quad \mathbf{E}_2 = \frac{-i}{k^2\varepsilon - \beta^2} \left(\mu_0\omega \frac{\partial v}{\partial x_1} - \beta \frac{\partial u}{\partial x_2} \right), \quad \mathbf{E}_3 = u, \quad (5)$$

$$\mathbf{H}_1 = \frac{i}{k^2\varepsilon - \beta^2} \left(\beta \frac{\partial v}{\partial x_1} - \varepsilon_0\varepsilon\omega \frac{\partial u}{\partial x_2} \right), \quad \mathbf{H}_2 = \frac{i}{k^2\varepsilon - \beta^2} \left(\beta \frac{\partial v}{\partial x_2} + \varepsilon_0\varepsilon\omega \frac{\partial u}{\partial x_1} \right), \quad \mathbf{H}_3 = v, \quad (6)$$

$$\begin{bmatrix} u(x) \\ v(x) \end{bmatrix} = \frac{i}{4} \int_{\gamma} H_0^{(1)} \left(\sqrt{k^2\varepsilon_{+/\infty} - \beta^2} |x - y| \right) \begin{bmatrix} f_{+/\infty}(y) \\ g_{+/\infty}(y) \end{bmatrix} dl(y), \quad x \in \Omega_{i/e}, \quad (7)$$

where unknown densities $f_{+/\infty}$ and $g_{+/\infty}$ belong to the space of Hölder continuous functions $C^{0,\alpha}$. The original problem (1)–(4) is reduced [7] by single-layer potential representation (5)–(7) to a nonlinear eigenvalue problem for a set of singular integral equations on boundary γ . This problem has the operator form

$$A(\beta)w \equiv (I + B(\beta))w = 0, \quad (8)$$

where I is the identical operator in the Banach space $W = (C^{0,\alpha})^4$ and $B(\beta): W \rightarrow W$ is a compact operator consisting particularly of the following boundary singular integral operators:

$$Lp = -\frac{1}{2\pi} \int_0^{2\pi} \ln \left| \sin \frac{t - \tau}{2} \right| p(\tau) d\tau, \quad t \in [0, 2\pi], \quad L: C^{0,\alpha} \rightarrow C^{1,\alpha}, \quad (9)$$

$$Sp = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{ctg} \frac{\tau - t}{2} p(\tau) d\tau + \frac{i}{2\pi} \int_0^{2\pi} p(\tau) d\tau, \quad t \in [0, 2\pi], \quad S: C^{0,\alpha} \rightarrow C^{0,\alpha}. \quad (10)$$

The original problem (1)–(4) is spectrally equivalent [7] to the problem (8). Namely, suppose that $w \in W$ is an eigenvector of the operator-valued function $A(\beta)$ corresponding to an eigenvalue $\beta \in \Lambda_0 \setminus D$, $D = \{\beta \in \mathbb{I}\} \cup \{\beta \in \mathbb{R}: \beta^2 < k^2 \varepsilon_\infty\}$. Then using this vector we can construct the densities of the single-layer potential representation (5)–(7) of an eigenmode $\{E, H\} \in U^6$ of the problem (1)–(4) corresponding to the same eigenvalue β . On the other side, any eigenmode of (1)–(4) corresponding to an eigenvalue $\beta \in \Lambda_0 \setminus D$ can be represented in the form of single-layer potentials. The densities of these potentials constitute an eigenvector $w \in W$ of the operator-valued function $A(\beta)$ corresponding to the same eigenvalue β .

Theorem 2 (see [7]). *For each $\beta \in \{\beta \in \mathbb{R}: \beta^2 \geq k^2 \varepsilon_+\}$ the operator $A(\beta)$ has a bounded inverse operator. The set of all eigenvalues β of the operator-valued function $A(\beta)$ can be only a set of isolated points on Λ . Each eigenvalue β depends continuously on $\omega > 0$, $\varepsilon_+ > 0$, and $\varepsilon_\infty > 0$ and can appear and disappear only at the boundary of Λ , i.e., at $\beta = \pm k\sqrt{\varepsilon_\infty}$ and at infinity.*

Theorem 2 generalizes the known results on the dependence of the propagation constants β of a step-index circular dielectric waveguide on wavenumber k and permittivity ε (see, for example [8]).

The statements similar to Theorems 1 and 2 for the scalar problem in weakly guiding approximation are proved in [9].

Describe a projection method for numerical solution of the problem (8). Denote by N the set of integers. We represent the approximate eigenvector of the operator-valued function $A(\beta)$ in the form

$$w_n = \left(w_n^{(j)}\right)_{j=1}^4, \quad w_n^{(j)}(t) = \sum_{k=-n}^n \alpha_k^{(j)} \exp(ikt), \quad n \in N, \quad j = 1, 2, 3, 4,$$

and look for unknown coefficients $\alpha_k^{(j)}$ by the Galerkin method

$$\int_0^{2\pi} (Aw_n)^{(j)}(t) \exp(-ikt) dt = 0, \quad k = -n, \dots, n, \quad j = 1, 2, 3, 4.$$

$\exp(ikt)$ are orthogonal eigenfunctions of the singular integral operators $L: C^{0,\alpha} \rightarrow C^{1,\alpha}$ and $S: C^{0,\alpha} \rightarrow C^{0,\alpha}$ corresponding to the following eigenvalues:

$$\begin{aligned} \lambda_m^{(L)} &= \{\ln 2 \quad \text{if } m = 0, \quad (2|m|)^{-1} \quad \text{if } m \neq 0\}, \\ \lambda_m^{(S)} &= \{i \quad \text{if } m = 0, \quad i \operatorname{sign}(m) \quad \text{if } m \neq 0\} \end{aligned}$$

for the operators L and S respectively. Hence, the action of the main (singular) parts of the integral operators in (8) on the basis functions is expressed explicitly.

Denote by W_n^T the set of all trigonometric polynomials of the orders up to n . Denote by $W_n \subset W$ the space of the elements $w_n = (w_n^{(j)})_{j=1}^4$ where $w_n^{(j)} \in W_n^T$. Using the Galerkin method for numerical solution of the problem (8), we get a finite-dimensional nonlinear spectral problem

$$A_n(\beta)w_n = 0, \quad A_n: W_n \rightarrow W_n. \quad (11)$$

Theorem 3 (see [10]). *If β_0 belongs to the spectrum $\sigma(A)$ of the operator-valued function $A(\beta)$, then there exists a sequence $\{\beta_n\}_{n \in N}$ with $\beta_n \in \sigma(A_n)$ such that $\beta_n \rightarrow \beta_0$, $n \in N$. If $\{\beta_n\}_{n \in N}$ is a sequence such that $\beta_n \in \sigma(A_n)$ and $\beta_n \rightarrow \beta_0 \in \Lambda$, then $\beta_0 \in \sigma(A)$. If $\beta_n \in \sigma(A_n)$, $A_n(\beta_n)w_n = 0$, and $\beta_n \rightarrow \beta_0 \in \Lambda$, $w_n \rightarrow w_0$, $n \in N$, $\|w_n\| = 1$, then $\beta_0 \in \sigma(A)$ and $A(\beta_0)w_0 = 0$, $\|w_0\| = 1$.*

Figure 2 shows (a) the dispersion curves for complex modes and (b) surface guided modes of step-index waveguides of circular and square cross-sections. The numerical results obtained by the Galerkin method are marked by circles and squares in Fig. 2(a). The dispersion curves for the circular waveguide are plotted by a solid line, $\tilde{\beta} = \beta/(k\sqrt{\varepsilon_\infty})$ and $V = kR\sqrt{\varepsilon_+ - \varepsilon_\infty}$. Fig. 2(b) compares the experimental data [11] for surface waves of a square waveguide (marked by squares) with our numerical results (solid lines). Here a is one half of the square’s side.

The statement similar to Theorem 3 for a scalar problem in weakly guiding approximation is proved in [12].

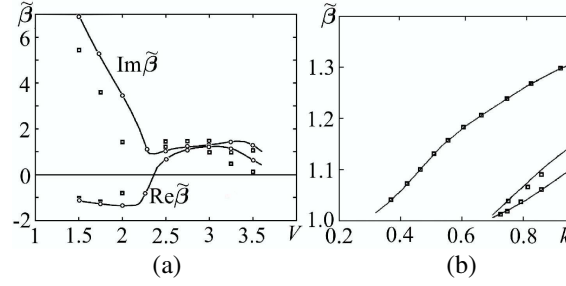


Figure 2: The dispersion curves for (a) the complex modes and (b) surface guided modes of the step-index waveguides of circular and square cross-section.

3. GENERALIZED NATURAL MODES OF AN INHOMOGENEOUS WAVEGUIDE

In this section, we consider the generalized natural modes of an inhomogeneous optical fiber without a sharp boundary. Let the permittivity ε belong to the space $C^2(\mathbb{R}^2)$ of twice continuously differentiable in \mathbb{R}^2 functions. Denote by ε_+ the maximum of the function ε in the domain Ω_i and let $\varepsilon_+ > \varepsilon_\infty > 0$. A nonzero complex vector $\{\mathbf{E}, \mathbf{H}\} \in (C^2(\mathbb{R}^2))^6$ is referred to as a generalized eigenvector (or eigenmode) of the problem corresponding to an eigenvalue $\beta \in \Lambda$ if the following relations are valid [5]:

$$\text{rot}_\beta \mathbf{E} = i\omega\mu_0 \mathbf{H}, \quad \text{rot}_\beta \mathbf{H} = -i\omega\varepsilon_0 \varepsilon \mathbf{E}, \quad x \in \mathbb{R}^2, \quad (12)$$

$$\begin{bmatrix} \mathbf{E} \\ \mathbf{H} \end{bmatrix} = \sum_{l=-\infty}^{\infty} \begin{bmatrix} A_l \\ B_l \end{bmatrix} H_l^{(1)}(\chi_\infty r) \exp(il\varphi), \quad r \geq R_0. \quad (13)$$

Theorem 4 (see [5]). *The imaginary axis \mathbb{I} and the real axis \mathbb{R} of the sheet Λ_0 except the set $G = \{\beta \in \mathbb{R}: k^2\varepsilon_\infty < \beta^2 < k^2\varepsilon_+\}$ are free of eigenvalues of the problem (12), (13). Surface and complex eigenmodes correspond to real eigenvalues $\beta \in G$ and complex eigenvalues $\beta \in \Lambda_0$, respectively. Leaky eigenmodes correspond to complex eigenvalues β belonging to an “improper” sheet of Λ for which $\text{Im}\chi_\infty(\beta) < 0$.*

If vector $\{\mathbf{E}, \mathbf{H}\} \in (C^2(\mathbb{R}^2))^6$ is an eigenvector of problem (12), (13) corresponding to an eigenvalue $\beta \in \Lambda$, then (see [5])

$$\mathbf{E}(x) = k^2 \int_{\Omega_i} (\varepsilon(y) - \varepsilon_\infty) \Phi(\beta; x, y) \mathbf{E}(y) dy + \text{grad}_\beta \int_{\Omega_i} (\mathbf{E}, \varepsilon^{-1} \text{grad} \varepsilon)(y) \Phi(\beta; x, y) dy, \quad x \in \mathbb{R}^2, \quad (14)$$

$$\mathbf{H}(x) = -i\omega\varepsilon_0 \text{rot}_\beta \int_{\Omega_i} (\varepsilon(y) - \varepsilon_\infty) \Phi(\beta; x, y) \mathbf{E}(y) dy, \quad x \in \mathbb{R}^2. \quad (15)$$

Using the integral representation (14) for $x \in \Omega_i$ we obtain a nonlinear eigenvalue problem for an IE in Ω_i which can be written in the operator form

$$A(\beta)\mathbf{F} \equiv (I - B(\beta))\mathbf{F} = 0, \quad (16)$$

where the operator $B(\beta): (L_2(\Omega_i))^3 \rightarrow (L_2(\Omega_i))^3$ corresponds to the right side of the integral representation (14) for $x \in \Omega_i$. For any $\beta \in \Lambda$ the operator $B(\beta)$ is compact [5].

It was proved in [5] that the original problem (12), (13) is spectrally equivalent to problem (16). Namely, suppose that $\{\mathbf{E}, \mathbf{H}\} \in (C^2(\mathbb{R}^2))^6$ is the eigenmode of problem (12), (13) corresponding to an eigenvalue $\beta \in \Lambda$. Then $\mathbf{F} = \mathbf{E} \in [L_2(\Omega_i)]^3$ is an eigenvector of the operator-valued function $A(\beta)$ corresponding to the same eigenvalue β . Suppose that $\mathbf{F} \in [L_2(\Omega_i)]^3$ is an eigenvector of the operator-valued function $A(\beta)$ corresponding to an eigenvalue $\beta \in \Lambda$ and that the same number β is not an eigenvalue of the following problem:

$$[\Delta + (k^2\varepsilon - \beta^2)] u = 0, \quad x \in \mathbb{R}^2, \quad u \in C^2(\mathbb{R}^2), \quad (17)$$

$$u = \sum_{l=-\infty}^{\infty} a_l H_l^{(1)}(\chi_\infty r) \exp(il\varphi), \quad r \geq R_0. \quad (18)$$

Let $E = B(\beta)F$ and $H = (i\omega\mu_0)^{-1}\text{rot}_\beta E$ for $x \in \mathbb{R}^2$. Then $\{E, H\} \in (C^2(\mathbb{R}^2))^6$ and $\{E, H\}$ is an eigenvector of the original problem (12), (13) corresponding to the same eigenvalue β .

Theorem 5 (see [5]). *For each $\beta \in \{\beta \in \mathbb{R}: \beta^2 \geq k^2\varepsilon_+\}$ the operator $A(\beta)$ has a bounded inverse. The set of all eigenvalues β of the operator-valued function $A(\beta)$ can be only a set of isolated points on Λ . Each eigenvalue β depends continuously on $\omega > 0$, $\varepsilon_+ > 0$, and $\varepsilon_\infty > 0$ and can appear and disappear only at the boundary of Λ , i.e., at $\beta = \pm k\sqrt{\varepsilon_\infty}$ and at infinity.*

Similar results for integrated optical guides are obtained in [13].

The scalar problem (17), (18) is a problem on eigenmodes of a nonhomogeneous optical fiber in weakly guiding approximation. The statements similar to Theorems 4 and 5 for scalar problem (17), (18) are proved in [14].

The initial problem (17), (18) for surface waves is reduced to a linear eigenvalue problem for an integral operator with a real-valued symmetric weakly singular kernel. The existence of the spectrum of this operator are proved in [15].

The collocation method for numerical approximation of weakly singular domain integral operators associated with problem (17), (18) is proposed in [15]. The statement similar to Theorem 3 concerning convergence of the collocation method is proved in [15].

As a numerical example Fig. 3 shows the isolines for real and imaginary parts of the fourth eigenfunction of a unite circular waveguide [15]. Here $\varepsilon = 2$, $x \in \Omega_i$, $\varepsilon_\infty = 1$, $\chi_\infty = 2.039 + i1.003$, $k^2 = 5.025$.

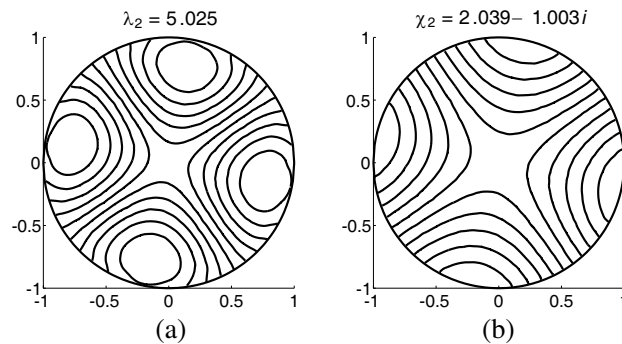


Figure 3: The isolines for (a) real and (b) imaginary part of the fourth eigenfunction of circular waveguide.

ACKNOWLEDGMENT

This work is supported by RFBR 12-01-97012-r-povolzh'e-a and by the Visby program of the Swedish Institute.

REFERENCES

1. Horikis, T. P., "Dielectric waveguides of arbitrary cross sectional shape," *Applied Mathematical Modeling*, Vol. 37, No. 7, 5080–5091, 2013.
2. Xiao, J. and X. Sun, "Full-vectorial mode solver for anisotropic optical waveguides using multidomain spectral collocation method," *Optics Communications*, Vol. 283, No. 14, 2835–2840, 2010.
3. Shestopalov, V. P. and Y. V. Shestopalov, *Spectral Theory and Excitation of Open Structures*, Peter Peregrinus Ltd., 1996.
4. Chernokozhin, E., Y. Smirnov, and Y. Shestopalov, *Logarithmic Integral Equations in Electromagnetics*, VSP Int. Science Publishers, Utrecht, Boston, Köln, Tokyo, 2000.
5. Kartchevski, E. M., A. I. Nosich, and G. W. Hanson, "Mathematical analysis of the generalized natural modes of an inhomogeneous optical fiber," *SIAM J. Appl. Math.*, Vol. 65, No. 6, 2033–2048, 2005.
6. Dautov, R. Z. and E. M. Karchevskii, *Integral Equations Method and Exact Nonlocal Boundary Conditions in the Theory of Dielectric Waveguides*, Kazan State University, Kazan, 2009.
7. Karchevskii, E. M., "The fundamental wave problem for cylindrical dielectric waveguides," *Differential Equations*, Vol. 36, 1109–1111, 2000.
8. Veselov, G. N. and S. B. Raevskiy, *Layered Metal-dielectric Waveguides*, Radio i Sviaz', Moscow, 1988.

9. Karchevskii, E. M., "On the investigation of the spectrum of natural waves in dielectric waveguides," *Comput. Math. Math. Phys.*, Vol. 39, No. 9, 1493–1498, 1999.
10. Karchevskii, E. M., "Mathematical analysis and numerical modeling of the guided modes of the step-index optical fibers," *SIAM Proc. in Applied Mathematics*, Vol. 102, 414–419, 2000.
11. Goncharenko, A. M. and V. A. Karpenko, *The Foundations of the Theory of Optical Waveguides*, Nauka i Tekhnika, Minsk, 1983.
12. Karchevskii, E. M., "Investigation of a numerical method for solving a spectral problem in the theory of dielectric waveguides," *Russian Math. (Iz. VUZ)*, Vol. 43, No. 1, 8–15, 1999.
13. Kartchevski, E. M. and G. W. Hanson, "Mathematical analysis of the guided modes of integrated optical guides," *Mathematical and Numerical Aspects of Wave Propagation — WAVES 2003*, 445–450, Springer, Berlin, 2003.
14. Karchevskii, E. M. and S. I. Solov'ev, "Investigation of a spectral problem for the Helmholtz operator in the plane," *Differ. Equ.*, Vol. 36, No. 4, 631–634, 2000.
15. Frolov, A. G. and E. M. Kartchevskiy, "Integral equation methods in optical waveguide theory," *Springer Proceedings in Mathematics and Statistics*, Vol. 52, 2013.

Coupled Electromagnetic Wave Propagation in Space and around Surfaces and Interfaces

Michael J. Underhill

Underhill Research Ltd, Lingfield, UK

Abstract— The discovery of Electromagnetic (EM) Coupling [1–7] has an impact on the understanding of propagation of waves in free space, in homogeneous and inhomogeneous solids, and at the interfaces between these. The maximum EM coupling factor is found to be $\kappa_0 = 1/2\pi$ and this implies that there are different types of free-space, bulk, and surface EM waves with perhaps a trio of different stable wave impedances depending on the initiating source type.

Waves impinging on the surface of a dielectric or magnetic material can be partially reflected and transmitted and also scattered depending on the relative impedances of the wave and the material. Partial EM coupling means that the wave impedances, scattering and reflections vary progressively on either side of the surface out to a (coupling) distance that is found to be inversely proportional to the square root of frequency.

EM coupling can also cause wave type and impedance to change progressively with distance in free space or in a medium such as the ionosphere or a solid material.

1. INTRODUCTION

Electromagnetic coupling has been found to have a maximum asymptotic value of about $1/2\pi$ or 16% [2, 7]. The fact that this is substantially less than 100% previously assumed has a dramatic effect on our understanding and representation of all electromagnetics and physics [1–8]. Here we address the impact on the propagation of waves in and just outside boundaries of homogeneous, composite or inhomogeneous materials.

2. WAVE IMPEDANCES IN THE COUPLED TRANSMISSION LINE MODEL OF ALL ELECTROMAGNETICS

Figure 1 shows the EM coupling between the overlapping power-flow filaments assumed in the Coupled Transmission Line (CTM) Model of EM [2]. There are four types of filament and a pair of these is shown. The maximum coupling factor is found to be $\kappa_0 = 1/2\pi = 0.16$ [2, 7]. Figure 2 shows the low-pass and high-pass equivalent circuits of the filaments [2]. Figure 3 below in Section 4 expands the series and shunt elements in this model in more detail.

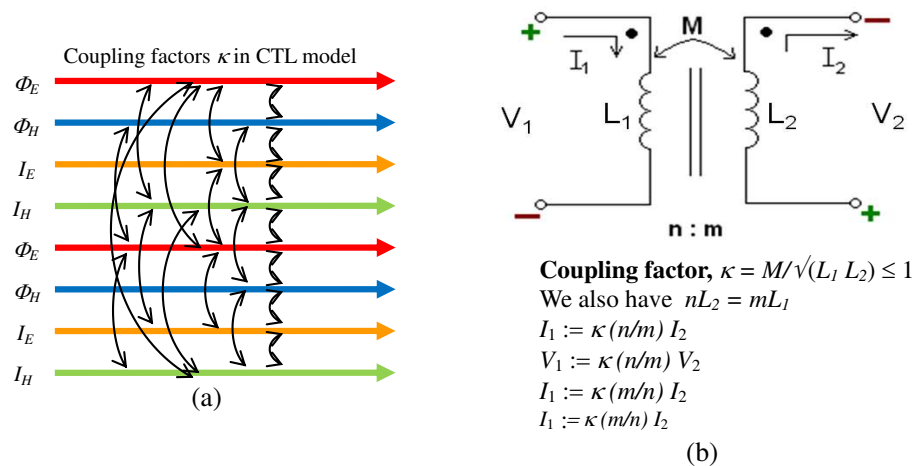


Figure 1: The coupling factors in the (four) Coupled Transmission Line model of Electro-Magnetic (EM) waves. (a) Shows equivalent power flow filaments two groups of the four main types of wave propagation in a small volume of space, with coupling between all filaments. The filaments may be adjacent and non-overlapping, if of the same type, or fully overlap, if of different types. (b) Shows the equivalent ‘transformer’ model of the coupling with relevant definitions and equations.

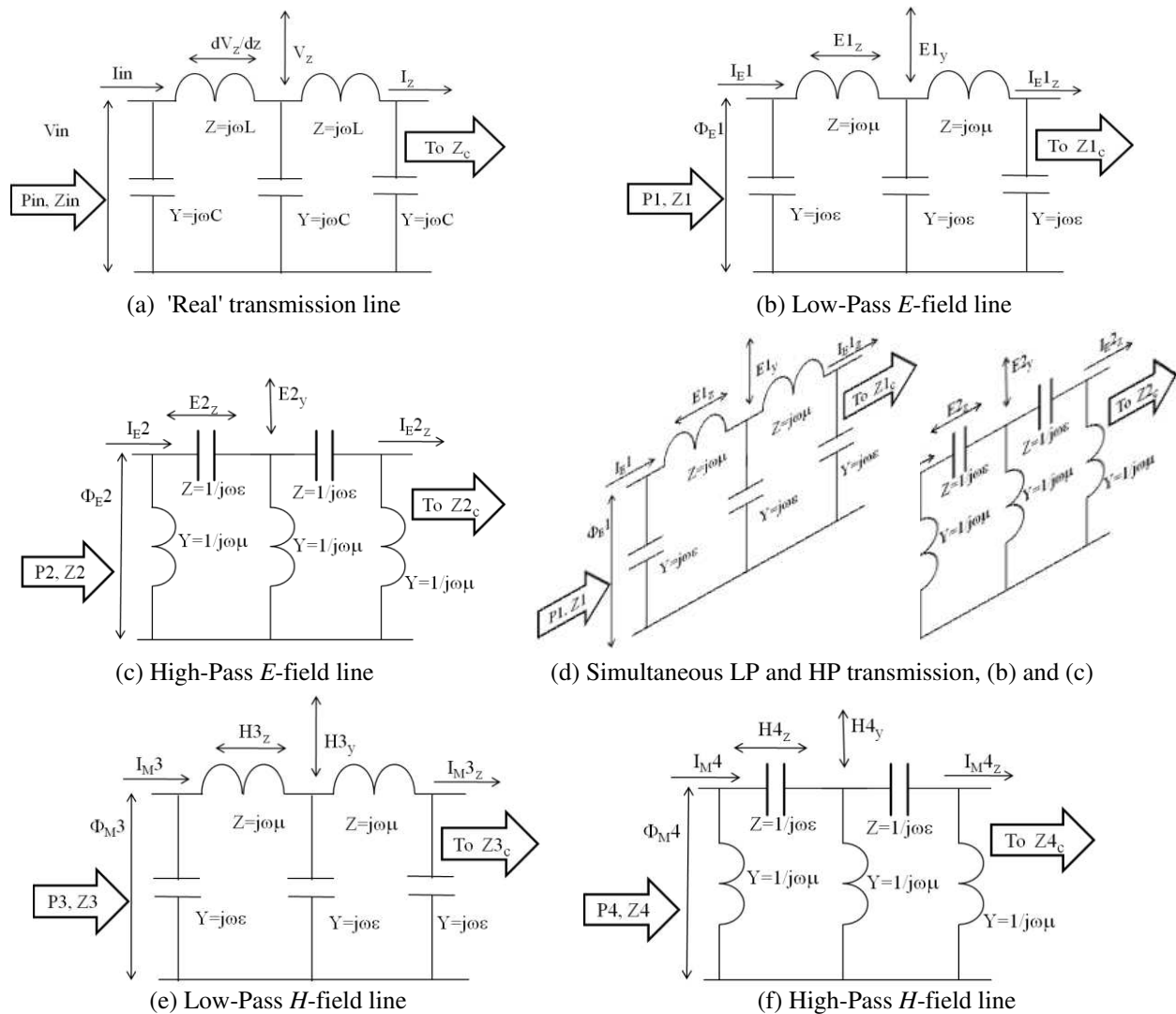


Figure 2: The four transmission line model of Electro-Magnetic (EM) waves in a local ether defined by permittivity ϵ and permeability μ . (a) is the basis transmission line lumped equivalent circuit, with voltage V and current I . L and C are values per unit length along the line. (b) is the free-space 'electric' transmission line low pass (LP) equivalent to the plane wave solution of Maxwell's equations, with L mapped to μ and C mapped to ϵ . V is mapped to the electric vector potential Φ_E and I is mapped to electric displacement current I_E . (c) is a novel discovery: it is the high-pass (HP) solution to Maxwell's equations that is found to dominate above about 47.75 MHz. (d) shows that both the HP and LP models are active around 47.75 MHz. (e) and (f) are the LP and HP magnetic current and magnetic vector potential equivalents, respectively of (b) and (c).

3. NEW WAVE TYPE DEFINITIONS AND IMPEDANCES

In free space or in uniform homogenous solids, we find that different types of plane wave travelling in the same direction can co-exist. The stable wave types have different proportions of E , H , D and B energy and these types can co-exist separately in a mixture even if travelling in the same direction.

For the Coupled Transmission Line model in Figures 1 and 2 the fields are more conveniently converted to electric and magnetic vector potentials, $\Phi_E = E$ and $\Phi_H = H$, and currents, $I_E = j\omega D$ and $I_H = j\omega B$ [2, 7].

Wave types can be defined by whatever component is strongest and dominant. For example E or D can be dominant in TE waves and H and B can be dominant in TE waves. TEM waves can then be considered to be a balanced mixture of TE and TM waves. impedances $Z_w = E/H$. As a consequence there are a number of valid definitions for wave impedance such as $E/j\omega D$, E/H ,

B/D , $H/\omega B$ and these encompass the TE, TEM, and TM waves that we see in waveguides and elsewhere.

Based on the assumption that the asymptotic value of EM coupling for free space is $\kappa_0 = 1/2\pi$ we find values for *stable* linearly polarised wave impedances that are not $Z_0 = 120\pi$ (377 ohms). The stable values of wave impedance in the case of TEM waves where E and H energies are equal are found to be $E/H = Z_0/2\pi = 60$ ohms, $E/H = Z_0 = 120\pi = 377$ ohms, and $E/H = 2\pi$, $Z_0 = 240\pi^2 = 2369$ ohms. There are similar trios of stable values possible for TE and TM waves.

Because of the axiom of ‘energy conservation’ all waves oscillate between two types of energy and there is no ‘double frequency energy pulsation’ as implied by the classical definition of the Poynting Vector as $P = E \times H$. The concept of quadrature components for potentials and currents solves this misconceived problem [3].

We know that pairs of orthogonally linear or circularly polarised waves may exist in the same volume. Even if these waves are very lightly coupled, there will be preferred asymptotically stable wave impedances, reached at some distance from the source. This opens up the possibility that there are many types of plain-wave free-space EM waves that have not yet been identified and investigated. Any mathematical analysis has to include the EM coupling factor for the wave components under consideration.

The Propagation through a non-uniform in-homogenous material such as in the ionosphere also depends on the wave impedance of the impinging wave. An unsolved issue is whether all waves in free space morph with distance eventually to only one value of wave impedance, or can maintain the initial ‘stable’ wave impedance out to a large or infinite distance. More practical measurements are needed.

4. DISTRIBUTED CIRCUIT THEORY AND WAVE PROPAGATION CONSTANTS

Figure 3 shows series elements Y_s and shunt elements Z_t of Figure 2 expanded in more detail to give the ‘dissipative or lost power’ radiation, reflection and heat-loss components. For propagation constant $k = 2\pi/\lambda$, the other parameters for the transmission line are:

$$\text{Velocity } v = \omega/k = f\lambda = 1/\sqrt{LC} \quad \text{and} \quad \text{Impedance } Z_w = \sqrt{(Z_s/Y_t)} \quad (1)$$

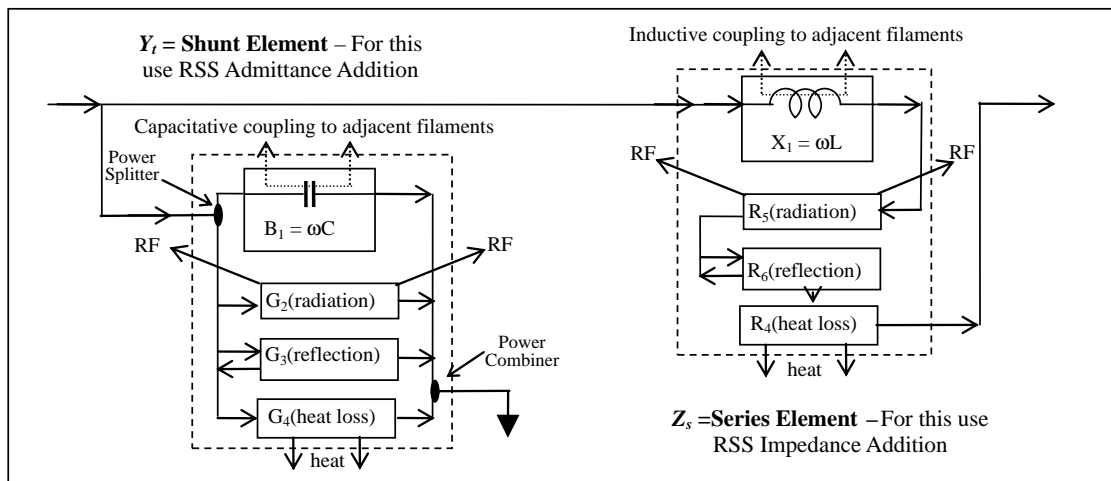


Figure 3: Distributed Circuit Theory block diagram of series and shunt lumped equivalent finite elements of a Filamentary Coupled Transmission Line EM Model shown in Figure 2. The use of the RSS (Root-Sum-of-the-Squares) Process Addition Rule is essential. Traditional Lumped Circuit Theory and rules are not valid for this distributed circuit model. Energy conservation is strictly observed.

To ensure energy conservation for distributed circuits with no physical circuit nodes, the use of the RSS rule for combining processes and their components, is essential. RSS means Root-Sum-of-the-Squares. It is the central and essential rule of Distributed Circuit Theory. It is the key for improved Electromagnetic Theory and improved EM Modelling in the future. It is an essential part of ‘Analytic Region Modelling’ [9]. For the elements in Figure 2 the RSS Rule gives a number of

possible algebraic definitions of which the best choice found so far is

$$Y_t = \frac{(G_2^2 + G_3^2 + G_4^2) + jB_1^2}{\sqrt{(G_2^2 + G_3^2 + G_4^2 + B_1^2)}} \quad \text{and} \quad Z_S = \frac{(R_6^2 + R_7^2 + R_8^2) + jX_5^2}{\sqrt{(R_6^2 + R_7^2 + R_8^2 + X_5^2)}} \quad (2)$$

In the operation of Equation (2) it is best to combine resistive components separately from reactive components and then combine them as shown in the equation. This recognises the fact that all the circuit components actually represent separate processes that are only partially coupled because they are all distributed processes albeit overlapping in the same space.

5. WAVE PROPAGATION CONSTANTS IN AND ABOVE MATERIALS

Surface waves above a conducting material or either side of a dielectric material can be in layers. One mode or type of propagation can dominate in each layer. The layers and modes are always lightly coupled with a maximum coupling factor of $\kappa_0 = 1/2\pi$. As a consequence each mode may be considered separately in the first place but then successively modified by adding in the effects of any significantly coupled adjacent modes.

In Figure 3 mode coupling is capacitive to $B_1 = \omega C$ and inductive to $X_5 = \omega L$. As a consequence of the maximum $\kappa_0 = 1/2\pi = 0.16$, the maximum effect on C is an increase to $1.16C$, and the maximum effect on L is a decrease to $0.84L$.

A consequence of the limited coupling is that a line carrying power excites a small amount of power in an adjacent coupled line. A stable equilibrium over a short distance can be achieved for two identical lines when a maximum of about $1/(2\pi)^2 = 2.5\%$ of the original power has been transferred to the secondary line. But for longer distances the power can transfer progressively until there are equal powers in the two lines.

The resistive and conductive components also couple through B_1 and X_5 . If they are unequal in a pair of coupled lines, one line will be lossier than the other. Then a stable equilibrium will be achieved at a distance when a maximum of 2.5% of the lossy line loss is coupled into the low loss line.

6. PROGRESSIVE REFLECTION AT MATERIAL BOUNDARIES

One consequence of partial EM coupling is that the reflection of a wave from a partially transparent material occurs progressively either side of the boundary. It is as if there is a smooth change of permittivity, permeability, or refractive index from one side of the boundary to the other.

The changes appear as two sided exponential functions as for example represented by the hyperbolic tangent function \tanh used in reference [5]. Figure 4 gives an example using a symmetric \tanh function. But because the wave velocity is usually lower in the material an asymmetric type of \tanh function should be used as shown in Figure 4 with material to the right. A suitable function is under investigation.

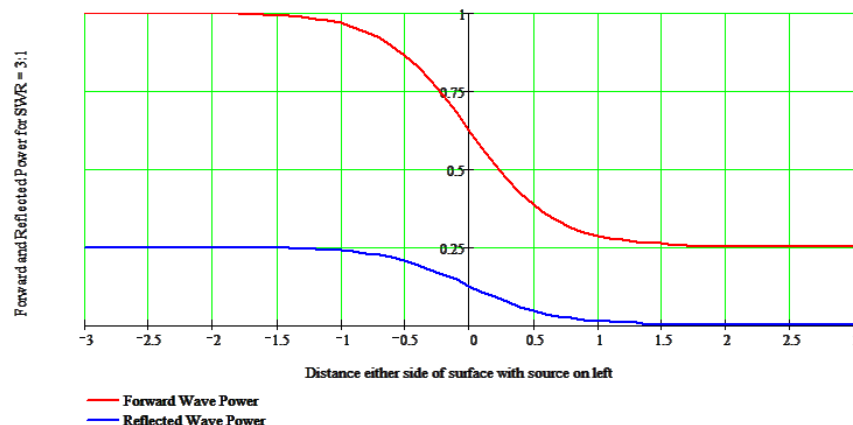


Figure 4: Progressive reflection of a wave at a material boundary. Wave source is at left and material to the right of centre.

EM coupling means that the exponential terms in Figure 4 will vary progressively on either side of the surface out to a (coupling) distance that is inversely proportional to the square root of frequency.

7. CONCLUSIONS AND FUTURE WORK

The impact of (partial) EM coupling on the propagation and reflection of waves as been outlined with some examples.

The RSS (Root-Sum-of-the-Squares) Process Combination Rule has been found to be essential for combining the components that make up the series and shunt elements of the CTL (Coupled Transmission Line) Universal Model of Electromagnetics [2].

Another conclusion is that there may be additional types of propagating (plane) wave that are stable to useful distances. How these might propagate in an ionosphere needs further investigation. Another possibility under investigation is whether it is possible to have a stable ‘tube’ of radiation where one field component is radially directed and the complementary field is solenoidal [1, 2, 8]. Such a (self-focussed) wave could have much less spreading and expansion with distance. The source antenna for such a wave would then effectively be a high gain small antenna. One possible difficulty is that there is likely to be an additional self focussing or defocusing action that is wave power dependent. Further experimental work should be able to demonstrate this.

A further conclusion is that partial EM coupling causes the wave impedances, scattering and reflections vary progressively on either side of the surface out to a (coupling) distance that has been found to be inversely proportional to the square root of frequency.

REFERENCES

1. Underhill, M. J., “The phase noise spectrum and structure of photons?,” *Proc. 16th EFTF-2010*, 8 pages, Noordwijk, Netherlands, Apr. 13–16, 2010.
2. Underhill, M. J., “A physical model of electro-magnetism for a theory of everything,” *PIERS Online*, Vol. 7, No. 2, 196–200, 2011.
3. Underhill, M. J., “Maxwell’s transfer functions,” *PIERS Proceedings*, 1766–1770, Kuala Lumpur, Malaysia, Mar. 27–30, 2012.
4. Underhill, M. J., “A local ether lens path integral model of electromagnetic wave reception by wires,” *PIERS Proceedings*, 1005–1008, Moscow, Russia, Aug. 19–23, 2012.
5. Underhill, M. J., “Antenna pattern formation in the near field local ether,” *PIERS Proceedings*, 1009–1012, Moscow, Russia, Aug. 19–23, 2012.
6. Underhill, M. J., “Anomalous ground wave tilt measured over wet ground,” *IET Conf. on Ionospheric Radio Systems and Techniques 2012*, Royal York Hotel, York, UK, May 15–17, 2012.
7. Underhill, M. J., “Wideband small loop-monopole HF transmitting antenna with implications for Maxwell’s equations and the Chu criterion,” *PIERS Proceedings*, 764–768, Taipei, Mar. 25–28, 2013.
8. Underhill, M. J., “Electromagnetic structures and inertias of particles including the higgs boson,” *PIERS Proceedings*, 401–405, Taipei, Mar. 25–28, 2013.
9. Underhill, M. J., “Novel analytic EM modelling of antennas and fields,” *PIERS Proceedings*, 1771–1775, Kuala Lumpur, Malaysia, Mar. 27–30, 2012.

The Stability of EM Particles and Predicted Mass Ratios

Michael J. Underhill

Underhill Research Ltd., Lingfield, UK

Abstract— Feasible electromagnetic structures for all particles have been proposed [1–3]. These support the proposition that there is an ‘Electromagnetic Model for the Electromagnetic Theory of Everything’ [3]. The prime basis for the Electromagnetic Model for Everything is the discovery of Electromagnetic (EM) Coupling [1–8]. One mechanism to be explained in electromagnetic terms is the physical process that gives rise to the measured stable masses of all the sub-atomic particles in the Standard Model [9] of particle physics. This paper is an examination of the factors that may contribute to the stability of the EM models of particles. An empirical formula for mass ratios within elementary particle classes is derived.

Any form of energy has a *gravitational* mass in accordance with $E = mc^2$. This includes the energy in fields or low density particles such as photons and neutrinos [1, 2]. Also in [1] an EM coupling mechanism was given for the inertial mass of particles having a dense central core. Then inertial mass is defined to be equal to gravitational mass. Low density forms of energy such as heat light and EM and gravitational fields are considered to be ‘dark matter’ having a low inertia mass relative to gravitational mass.

The ‘core-atmosphere’ particle model in [1] is the basis for finding particle stability conditions. The core consists of concentrated ‘substance’ that creates the surrounding atmosphere. All stable particles can be considered to have line spectra with one main dominant line. Stability then requires that the potential of the atmosphere is exactly in phase quadrature with the oscillating core ‘substance’. Positive deviation in phase signifies that the particle has excess energy which will be radiated. Negative deviation in phase signifies that the particle has a deficit of energy and it will scavenge surrounding (dark matter or field) energy until it reaches the stable quadrature phase condition. The energy-frequency relation $E = hf$ is assumed to hold for particles, but not for dark matter.

1. INTRODUCTION

The underlying thesis is that all particles and fields are electromagnetic and have defined shapes and density profiles in the three spatial dimensions and in the two dimensions of time and frequency [1]. Stability defines the evolution of the particle in time, whether its energy grows, as in its formation, or decays by radiation or fission. The concepts and mathematics of ‘Transfer Functions’, as used by control engineers and circuit engineers, are therefore particularly applicable [5].

All matter and particles in [1–8] have so far been defined in terms of potentials and ‘*substances*’. *Substance* has been defined to include, plain charge, polarisation (charge), electric and magnetic currents, particle and sub-atomic substance cores, and any source material that generates, or is generated by, a potential. It should be noted that recent Physics convention in general now appears to call all these ‘currents’ but differentiated into a number of types. However it is not clear whether all these types of currents have been defined in the past as capable of being generated, created or induced by potentials, as required by the new concept of ‘Electromagnetic Coupling’. Therefore the concept of *substance* as defined above will be continued here.

The symbol ψ has been adopted for substance to signify that it has a strong relationship to what is defined in wave function equations. The main difference is that it is defined as a density that in general is oscillatory with a line spectrum, but is otherwise a constant density that does not fluctuate. Thus ψ is not a probability function. Substance exists all the time with a stable shape and a stable radial distribution profile.

Electromagnetic string structures suitable for photons [2] and the main sub-atomic particles [3] have previously been presented. A photon is an ‘electromagnetic string arrow’ with a single line frequency spectrum as in Figures 1(a) and 1(b). It is envisaged to be ‘dark matter’ with low inertia. A photon is a boson and has zero or unit spin. The unit spin can be taken to correspond to left-hand or right-hand circular polarisation in the direction of energy flow. Protons, electrons and neutrons are described by two or more interlaced toroidal string loops arguably having two or more dominant line frequency components in the particle’s spectrum [3]. Defined along the direction of energy flow electron core string structures are shown in Figures 1(c) to 1(g). The interlacing can be regarded as a representation of circular polarisation and ‘intrinsic spin’. These

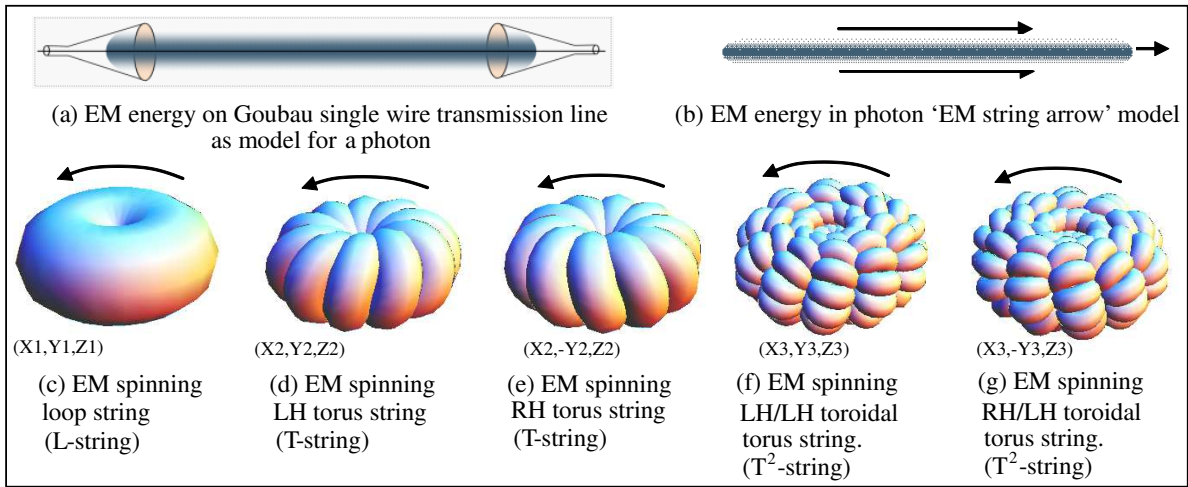


Figure 1: EM String models for photons and sub-atomic particles: (a) EM energy on Goubau single wire transmission line. (b) EM energy in photon ‘EM string arrow’ model, with similar energy distribution to (a). (c) EM spinning loop string; an L-string. (d) EM spinning LH torus string; a T-string. (e) EM spinning LH torus string; a T-string and a mirror image of (d). (f) EM spinning LH/LH toroidal torus string; a T^2 -string. (g) EM spinning RH/LH toroidal torus string; a mirror image of (f).

three particles have half unit spin and so are classified as ‘fermions’. All these particle models have an electromagnetic frequency spectrum with discrete lines. The total energy of the spectral lines is the mass m of a particle according to $E = mc^2$. The Physical EM model has been extended to include mechanisms for particle inertia, particle beam diffraction, charged particles and a possible EM neutrino structure [1].

Figures 1(d) and 1(e) are possible EM string models for electrons with opposite (intrinsic) spins. The spin multiplicity is 2 and can be considered as two extra dimensions. Figures 1(f) and 1(g) are possible models for quarks. They allow three dimensions or independent variables of spin that in combination could define the six flavours of up, down, strange, charm, bottom and top [3].

2. GENERAL MODEL OF PARTICLES

Particle stability conditions can be deduced from the General Model of Particles given in [1]. All particles are defined as a compact high density ‘substance’ core surrounded by a low density potential (evanescent wave) ‘atmosphere’. Substance ψ is essentially all forms of electromagnetic currents and charges and RSS summed in any volume according to their energies. (RSS means ‘Root-Sum-of-the-Squares’). The core substance creates a surrounding (spherical) potential distribution Φ that decays as $1/r$ outside the core, where r is the radial distance from the core centre. Particle cores and atmosphere potentials have the same (line) spectrum.

At any point potential Φ and substance ψ are partially coupled by the asymptotic local coupling factor κ_0 . For EM strings and current filaments $\kappa_0 = 1/2\pi$ [4, 5], and for compact quasi-point sources $\kappa_0 = 1/4\pi$. (The latter value can be derived from the formula that defines the directivity D of an antenna of capture aperture area A as $D = 4\pi A/\lambda^2$).

For convenience we can define both Φ and ψ in units of the square root of energy. Then we find that just outside the surface of the core $\Phi_2 = \kappa_0\psi_1$ and $\Phi_1 \propto 1/r$. In turn the potential Φ_2 induces or creates substance with density ψ_2 . This means that the atmosphere has energy density $U = \Phi_2\psi_2$.

A more accurate representation of the combined core and atmosphere potential distribution for the dominant line in the particle spectrum is given by Equation (1) of reference [2] as

$$\Phi_n = \left(1 - e^{-\kappa_1/(\sqrt{f \times r})}\right) \quad (1)$$

This function of r is meromorphic and so advantageously has no central singularity. Away from the centre a $1/r$ law holds.

Figure 2 shows that the origin of particle inertia is a time delay τ in the coupling between core and atmosphere. When acceleration a occurs, the offset d is given by:

$$d = a\tau_a \quad (2)$$

The atmosphere is primarily a potential distribution created by the core substance. It is taken to be ‘dark matter’ which has gravitational mass in accordance with $E = mc^2$. But because it is diffuse with no dense point-like cores dark matter has a much lower inertial mass for the same amount of energy or gravitational mass.

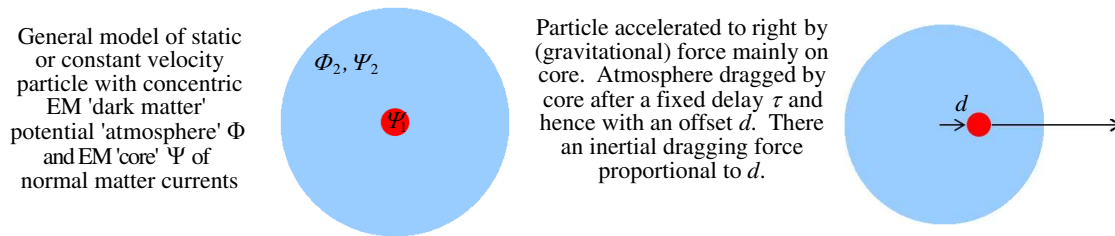


Figure 2: General model of a particle with an explanation for the inertia of normal matter.

3. FINITE ENERGY REGIONS AND PARTICLE DENSITY PROFILES

The stability conditions of particles are expected to be a function of the particle size, profile and dominant line frequency f_p as given in [1].

The atmosphere of a particle has energy density $U = \Phi_2 \psi_2 \propto 1/r_2$. It follows that the energy added per unit radial distance is constant. The energy of the atmosphere has to be a finite value. We therefore assume that the atmosphere has a finite energy that is less than the core energy by a coupling factor κ_2 and therefore terminates at a finite radial distance $r_2 = r_1/\kappa_2$ where r_1 is the radius of the core. At a further distance the energy density decays more rapidly than the inverse square law to avoid the total field energy becoming greater than the core energy. This fundamental process has to hold at all frequencies down to static fields as from a static charge. The energy of the field atmosphere can never become infinite. For the ‘coherent’ coupling that occurs for a particle with a line spectrum, the termination distance can be expected to be κ_2 times the Goubau distance [1, 2].

Figure 3 shows the particle energy density spherical ‘regions’ for two cases having different (EM) coupling factors. Calibration of the coupling and sizes of these two types of particles is under further investigation. More measurement data is needed. Existing theoretical assumptions may have to be modified.

4. PHASE DIFFERENCE BETWEEN SUBSTANCE AND POTENTIAL AS THE PARTICLE STABILITY CONDITION

The ‘core-atmosphere’ particle model in [1] is the basis for finding particle stability conditions. All particles are taken to be electromagnetic structures with a line frequency spectrum. For the

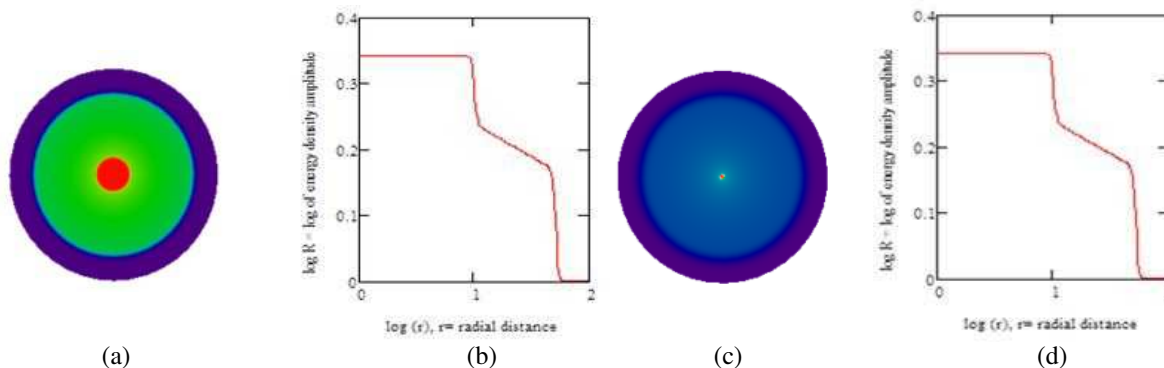


Figure 3: Particle spherical energy density ‘regions’ consisting of central dense ‘substance’ dominated ‘core’ (red) surrounded by $1/r$ potential density dominated ‘atmosphere’, which is size limited by a defined energy constraint, and abruptly descends to the free-space energy density level (violet): (a) Profile of highly self-coupled particles with dominant line spectra with cores as in Figures 1(c) to 1(f). (b) Plot of log of energy density amplitude against log of radial distance for (a). (c) Profile of lightly self-coupled Higgs particle with broadband energy spectrum. (d) Plot of log of energy density amplitude against log of radial distance for (c).

sub-atomic particles such as the electron, proton and neutron, the major particle mass component is assumed to the energy in the main dominant frequency spectrum line f_{p1} . First level stability conditions can then be derived from examination of the phase difference between derived the particle core substance ψ_1 and the potential Φ in which the core is placed at the core main frequency.

Stability requires that the potential of the atmosphere, where the potential Φ dominates, is exactly in phase quadrature with the oscillating core, where the ‘substance’ dominates. Positive deviation in phase signifies that the particle has excess energy which will be radiated. Negative deviation in phase signifies that the particle has a deficit of energy and it will scavenge surrounding (dark matter or field) energy until it reaches the stable quadrature phase condition.

The core actually has various types of string structure. Those shown in Figure 1 are a subset of stable particle configurations. And the structure affects the effective mean coupling between the core and its own potential. In a class of particles the main loops are assumed to be one, two, three, n_1 or more wavelengths of the particle main frequency. Also there may be subsidiary loops pinned either side or inside the main loop. Which possible structures apply to the various classes of particles in the Standard Model [9] is yet to be determined.

The energy-frequency relation $E = hf$ is assumed to hold for particles, but not for dark matter. Another important factor is that EM coupling is found to be inversely proportional to the square root of the main particle frequency f_{p1} . Thus the linear size of the particle atmosphere is also inversely proportional to the square root of the frequency f_{p1} . A further factor is that there is a time delay τ_{ca} between the core and its atmosphere that accounts for particle inertia as shown in Figure 2. In principle the delay can be calculated from the Gravitational Constant G but it also requires that the EM coupling is known accurately, which at the moment is not the case. However it is assumed to be a definite value, and this enables an empirical formula for the ratio of the masses of elementary particles described below. This delay sets the phase difference θ between the particle core and its atmosphere depending on the particle main spectrum line frequency f_{p1} . We thus have:

$$\theta = 2\pi\tau_{ca}f_{p1} \quad (3)$$

We also have the particle energy E to frequency f_{p1} relation in terms of the Planck Constant h and combined with Einstein particle mass m to energy equivalence equation in terms of the velocity of light c :

$$E = hf_{p1} = mc^2 \quad (4)$$

The radial size R_a of the particle (vector potential dominant) atmosphere is assumed to be the same as the EM coupling distance [1, 2]. The calibration factor for this has so far only been estimated approximately ($\sim 10\%$) as $f_g = 14\text{MHz}$ to give

$$R_a = (f_g/f_{p1})^{1/2} = (14\text{MHz}/f_{p1})^{1/2} \quad (5)$$

Further subsidiary components of lesser energy may be necessary for total particle stability. These can be assumed to play an important role in the stability of the more complex particles such as the nuclei of atoms with respect to radioactive decay lifetimes. These components also give deviations of mass from the main mass formula denoting mass relationships in a particular class.

5. MASS RELATIONSHIP IN AN ELEMENTARY PARTICLE CLASS

The Standard Model groups elementary particles into a four by five matrix of shared properties [9]. Tables 1 and 2 compare known and predicted masses for the electron and neutrino rows in the matrix. A formula has been derived by trial-and-error for the ratios of the masses in these two rows. The formula has been derived from what is known about electromagnetic coupling and the principle of particle stability given above.

The choice of the formula has been made on a trial and error basis as a compromise between simplicity and best fit for the ratios of the masses. To reduce the number of possibilities string loops were taken to have a circumference that is a whole number of wavelengths n . Some elementary particles may have two, three or more loops, each with circumferences that are integer number of wavelengths n_i of the same frequency. The consequence is that the particle frequency f_{p1} in Equations (3) and (4) is reduced by $n = \Sigma n_i$ to give:

$$E = mc^2 = nhf_{p1} = \Sigma n_i hf_{p1} \quad (6)$$

By trial and error the following formula has been derived for the mass ratios in Tables 1 and 2:

$$m_i = m_{ref} \times (x_i)^p \quad (7)$$

Table 1: Lepton and W boson mass ratio prediction.

	Electron	Muon	Tau	W boson
Known Mass	0.511 MeV/c ²	105.7 MeV/c ²	1.777 GeV/c ²	80.4 GeV/c ²
Mass Prediction	0.511	$0.511 \times 2.905^5 = 106.63$	$0.511 \times 5.1^5 / 1000 = 1.798$	$0.511 \times 10.95^5 / 1000 = 80.44$

Table 2: Neutrino and Z boson mass ratio prediction.

	Electron neutrino	Muon neutrino	Tau neutrino	Z boson
Known Mass	< 0.22 eV/c ²	< 0.17 MeV/c ²	15.5 MeV/c ²	91.2 GeV/c ²
Mass Prediction	0.22	$0.22 \times 6.95^7 / 10^6 = 0.173$	$0.22 \times 13.2^7 / 10^6 = 15.36$	$0.22 \times 17.03^7 / 10^6 = 91.4$

The values of x_i are fine tuned to give agreement with the measured mass values. Note that these are nearly integer values. The small deviations from exact values are thought to be akin to end shortening effect of resonant dipole antennas and a similar lengthening effect observed in loop antennas. The values of p as 5 and 7 respectively were chosen for best fit to integer values.

One possible particle shape class is a central loop with two smaller loops either side of it, each loop having an integer number of wavelengths. For example the W boson in Table 1 may have a 5 wavelength central loop with a pair of 3 wavelength loops either side of it, making a total of 11 as in Table 1. The atmosphere could then be an appropriate higher order spherical harmonic mode [10].

6. CONCLUSION

Assuming that all particles have a line frequency spectrum usually with one dominant line component, the particle stability can be taken to be dependent on an exact quadrature phase difference between the particle ‘core (mass)’ and the surrounding ‘(field potential)’ atmosphere. A positive phase deviation from quadrature signifies radiation of energy and mass from the particle. Negative deviation in phase signifies that the particle has a deficit of energy and it will scavenge surrounding (dark matter or field) energy until it reaches the stable quadrature phase condition. It acts as a temporary miniature ‘black-hole’.

Some mass ratios of fundamental particles have been arranged into a formula derived empirically from the new principles of ‘electromagnetic coupling’. Direct computation of masses from a formula requires the coupling factor variation inversely with the square root of frequency to be measured very accurately. Further work is needed.

REFERENCES

1. Underhill, M. J., “Electromagnetic structures and inertias of particles including the higgs boson,” *PIERS Proceedings*, 401–405, Taipei, Mar. 25–28, 2013.
2. Underhill, M. J., “The phase noise spectrum and structure of photons?,” *Proc. 16th EFTF-2010*, 8 pages, Noordwijk, Netherlands, Apr. 13–16, 2010.
3. Underhill, M. J., “A physical model of electro-magnetism for a theory of everything,” *PIERS Online*, Vol. 7, No. 2, 196–200, 2011.
4. Underhill, M. J., “Maxwell’s transfer functions,” *PIERS Proceedings*, 1766–1770, Kuala Lumpur, Malaysia, Mar. 27–30, 2012.
5. Underhill, M. J., “A Local ether lens path integral model of electromagnetic wave reception by wires,” *PIERS Proceedings*, 1005–1008, Moscow, Russia, Aug. 19–23, 2012.
6. Underhill, M. J., “Antenna pattern formation in the near field local ether,” 1009–1012, *PIERS Proceedings*, Moscow, Russia, Aug. 19–23, 2012.

7. Underhill, M. J., “Wideband small loop-monopole HF transmitting antenna with implications for maxwell’s equations and the chu criterion,” *PIERS Proceedings*, 764–768, Taipei, Mar. 25–28, 2013.
8. Underhill, M. J., “Coupled electromagnetic wave propagation in space and around surfaces and interfaces,” *PIERS Proceedings*, Stockholm, Accepted for Publication.
9. http://en.wikipedia.org/wiki/Standard_Model.
10. http://en.wikipedia.org/wiki/Spherical_harmonics.

A Novel Approach for the Construction of Dirac's Delta Functions

A. R. Baghai-Wadji

University of Cape Town, South Africa

Abstract— In recent works a series of Dirac delta functions were constructed from Poisson's equations in isotropic or an-isotropic dielectrics subject to simple boundary conditions. The construction of delta functions was also briefly touched upon in connection with Maxwell's equations in unbounded free space. The present paper generalizes the previously introduced ideas and applies them to more general cases; provides physical interpretation of the results; and offers a recipe on how to employ the constructed Dirac delta functions in regularizing singular surface integrals arising in computational electromagnetics in isotropic, anisotropic, or bi-anisotropic media.

1. INTRODUCTION

To sketch the underlying ideas for constructing physics-based problem-tailored Dirac delta functions, consider the following generic problem. Let the system of partial differential equations $\mathcal{L}\Psi(\mathbf{x}) = \mathbf{s}(\mathbf{x})$ be given. Solve this system subject to given boundary conditions. Here, \mathcal{L} is a matrix differential operator; $\Psi(\mathbf{x})$ is the solution to be determined; $\mathbf{s}(\mathbf{x})$ is a given source function (direct problem); and $\mathbf{x} = (x_1, x_2, x_3)$. In applications we are mostly faced with indirect problems. This distinction is, however, not relevant in the present discussion. To solve this problem consider $\mathcal{L}\mathbf{G}(\mathbf{x}) = \mathbf{u}\delta(\mathbf{x} - \mathbf{x}')$, subject to the boundary conditions of the original problem, or simplified boundary conditions, as it is the case in the majority of practical problems. Here, $\mathbf{G}(\mathbf{x})$ refers to the dyadic Green's functions (DGFs); \mathbf{u} is a unit vector in space; and $\delta(\mathbf{x} - \mathbf{x}')$ is the Dirac delta function in symbolic form. The specification "symbolic" shall play a significant role in the discussion. The construction of DGFs usually requires the transformation of the equations into spectral domain. Utilizing the sifting property of the symbolic delta function, the simple exponentially harmonic Fourier transform of the delta function is immediate. Hereby, the major point is that no integral representations for the delta function are required in order to construct the DGFs.

Assume next that the DGFs have been determined. As a rule, $\mathbf{G}(\mathbf{x})$ is obtained in terms of Fourier-type integrals. Substitute the expression for $\mathbf{G}(\mathbf{x})$ into $\mathcal{L}\mathbf{G}(\mathbf{x}) = \mathbf{u}\delta(\mathbf{x} - \mathbf{x}')$. Thus, the LHS generates integral representations for the symbolic delta function $\delta(\mathbf{x} - \mathbf{x}')$ appearing at the RHS. It can be shown that not only integral representations for the Dirac delta functions, but also for Cauchy Principle Values, associated with the problems at hand, can be constructed.

The present paper details the underlying ideas along with the implications for the regularization of the associate singularities of the DGFs. The aforementioned discussion should justify the claim that the proposed method is a problem-tailored regularization technique.

Remark. The discussion in the following sections critically depends on the employment of inhomogeneous rather than homogeneous plane wave expansions for the fields. To this end, $\mathcal{L}\Psi(\mathbf{x}) = \mathbf{s}(\mathbf{x})$, introduced above, will be diagonalized [1]. Indeed diagonalization seems to be the single most powerful mathematical tool required for the machinery developed in this paper and for many other related ideas and concepts [2–4].

2. ELECTROSTATIC FIELDS IN ANISOTROPIC MEDIA

Statement of the problem. Consider an unbounded homogeneous anisotropic dielectric medium characterized by the positive-definite 3×3 symmetric permittivity matrix $\boldsymbol{\varepsilon}$. Assume there is no variation in the y -direction, with the y -axis pointing into this plane. Let the x -axis to run horizontally from left to the right! Let the z -axis run vertically from south to north! Assume a line-charge source (modeled by a Dirac delta function) with the coordinates (x', z') to run parallel to the y -axis. Determine the induced electric potential distribution in entire space, the infinite-domain Green's function associated with this problem.

Formal Statement of the Problem. The Poisson's equation $\nabla \cdot \mathbf{D} = \delta(x - x', z - z')$ establishes the relationship between divergence of the dielectric displacement $\mathbf{D}(x, z)$ and the ideally-localized charge distribution $\delta(x - x', z - z')$. The relationship between the electric field vector $\mathbf{E}(x, z)$ and the electric potential field $\varphi(x, z)$ can be deduced from the fact that $\mathbf{E}(x, z)$ is curl-free in static limit: $\mathbf{E} = -\nabla\varphi$. The linear constitutive equation $\mathbf{D} = \boldsymbol{\varepsilon}\mathbf{E}$ governs the relationship between $\mathbf{D}(x, z)$ and $\mathbf{E}(x, z)$.

2.1. On the Construction of the Infinite Domain Green's

Field- and constitutive equations in component form. Assume the fictitious plane $z = z'$, passing through the point (x', z') , partitions space into two semi-spaces $z > z'$ (Domain I) and $z < z'$ (Domain II). The following component-wise field- and constitutive equations hold valid in each of the Domains I and II:

$$\partial_x D_1 + \partial_z D_3 = 0, \quad D_1 = \varepsilon_{11} E_1 + \varepsilon_{13} E_3, \quad D_3 = \varepsilon_{13} E_1 + \varepsilon_{33} E_3, \quad E_1 = -\partial_x \varphi, \quad E_3 = -\partial_z \varphi \quad (1)$$

Interface conditions. $\varphi(x, z)$ is continuous across the fictitious interface plane $z = z'$. $D_3(x, z)$ has a jump discontinuity equal to $\delta(x - x')$.

Solution procedure. Expressing D_1 and D_3 in terms of φ in (1) results in the Laplace's equation. Solutions of the type $\varphi(x, z|k) \propto e^{jkx} e^{\gamma z}$ suggest themselves leading to:

$$\gamma^\pm = -\frac{\varepsilon_{1e}}{\varepsilon_{33}} jk \pm \frac{\varepsilon_P}{\varepsilon_{33}} |k| \quad \text{with} \quad \varepsilon_P = \sqrt{\varepsilon_{11}\varepsilon_{33} - \varepsilon_{13}^2} \quad (2)$$

Standard procedure: solution in domains I and II. The potential function in domain I, $\varphi^+(x, z|k)$, has the general form: $\varphi^+(x, z|k) = A^+(k)e^{jkx} e^{\gamma^-(k)z} + B^+(k)e^{jkx} e^{\gamma^+(k)z}$. Since the potential must vanish for $z \rightarrow \infty$, $B^+(k) = 0$. Similar considerations can be made for $\varphi^-(x, z|k)$:

$$\varphi^+(x, z|k) = A^+(k)e^{jkx} e^{\gamma^- z}; \quad \varphi^-(x, z|k) = B^-(k)e^{jkx} e^{\gamma^+ z} \quad (3)$$

The interface condition for $\varphi(x, z)$. The continuity condition for the electric potential function at the fictitious interface plane $z = z'$; i.e., $\varphi^+(x, z|k)|_{z=z'_+} = \varphi^-(x, z|k)|_{z=z'_-}$ implies

$$A^+(k)e^{\gamma^-(k)z'} = B^-(k)e^{\gamma^+(k)z'} \stackrel{\text{def}}{=} P(k), \quad (4)$$

Expressing $A^+(k)$ and $B^-(k)$ in terms of $P(k)$, results in $\varphi^\pm(x, z|k)$ and consequently $D_3^\pm(x, z|k)$:

$$\varphi^\pm(x, z|k) = P(k)e^{jkx} e^{\gamma^\mp(z-z')} = \varphi^\pm(x, z'_\pm|k) e^{\gamma^\mp(z-z')} \quad (5a)$$

$$D_3^\pm(x, z|k) = \pm \varepsilon_P |k| P(k) e^{jkx} e^{\gamma^\mp(z-z')} = D_3^\pm(x, z'_\pm|k) e^{\gamma^\mp(z-z')} \quad (5b)$$

where the values of the displacement normal components at the two banks of the interface plane; i.e., $D_3^+(x, z'_+|k)$ and $D_3^-(x, z'_-|k)$, are defined as:

$$D_3^\pm(x, z'_\pm|k) = D_3^\pm(x, z|k)|_{z=z'_\pm} = \pm \varepsilon_P |k| P(k) e^{jkx} \quad (6)$$

Before embarking on the interface condition for the normal component of the displacement vector; i.e., $D_3(x, z)$, the Fourier transform of the charge distribution on the interface $z = z'$ is needed:

$$\rho(k) = \int_{-\infty}^{\infty} dx \delta(x - x') e^{-jkx} = e^{-jkx'}. \quad \text{Consequently, } \delta(x - x') = \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{-jkx'} e^{-jkx}.$$

The interface condition for $D_3(x, z)$ in real space reads $D_3^+(x, z'_+) - D_3^-(x, z'_-) = \delta(x - x')$. Or, in spectral domain $\varepsilon_P |k| P e^{jkx} + \varepsilon_P |k| P(k) e^{jkx} = e^{-jkx'}$, leading to $P(k) = \frac{1}{2\varepsilon_P |k|} e^{-jkx'}$. Having determined $P(k)$, the field problem has been solved in spectral domain:

$$\varphi^+(x, z|k) = \frac{1}{2\varepsilon_P |k|} e^{jk(x-x')} e^{\gamma^-(z-z')}; \quad \varphi^-(x, z|k) = \frac{1}{2\varepsilon_P |k|} e^{jk(x-x')} e^{\gamma^+(z-z')} \quad (7a)$$

$$D_3^+(x, z|k) = \frac{1}{2} e^{jk(x-x')} e^{\gamma^-(z-z')}; \quad D_3^-(x, z|k) = -\frac{1}{2} e^{jk(x-x')} e^{\gamma^+(z-z')} \quad (7b)$$

$$D_1^+(x, z|k) = -\frac{1}{2} \frac{\gamma^-}{jk} e^{jk(x-x')} e^{\gamma^-(z-z')}; \quad D_1^-(x, z|k) = \frac{1}{2} \frac{\gamma^+}{jk} e^{jk(x-x')} e^{\gamma^+(z-z')} \quad (7c)$$

Consistency. It is gratifying to experience the intricate interplay between various terms, and the cancellation of certain terms, in making sure that $\partial_x D_1 + \partial_z D_3$ vanishes in Domains I and II.

Remarks. Several comments are in place. i) $\lim_{|k| \rightarrow \infty} \gamma^\pm / (jk) = \text{constant}$. ii) In region I: $z - z' = |z - z'|$, while in region II: $z - z' = -|z - z'|$. These relationships will frequently be used, to express the fields in entire space in a unified form. iii) The expression for $\varphi(x, z|k)$ has

a singularity at $k = 0$. Elsewhere has been shown that realistic systems satisfying the charge neutrality condition compensate for this singularity. The main concern in the present discussion is the slow, or the absence of any exponential decay whatsoever, at the limit $z \rightarrow z'$; the inverse Fourier integrals representing the fields diverge in the near-field even if the system, as a whole, is charge neutral. The concern in the present discussion is the behavior of integrands for $|k| \rightarrow \infty$. iv) $D_1(x, z|k)$ and $D_3(x, z|k)$ do not have any singularities at $k = 0$! The decaying behavior in the exponential terms $e^{\gamma^\pm(z-z')}$ seems to be reassuring! However, these expressions are plagued with difficulties as well. For $z \rightarrow z'$ the terms $e^{\gamma^\pm(z-z')}$ approach unity and thus the integrals do not converge in conventional sense. The limiting integrals are distributions. Recall that $D_3(x, z)$ had to satisfy the interface condition which involved the Dirac's delta function. This fact suggests that $D_3(x, z)$ in the limits $z \rightarrow z'_\pm$ must be proportional to a Dirac's delta function! This will be shown affirmatively.

2.2. Green's Function in Real Space

Substituting for γ^\pm into $\varphi^+(x, z|k)$ and $\varphi^-(x, z|k)$, along with a symmetry argument allow combining the two integral representations for $\varphi^+(x, z|k)$ and $\varphi^-(x, z|k)$ arriving at

$$\varphi(x, z) = \frac{1}{2\pi\varepsilon_P} \int_0^\infty dk \frac{1}{k} \cos \left(k \left[(x - x') - \frac{\varepsilon_{13}}{\varepsilon_{33}}(z - z') \right] \right) e^{-\frac{\varepsilon_P}{\varepsilon_{33}}k|z-z'|}, \quad (x, z) \in \mathbb{R} \times \mathbb{R} \quad (8)$$

This integral can be calculated in closed-form. Adopting $G(x - x', z - z')$ (Green's function) for $\varphi(x, z)$ and ignoring the irrelevant constant term $(-1/4\pi\varepsilon_P)\ln(\varepsilon_{33}/\varepsilon_P)$, result in:

$$G(x - x', z - z') = -\frac{1}{4\pi\varepsilon_P} \ln \left| \frac{\varepsilon_{33}}{\varepsilon_P}(x - x')^2 - 2\frac{\varepsilon_{13}}{\varepsilon_P}(x - x')(z - z') + \frac{\varepsilon_{11}}{\varepsilon_P}(z - z')^2 \right| \quad (9)$$

Remark. Let \mathbf{r} denote the vector with the components $x - x'$ and $z - z'$, along the x - and z -axis, respectively. It is important to note that the argument of the logarithm in (9) is given by $\mathbf{r}^T \boldsymbol{\varepsilon}^{-1} \mathbf{r}$, with \mathbf{r}^T being \mathbf{r} transposed. For $\mathbf{r} \neq \mathbf{0}$, $\mathbf{r}^T \boldsymbol{\varepsilon}^{-1} \mathbf{r} > 0$ in virtue of the positive-definiteness of $\boldsymbol{\varepsilon}$ and thus $\boldsymbol{\varepsilon}^{-1}$. In the near-field zone; i.e., in the neighborhood of $\mathbf{r} = \mathbf{0}$, the Green's function in (9) has an (integrable) logarithmic singularity. The discussion in the next section presents a novel method for the regularization of singularities in the vicinity of the source. The implication of this property for numerical calculations is that the diagonal and near-diagonal elements of the discretized system matrix are renormalized.

3. A METHOD FOR THE REGULARIZATION OF GREEN'S FUNCTION SINGULARITIES IN THE NEAR-FIELD ZONE

The proposed method is based on the realization that $G(x - x', z - z')$ is conventionally constructed as the potential response to an idealized localized Dirac's delta function. As demonstrated in the preceding section the line charge distribution $\delta(x - x')$, or, equivalently, its Fourier transform $e^{-jkx'}$, enters the calculation of $G(x - x', z - z')$. Given a BVP several questions arise: (i) Is it possible to construct a problem-specific representation for the Dirac's delta function $\delta_\eta(x - x')$ depending on a "small control parameter" η ? If such a representation exists, it would define a continuous series of parametrized functions which would converge to $\delta(x - x')$ at the limit $\eta \rightarrow 0$. The delocalized function $\delta_\eta(x - x')$ in spectral domain would then be integrable in ordinary, rather than distributional sense. (ii) Is it possible to construct the electric potential response $G_\eta(x - x', z - z')$ to $\delta_\eta(x - x')$? If such a function exists, and provided it can be constructed, it can be expected to be regular and free from singularities in the near-field zone. Then $G_\eta(x - x', z - z')$ can be used to carry out field analysis in accordance to the recipes available in the BEM. Let $f_\eta(x, z)$ denote any of the components of the solution vector. It can be expected that $f_\eta(x, z)$ in the limit $\eta \rightarrow 0$ converges to the field solution corresponding to the Dirac's delta function $\delta(x - x')$, thereby, bypassing the difficulties associated with the Green's function singularities.

In what follows the above questions will all be answered affirmatively. It should also be pointed out that the stream of the arguments will be valid for the electromagnetic waves and fields in general bi-anisotropic media, and even more generally, for problems in mathematical physics. It is instructive to point out that the underlying principle is the diagonalization of PDEs in mathematical physics which already has been covered in several earlier occasions and will not be addressed here.

3.1. The Limit of the Dielectric Displacements at the Fictitious Interface

Substituting for γ^\pm into the expressions for $D_3^+(x, z|k)$ and $D_3^-(x, z|k)$ and rearranging gives:

$$D_3^\pm(x, z|k) = \pm \frac{1}{2} e^{jk[(x-x') - \frac{\epsilon_{13}}{\epsilon_{33}}(z-z')]} e^{\mp \frac{\epsilon_P}{\epsilon_{33}} |k|(z-z')} \quad (10)$$

Integrating over k from $-\infty$ to ∞ , along with a symmetry argument, lead to:

$$D_3^\pm(x, z) = \pm \frac{1}{2\pi} \frac{\frac{\epsilon_P}{\epsilon_{33}} |z - z'|}{\left[(x - x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} |z - z'| \right]^2 + \left[\frac{\epsilon_P}{\epsilon_{33}} |z - z'| \right]^2} \quad (11)$$

Integrating with respect to x from $-\infty$ to ∞ and writing η for $|z - z'|$, yield:

$$\lim_{z \rightarrow z'_\pm} \int_{-\infty}^{\infty} dx D_3^\pm(x, z) = \pm \frac{1}{2\pi} \lim_{\eta \rightarrow 0} \int_{-\infty}^{\infty} dx \frac{\frac{\epsilon_P}{\epsilon_{33}} \eta}{\left[(x - x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} \eta \right]^2 + \left(\frac{\epsilon_P}{\epsilon_{33}} \eta \right)^2} \quad (12)$$

Proceed as follows: i) Substitute $u = x - x'$. ii) Substitute $x = \eta u$. Observe that η^2 cancels out, making $\lim_{\eta \rightarrow 0^+}$ irrelevant. iii) Substitute $u = x \mp \frac{\epsilon_{13}}{\epsilon_{33}}$ as the case maybe. iv) Finally substitute $u = \frac{\epsilon_P}{\epsilon_{33}} x$ and cancel the term $(\epsilon_P/\epsilon_{33})^2$. The resulting integrals both give π , thus leading to:

$$\lim_{z \rightarrow z'_+} \int_{-\infty}^{\infty} dx D_3^+(x, z) = \frac{1}{2}; \quad \lim_{z \rightarrow z'_-} \int_{-\infty}^{\infty} dx D_3^-(x, z) = -\frac{1}{2} \quad (13a)$$

Thus it is proved that the following relationships hold valid individually:

$$\lim_{z \rightarrow z'_+} D_3^+(x - x', z - z') = \frac{1}{2} \delta(x - x'); \quad \lim_{z \rightarrow z'_-} D_3^-(x - x', z - z') = -\frac{1}{2} \delta(x - x') \quad (14)$$

In view of this result and (10) it can be concluded that

$$\lim_{z \rightarrow z'_\pm} \int_{-\infty}^{\infty} \frac{dk}{2\pi} D_3^\pm(x, z|k) = \pm \frac{1}{2} \lim_{z \rightarrow z'_\pm} \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{jk[(x-x') - \frac{\epsilon_{13}}{\epsilon_{33}}(z-z')]} e^{\mp \frac{\epsilon_P}{\epsilon_{33}} |k|(z-z')} = \pm \frac{1}{2} \delta(x - x') \quad (15)$$

In terms of $\eta = |z - z'|$ these equations read:

$$\lim_{z \rightarrow z'_\pm} \int_{-\infty}^{\infty} \frac{dk}{2\pi} D_3^\pm(x, z|k) = \pm \frac{1}{2} \lim_{\eta \rightarrow 0} \int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{jk[(x-x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} \eta]} e^{-\frac{\epsilon_P}{\epsilon_{33}} |k|\eta} \quad (16a)$$

$$= \pm \frac{1}{2} \lim_{\eta \rightarrow 0} \delta_\eta^\pm(x - x') = \pm \frac{1}{2} \delta(x - x') \quad (16b)$$

Interpretation. From (16) it can be deduced that:

$$D_3^\pm(x, z'_\pm|k) = \pm \frac{1}{2} \lim_{\eta \rightarrow 0} e^{jk[(x-x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} \eta]} e^{-\frac{\epsilon_P}{\epsilon_{33}} |k|\eta} \quad (17)$$

Equation (17) suggest the introduction of $D_{3,\eta}^+(x, z'_+|k)$ and $D_{3,\eta}^-(x, z'_-|k)$ according to

$$D_{3,\eta}^\pm(x, z'_\pm|k) = \pm \frac{1}{2} e^{jk[(x-x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} \eta]} e^{-\frac{\epsilon_P}{\epsilon_{33}} |k|\eta} \quad (18)$$

with

$$D_3^\pm(x, z'_\pm|k) = \lim_{\eta \rightarrow 0} D_{3,\eta}^\pm(x, z'_\pm|k) = \pm \frac{1}{2} e^{jk(x-x')} \quad (19)$$

Remark. Comparing (19) with (18), the following conclusions can be made. The terms in Eq. (19) are “conventionally” used in interface conditions involving (localized) Dirac’s delta functionals. In contrast, the proposed terms in Eq. (18) arise of “de-localized” Dirac’s delta functions. Further above it was demonstrated that (19) leads to an ill-conditioned Green’s function with a singularity in the near-field zone. It can be expected that (18) induces an η -dependent Green’s function which is well-conditioned and devoid from any singularity in the near-field zone. Consider next the system matrix which would arise in the BEM applications by employing the “regularized,” rather than the singular Green’s function. It can be expected that the η -dependent diagonal- and near-diagonal terms are well-behaved and amenable to numerical calculations. Consequently, extrapolating the results in the limit $\eta \rightarrow 0$, can be expected to lead to regularized self-action and near-field interaction terms. This is the essence of the proposed regularization (renormalization) method.

Remark. To further emphasize the aforementioned conclusions (anticipations) focus on (16). The terms at LHS are the limits of $D_3^\pm(x, z)$ at z'_\pm . On the other hand these limits are equal to $\pm \frac{1}{2} \delta(x - x')$. Furthermore, (16) shows that $\lim_{\eta \rightarrow 0} \delta_\eta^\pm(x - x')$ is equal to $\pm \delta(x - x')$ with:

$$\int_{-\infty}^{\infty} \frac{dk}{2\pi} e^{jk \left[(x-x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} \eta \right]} e^{-\frac{\epsilon_P}{\epsilon_{33}} |k| \eta} = \delta_\eta^\pm(x - x') \quad (20)$$

This equation suggests a parametrized (smeared out) representation for the Dirac’s delta function. This result is for its own sake interesting since its functional form has inherited the properties of the medium in which it was constructed. *It is a problem-specific Dirac’s delta function.*

The major insight which can be deduced from (20) is the fact that instead of exciting the medium with $\delta(x - x')\delta(z - z')$, the medium can be excited by the source functions $\delta_\eta^\pm(x - x')\delta(z - z')$. The implications of this idea may prove to be far reaching. It proposes a constructive process for the problem-specific regularization of singularities.

Remark. Note that the role played by $\delta(z - z')$ in the above formulation was the introduction of the fictitious plane $z = z'$, which demarcates regions I and II.

Denoting the Fourier transforms of $\delta_\eta^\pm(x - x')$ by $\delta_\eta^\pm(k)$, from (20), it can be concluded that:

$$\delta_\eta^\pm(k) = e^{-jk \left(x' \pm \frac{\epsilon_{13}}{\epsilon_{33}} \eta \right)} e^{-\frac{\epsilon_P}{\epsilon_{33}} |k| \eta} \quad (21)$$

Denoting the Fourier transform of $\delta(x - x')$ by $\delta(k)$, the relationship $\delta(k) = e^{-jkx'}$ is valid. A comparison between $\delta(k)$ and $\delta_\eta^\pm(k)$ shows that while the inverse Fourier transform of $\delta(k)$ exists in distributional sense only, the inverse Fourier transforms of $\delta_\eta^\pm(k)$ exist in ordinary sense for any $\eta \neq 0$. Furthermore in the limit $\eta \rightarrow 0$ they converge to $\delta(x - x')$.

This insight is of great conceptual and practical value. In the previous section it was shown that an idealized localized charge distribution with its Fourier transform being of the form $\delta(k) = e^{-jkx'}$, results in a Green’s function which is singular when the observation point coincides with the source point. Whenever Green’s functions are available in closed-form this is not critical since certain regularization (renormalization) techniques can be applied, even though they might not be ideally suited to the problem. The above analysis suggests that replacing the localized line charge distribution $\delta(x - x')$ by de-localized charge distributions $\delta_\eta^\pm(x - x')$ might resolve problems associated with divergence behavior of the problem’s Green’s function.

4. A PROBLEM-TAILORED REGULARIZATION TECHNIQUE

4.1. Distributed Dirac’s Delta Functions and Interface Condition for $D_3(x, z)$

The determination of the regularized Green’s function requires a problem-specific η -parametrized Dirac’s delta function. Consequently, the fields and the involved coefficients are all η -dependent. To appreciate the underlying details relevant equations have been collected as follows:

$$\varphi_\eta^\pm(x, z|k) = \frac{1}{2\epsilon_P |k|} e^{jk \left[(x-x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} \eta \right]} e^{-\frac{\epsilon_P}{\epsilon_{33}} |k| \eta} e^{\gamma^\mp (z-z')} \quad (22a)$$

$$D_{3,\eta}^\pm(x, z|k) = \pm \frac{1}{2} e^{jk \left[(x-x') \mp \frac{\epsilon_{13}}{\epsilon_{33}} \eta \right]} e^{-\frac{\epsilon_P}{\epsilon_{33}} |k| \eta} e^{\gamma^\mp (z-z')} \quad (22b)$$

Using the expressions for γ^\mp , and rearranging lead to:

$$\varphi_\eta^\pm(x, z|k) = \frac{1}{2\varepsilon_p|k|} e^{jk \left[(x-x') \mp \frac{\varepsilon_{13}}{\varepsilon_{33}} (\eta + |z-z'|) \right]} e^{-\frac{\varepsilon_p}{\varepsilon_{33}} |k| (\eta + |z-z'|)} \quad (23a)$$

$$D_{3,\eta}^\pm(x, z|k) = \pm \frac{1}{2} e^{jk \left[(x-x') \mp \frac{\varepsilon_{13}}{\varepsilon_{33}} (\eta + |z-z'|) \right]} e^{-\frac{\varepsilon_p}{\varepsilon_{33}} |k| (\eta + |z-z'|)} \quad (23b)$$

4.2. Regularized Green's Function

Proceeding analogous to (8) and (9), $\varphi_\eta^\pm(x, z|k)$ can be transformed into real space. Thereby, it is instructive to write G_η^\pm rather than φ_η^\pm to emphasize the Green's function nature of the solution. The regularized Green's function is (the argument is positive even for vanishing $x - x'$ and $z - z'$):

$$G_\eta^\pm(x - x', z - z') = -\frac{1}{4\pi\varepsilon_p} \ln \left| \frac{\varepsilon_{33}}{\varepsilon_p} (x - x')^2 \mp 2 \frac{\varepsilon_{13}}{\varepsilon_p} (x - x') [\eta + |z - z'|] + \frac{\varepsilon_{11}}{\varepsilon_p} [\eta + |z - z'|]^2 \right| \quad (24)$$

5. CONCLUSION

A general method for physics-based, problem-tailored regularization of singular Green's functions has been introduced. The method utilizes distributed sources rather than "sharply localized" Dirac's delta functions, as it is customarily the case. The distributed source, chosen for the analysis, was derived from a delta function integral representation, associated with the underlying partial differential equation. The Fourier-type integral representation involved "inhomogeneous" rather than "homogeneous" plane-waves, and was optimally matched to the problem. It remains to be investigated how Dirac delta functions would perform, which are not problem-specific. Author's current work concerns the performance analysis of the constructed regularized Green's functions.

ACKNOWLEDGMENT

This work is based on the research supported in part by the National Research Foundation (UID: 85889). A start-up grant, provided by the Engineering and Built Environment (EBE) Faculty at UCT, is gratefully acknowledged. The initial stage of the work was carried out under an Australian Research Council (ARC) Linkage Grant: LP0775463.

REFERENCES

1. Baghai-Wadji, A. R., "A symbolic procedure for the diagonalization of linear PDEs in accelerated computational engineering," *Symbolic and Numerical Scientific Computation*, F. Winkler and U. Langer, Eds., 347–359, Springer Verlag, 2002.
2. Baghai-Wadji, A. R., "Zooming into the near field: Shear-horizontal polarized acoustic waves," *Proceedings 12th International Symposium on Integrated Circuits, ISIC*, Singapore, 2009.
3. Baghai-Wadji, A. R., "Zooming into the near field: A novel formulation of the BEM as applied to EMC modelling and simulation problems," *Proceedings APEMC*, Beijing, China, 2010.
4. Baghai-Wadji, A. R., "Zooming into the near-fields: Embedding generalized- and pseudo-functions in the solution space of Maxwell's equations," *Proceedings ACES*, Monterey, USA, 2013.

Electromagnetic Sources and Observers in Motion IX — Nature of Gravity

Selwyn E. Wright

Moor Lane Laboratory, ECASS Technologies Ltd, HD8 0QS, UK

Abstract— This is the ninth paper in the series of EM sources and observers in Motion. Two given in Xi'an China, two in Cambridge USA, two in Marrakech Morocco and two in Moscow Russia. This Paper IX is a companion paper to Paper X, the Unification of Electromagnetism and Gravity also presented at this symposium. Gravity has similar properties to the electric field. Both are attractive, their intensity is dependent on the inverse square law with distance, they use the same medium and have a propagation (retarded) time delay. Generally, a distribution of positive and negative charges will produce internally an attractive force and externally a residual difference attractive electric field. This residual field will tend to zero as the distribution size goes to zero. But for a finite distribution this field remains finite, capable of attracting similar multipole charge distributions contained in matter. This appears to be the basis of gravity explaining how dissimilar charges from gravitational matter always attract. Thus unsteady electric (electromagnetic), steady difference electric (gravitational), and residual gravity fields throughout the universe (inertial) are all electrical and use the same propagation medium. This allows the Lorentz transform, gravity and accelerating frames to be compared directly. Einstein's invariant inertial frame, which Einstein based his ether-less predictions, is shown to be non causal. A medium based variant propagation frame supports a Propagation Time Asymmetric model predicting all wave behaviour both classical and EM.

1. INTRODUCTION

Most of the matter in the universe is thought to be comprised of just negatively charged electrons and two types of charged quarks, having fractional values of charge q , an up quark 'u' ($q = 2/3$) and a down quark 'd' ($q = -1/3$). The main field contributors are therefore the proton $q = u + u + d = 2/3 + 2/3 - 1/3 = 1$ and the electron $q = -1$. Even possibly the neutron $q = d + d + u = -1/3 - 1/3 + 2/3 = 0$, if its quarks, under high pressure, can be reconstructed into positive and negative charges. Using a point atomic model an equal number of negative electrons and positive protons, with their corresponding electric fields cancel, creating a neutral atom. However, a finite distribution of dissimilar charges will always produce an attractive steady residual difference field. This can be easily demonstrated with dissimilar magnetic or charged poles, as illustrated in Figure 1. These basic fields appears to play an important part in the attraction of atoms and molecules and therefore gravity.

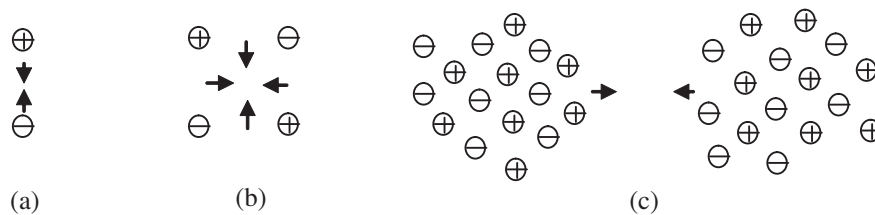


Figure 1: A finite distribution of dissimilar charges will always attract, forming the basis of gravity. (a) Dipole. (b) Quadrupole. (c) Clouds of multipoles.

2. ATOMIC RESIDUAL DIFFERENCE FIELD (ARDF)

According to [1] induced dipoles from molecules with permanent bipolar fields and induced dipoles, causes an attraction between individual atoms and molecules. These attractive dipole fields, in the near field, determine their liquid and gaseous phases and the pulling together of mercury beads and soap films. They are local, having a very rapidly reducing field strength with distance. Also [8], predicted that electrons orbiting at large distances around atomic nuclei repel (displace) electron orbits of neighbouring atoms and molecules. These unbalanced orbits then create a mass/spring induced dipole field, oscillating at imaginary non propagating frequencies, resulting in a rapidly

decaying reactive field with distance. However, low atomic number molecules can have high and low energy levels. These result in strong and weak attractive near fields creating both solids and gases with similar atomic numbers, i.e., their attraction is not proportional to the atomic number.

Although these van der Waals type attractive fields are electrostatic, they do not appear to be gravitational, they are local fields holding the molecules together. For gravity we are looking for attractive electrical dipole fields that propagate into the radiation far field, i.e., decaying according to the inverse square law and whose strength (weight) tends to be proportional to the atomic number. The model that best fits this description is the Atomic Residual Difference Field (ARDF) whose attractive field is proportional to the atomic number (number of pairs of dissimilar charges in the atom). From quantum mechanical grounds, the difference residual field should form in the reactive near field, at the atomic level. The field summation from all atoms and molecules in the vicinity would then be attractive and decay into the propagation far field according to the inverse square law, the same as gravity. Equating the electric field E , given by $E = kq^*r'/r^2$ where $k = 8.99 \times 10^9 \text{ Nm}^2/\text{C}^2$ and q^* is an attractive dissimilar charge in Coulombs. And the gravity field g , given by $g = Gmr'/r^2$ where $G = 6.67 \times 10^{-11} \text{ Nm}^2/\text{kg}^2$, and the attractive mass m is in kg. Also r' is the unit vector along the distance radius r . Therefore gravitational mass has an equivalent charge of:

$$q^* = (G/k)m = (6.67 \times 10^{-11}/8.99 \times 10^9)m = (0.74 \times 10^{-20})m \quad (1)$$

which creates a potentially very weak attractive field. Other equivalents are then:

$$\begin{aligned} g &\equiv E, \text{ force } F = mg \equiv q^*E, \text{ potential } v = \int gdr \equiv \int Edr, \\ \text{energy } U &= Fdr = \int mgdr = mv \equiv \int q^*Edr = q^*v \end{aligned} \quad (2)$$

3. UNIVERSAL GRAVITATIONAL REFERENCE FIELD (UGRF)

Whatever the actual details of the ARDF, it can be considered to be finite from the almost limitless gravitational matter, seen and unseen throughout the universe. Although this field reduces with the inverse square of the distance from its gravitational source, the contribution at any point in the universe, from all known and inferred gravitational matter q^* , of equivalent volume charge density ρ^* lies on an expanding sphere whose surface area increases as the square of the distance ($4\pi r^2$). The modulus of the total field (inertial field) contributing to that point then becomes:

$$|E| = k\Sigma|q^*|/r^2 \text{ where } \Sigma|q^*| = \int(\rho^*4\pi r^2)dr, |E| = \int(k\rho^*4\pi)dr = k\rho^*4\pi r, \text{ where } r = \infty \quad (3)$$

Only through symmetry does the ARDF tend to zero, $E = 0$. but the modulus of the inertial field $|E|$ is *enormous*. Thus from the total matter in the universe, the ARDF provides a non zero Universal Gravitational Reference Field (UGRF) over all space. This scalar potential energy field is available to resist motional change of mass (energy) in motion, creating mass inertia. The $|E|$ field is also a measure of the magnitude of the attractive field (attractive ‘dark’ energy) slowing down the gravitational mass expansion in the universe, relative to the propagation medium. The medium could then provide complimentary repulsive ‘dark’ energy. This could overcome the global gravitational attraction, through the repulsive expansion of the medium, accelerating further the gravitational matter of the universe. Unlike unsteady EM and gravity waves, the UGRF, from steady charges throughout the universe behaves as a steady field. The ARDF, being a difference E field, has no mass and spin 1. Whereas, the UGRF, being a resultant scalar potential field from difference E fields from all directions across the universe, has zero intensity (force), no mass and no spin, making it difficult to detect, but it has finite potential energy. This field resembles the Higgs field [9].

4. INERTIAL AND GRAVITATIONAL MASS

As Einstein kept reminding us, all forms of energy (potential and kinetic) are equivalent to inertial mass. It appears that inertial mass of subatomic particles, complete atoms and molecules, results through the binding energy of the system, the energy required to create the structure. For example, in the proton only 2% of its binding mass is from its individual components, the other 98% is accountable through the strong binding force (energy) of the gluons holding the proton together. Inertia is the resistance to accelerating binding mass. As in electrons in orbit around an atom and J. J. Thompson’s e/m experiment [3], where gravity has no part, only inertial mass (no g or G constants involved).

It appears that it is not the inertial or binding mass m_b that causes gravity. It is a quantity of energy, having no substance or solidity in itself and no attractive capability. Attraction of matter (gravity), appears to be the Atomic Residual Difference Fields (ARDF) between finite distribution of dissimilar charges within gravitational matter (atoms and molecules) having a total gravitational mass m_g . Here $m_b = m_g$, but m_b can exist without m_g , as in the electron having mass but no attraction. Although neutron stars are nominally neutral, it appears, under the huge compacting pressures, their gravitational matter, including neutrons (which are comprised of positive and negative fractional charges) must undergo some kind of charge reconstitution of their atomic quarks into positive and negative free charges (currents), creating their intense magnetic and gravitational fields.

5. RELATION BETWEEN LORENTZ, GRAVITY AND ACCELERATION

Now that gravity (steady difference electric) and light (unsteady electric) are both parts of the same field, sharing the same propagation medium. It provides a common link between Lorentz's transform, gravity and accelerating frames, as the following examples show.

5.1. Constant Motion

According to [4], a system moving at constant velocity ' v_s ' and interacting with the medium, contracts its structural time τ_s and space x_s , compared to the propagation medium values τ_p and x_p . Thus:

$$\tau_s = \alpha_s \tau_p, \quad x_s = \alpha_s x_p, \quad \alpha_s = (1 - M_s^2)^{1/2}, \quad M_s = v_s/c \quad (4)$$

M_s (or β_s) is the system (source) [2] number. For small speed changes $M_s \ll 1$, fractional changes are:

$$\Delta\tau_s/\Delta\tau_p = \Delta\alpha_s = -M_s^2/2, \quad \Delta x_s/\Delta x_p = \Delta\alpha_s = -M_s^2/2 \quad (5)$$

For example, the amount of time slowing for a passenger travelling on a modern airliner for 24 hours relative to the stationary medium surrounding the Earth, where $v_s = 300$ m/s, (670 miles/hour), $c = 3 \times 10^8$, $M_s = 10^{-6}$, time slowing (24 hours) $\Delta\tau_s = -(10^{-12}/2)\Delta\tau_p = -(10^{-12}/2) \times 8.6 \times 10^4 = -43$ ns/day. Time slowing at Earth's rotational speed at equator (460 m/s) is -95 ns/day relative to surrounding stationary medium. Flying supersonically (460 m/s, $M = 1.5 \times 10^{-6}$) in the direction of the Earth's rotation (920 m/s relative to surrounding stationary medium) is $-4 \times 95 = -380$ ns/day and relative to the Earth's surface $-3 \times 95 = -285$ ns/day. Flying against the Earth's rotation (0 m/s relative to the surrounding stationary medium) is 0s/day, and $+95$ ns/day relative to the Earth's equator.

5.2. Gravitational Acceleration

According to Schwarzschild's dominant term [6], the propagation medium and its contents in a gravi-tational field ' g ' is compressed both time τ_g and space x_g , compared to their free-field values τ_p and x_p . Thus:

$$\begin{aligned} \tau_g &= \delta_g \tau_p, \quad x_g = \delta_g x_p, \quad \delta_g = (1 - M_g)^{1/2}, \\ M_g &= v_g/c = 2Gm/c^2 R_g = 2gR_g/c^2 \text{ as } g = Gm/R_g^2 \text{ and } v_g = 2gR_g/c \end{aligned} \quad (6)$$

For $M_g \ll 1$, fractional changes are:

$$\Delta\tau_g/\Delta\tau_p = \Delta\delta_g = -M_g/2, \quad \Delta x_g/\Delta x_p = \Delta\delta_g = -M_g/2 \quad (7)$$

For example, the amount of time slowing on the Earth's surface due to the Earth's gravity compared with free space, where $G = 6.67 \times 10^{-11}$ Nm²/kg² is the gravitational constant, $m = 6 \times 10^{24}$ kg is the gravitational mass of the Earth, $c = 3 \times 10^8$ m/s, $g = 9.8$ m/s², $G/c^2 = 7.4 \times 10^{-28}$, $R_g = 6.4 \times 10^6$ is the radial distance from the Earth's centre, gives $M_g \approx 1.3 \times 10^{-9}$ for the Earth's surface equivalent gravitational Mach number.

For an accumulative time (24 hours) $\Delta\tau_g = -(M_g/2)\Delta\tau_p = -(1.3 \times 10^{-9}/2)(8.6 \cdot 10^4) = -56$ μ s/day.

5.3. Observer Acceleration

For an accelerating observer a_o moving relative to the propagation medium, over a time t_o , and distance d_o and wave propagation distance d_w we have:

$$\begin{aligned} \Delta\tau_o/\Delta\tau_p &= \Delta\delta_o = M_o, \quad M_o = v_o/c = a_o t_w/c = a_o d_w/c^2, \\ \text{where } v_o &= a_o t_w, \quad t_w = d_w/c, \quad d_o = a_o t_w^2/2 \end{aligned} \quad (8)$$

For example, the special case of an observer falling under gravity. M_o is positive, observer's surrounding time and space relatively expand, neutralizing gravity's time loss. According to the Equivalence Principle $M_o = M_g/2$, for $M_o = 1.3 \times 10^{-9}/2$, time quickens $\Delta\tau_o = M_o\Delta\tau_p = (1.3 \times 10^{-9}/2)$ (8.6×10^4) = 56 $\mu\text{s}/\text{day}$ neutralizing the Earth's gravity time slowing. Also $a_o = 9.8 \text{ m/s}^2$, wave distance $d_w = M_o c^2/a_o = 6 \times 10^6 \text{ m}$, $t_w = d_w/c = 2 \times 10^{-2} \text{ s}$, observer velocity $v_o = a_o t_w = 2 \times 10^{-1} \text{ m/s}$ and observer distance $d_o = at_w^2/2 = 10 \times 4 \times 10^{-2}/2 = 0.2 \text{ m}$.

6. BASIC WAVE PROPAGATION

Before considering unification theories, considered in Paper X, it appears [5] believing there was no propagation medium, misunderstood the basic wave propagation process. It is true that Einstein's inertial frame, where the physics and speed of light are invariant, is non preferred (not unique). However, his ether-less frame cannot predict observed motion as Einstein believed. *To observe events from Einstein's inertial frame requires the propagation medium to be restored to allow the events to reach the observer.* Thus the ether-less, non preferred, omni-directional, inertial frame, plus the directional wave propagation, becomes the medium based, preferred, directional, reference frame, which propagates the source events to the far field. In spite of Einstein's uniqueness claim, there is nothing special regarding EM waves, they require a propagation medium the same as all causal wave theories.

7. PROPAGATION TIME ASYMMETRY

The effects of the preferred propagation frame is displayed in Figure 2. These are of course variant, not invariant as Einstein's inertial frame describes. It shows the reality of wave propagation that everyone is familiar with, illustrating Propagation Time Asymmetry (PTA) around a moving source. These motional directional waveforms, based on a propagation medium, are representative of all types of classical and EM wave propagation. It's irrational to accept that EM sources are somehow unique, they no longer require a propagation medium when set in motion. At low speeds the waveforms are dominated by classical PTA wave generation. Einstein's ether-less relativity, without a medium, cannot predict PTA. The conventional Galilean concept of PTA was extended by Lorentz, his Transform (LT) includes PTA as an integral part. It provides the foundation for Einstein's medium based causal SR.

For $M \ll 1$, the classical asymmetrical propagation displacement Δd , and time displacement Δt through motion, is given by:

$$\begin{aligned} \Delta t &= \pm \Delta d/c = \pm vt/c = \pm (vd/c)/c = \pm vd/c^2 = \pm Mt, \\ M &= v/c, \quad t = d/c, \quad \Delta t/t = \Delta d/d = M \end{aligned} \quad (9)$$

where v is the system velocity, \pm is for motion in and against the direction of the light propagation, and d is the propagation path between two fixed points moving with the system relative to the medium. Figure 2(a) is for a single frequency and system Mach number $M = v/c = 0.66$. The PTA wave pattern is virtually identical for classical and EM sources, for the same M . Figure 2(b) is a wave pattern generated by the PTA part of the Lorentz transform for an electron moving for

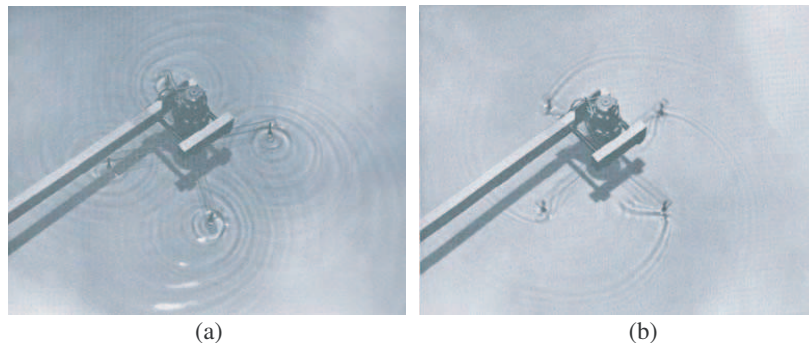


Figure 2: These Propagation Time Asymmetric (PTA) wave patterns, generated by all forms of source motion, including classical and EM sources, are caused by motion with respect to the propagation medium. These patterns, which are a result of a preferred frame of reference cannot exist without a propagation medium. The same basic theory predicts both classical and EM waves. (a) Unsteady sources in motion. (b) Steady sources exceeding wave speed.

$M > 1$. Again, the classical medium based PTA predicts the main structure for an impulsive wave front from an electron exceeding its wave speed in the medium, as in [7] radiation. Here a high speed electron leaves a vacuum and enters an electrically more dense medium (air).

On Earth, motional effects, without involving particle accelerators or particles from outer space, are predicted by the medium based classical wave theory, i.e., just the PTA part of the LT. Relativistic and gravitational contributions have little effect on the instantaneous PTA disturbance. Assuming realistic Earth speeds of $M_s \approx 10^{-6}$ and Earth equivalent gravitational speed of $M_g \approx 10^{-9}$, from Equations (5), (7) and (9), we have for the time ratios, compared to the unaffected values, Lorentz $\Delta\tau_s/\Delta\tau_p = M_s^2/2 \approx 10^{-12}/2$, gravity $\Delta\tau_g/\Delta\tau_p = M_g/2 \approx 10^{-9}/2$ and PTA $\Delta t_{pta}/t_p = M_s \approx 10^{-6}$. It is the fundamental PTA wave disturbance that is the dominant feature. The relativistic and gravitational effects cannot act alone; they are a small modifying effect of PTA. *At speeds and gravitational strengths on and around the Earth, and small measurement times, they cannot explain basic wave propagation, as sometimes claimed.*

8. CONCLUSION

Light is a varying E field, for example like charges (electrons) jumping orbit in atoms and molecules. The Atomic Residual Difference Field (ARDF) is the steady difference E field from dissimilar charge distributions, primarily from protons and electrons. Gravity is the ARDF from a large number of atoms and molecules. The Universal Gravitational Reference field (UGRF) is the ARDF, summed at a point from all over space, from the total gravitational matter in the universe. The UGRF provides a net zero force but a non zero potential energy reference field in the medium at rest in space, providing inertia for mass in motion.

All these effects have a common basis i.e., they are all electrical, they all depend on the electric charge and a common propagation medium. Sources, according to Lorentz, moving through the medium contract their time and structure. The medium, according to Schwarzschild, is compressed both time and space, by gravity. An accelerating observer contracts its time and structure, relatively expanding the surrounding medium with the ability of reducing gravity. Einstein's invariant, omnidirectional, inertial frame, upon which he based his ether-less predictions, is shown to be non causal. It is the Propagation Time Asymmetry (PTA), based on a propagation medium, which characterises the directionality around all sources in motion, both classical and electromagnetic.

REFERENCES

1. Van Der Waals, J. D., "Over de continuïteit van den gasen vloeistofoestand (on the continuity of the gas and liquid state)," Ph.D. thesis (excerpt), Leiden, The Netherlands, 1873.
2. Mach, E., *Mach's Principle: Mach's Archive*, Deutches Museum Munich, 1887.
3. Thomson, J. J., "Cathode rays," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Fifth Series, 296, Institute of Physics Pub., Bristol, 1897.
4. Lorentz, H. A., "Simplified theory of electrical and optical phenomena in moving systems," *Proc. Acad. Science Amsterdam*, Vol. 1, 427–442, 1899.
5. Einstein, A., "On the electrodynamics of moving bodies," *Annalen der Physik*, Vol. 17, 891–921, 1905.
6. Schwarzschild, K., "Über das gravitationsfeld eines massenpunktes nach der einstein'schen theorie," *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften*, Vol. 1, 189–196, 1915.
7. Cerenkov, P. A., "Visible emission of clean liquids by action of γ radiation," *Doklady Akad. Nauk SSSR*, Vol. 2, 451, 1934.
8. London, B. F., "The general theory of molecular forces," *Transactions of the Faraday Society*, Vol. 33, 8–26, 1937.
9. Higgs, P. W., "Broken symmetries, massless particles and gauge fields," *Phys. Lett.*, Vol. 12, 132, 1964.

The Uniqueness Theorems in the Electromagnetic Wave Theory and Quasi-periodical Solutions of the Periodical Diffraction Problems

N. B. Pleshchinskii

Kazan Federal University, Russia

Abstract— The simple method to prove the uniqueness of the solution of the electromagnetic wave diffraction problems is proposed in the case of loss-free media. The set of the diffraction problems on the thin conducting screens in the wave-guided structures are considered. The over-determined boundary value problem method is used. It is shown that the limited absorption principle is correct for the considered problems.

The uniqueness conditions of the solution of the integral equations with difference kernel are obtained as an auxiliary result. These conditions are used to prove the Floquet theorem: the solution of the quasi-periodical wave diffraction on the periodical set of the heterogeneities can be the quasi-periodical wave (Floquet wave) only.

The case of diffraction problem on the bi-periodical set of thin conducting screens in the opened space is considered in detail as an example. The boundary value problem is equivalent to the dual summatorial functional equation for the Floquet coefficients. It is proved that this problem is equivalent to the regular infinite set of the linear algebraic equations for the coefficients of decomposition of the electromagnetic field by Floquet harmonics. By this set of equations the algorithms are constructed for numerical solving the diffraction problem.

1. INTRODUCTION

Solution uniqueness theorems for diffraction problems of electromagnetic waves on heterogeneities of different types are used as for justification of correctness of their mathematical formulations, so for proof of convergence of approximate methods of solving such problems. As it is generally known, the diffraction problems can have only one solution only in the case when a radiation condition presents in the formulation of the problem. In present work the supposition of orientability of unknown solution is used as a radiation condition.

Let us call at electromagnetic wave positively-oriented in relation to the axis of waveguide structure, if it attenuates in direction of this axis or transfers energy in this direction.

It is expedient to use the over-determined boundary value problem method [1] (see also [2]) for reducing the diffraction problem of electromagnetic waves on the systems of thin conducting screens. This method is a variant of general partial domain method. As auxiliary boundary value problems such problems are considered in separate parts of waveguide structure, when more conditions than it is necessary for determination of the unique solution of the Maxwell set of equations (or of Helmholtz equation in two-dimensional case) are given on the boundary of domain. It is convenient to use the necessary and sufficient conditions of solvability of the over-determined problems for reducing conjugation problems to the one-sided boundary value problems or directly to integral equations.

It is proved that these solvability conditions can be used to prove the uniqueness of solution of the diffraction problems.

In the case, when screens form the periodic systems of heterogeneities, an important statement follows from the uniqueness theorem: if the quasi-periodical wave (in particular case — plane wave) falls down periodic system of screens, then by its diffraction the quasi-periodical field appears also. This statement is formulation of Floquet theorem that is often used for research of the electromagnetic fields in periodic structures. But proof of this obvious from the physical point of view fact is usually absent.

2. DIFFRACTION OF ELECTROMAGNETIC WAVE ON CONDUCTING THIN STRIPS

Two-dimensional diffraction problem of TE-polarized electromagnetic wave on plane system of metallic strips is being reduced to the conjugation problem for Helmholtz equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2} + k^2 u(x, z) = 0, \quad (1)$$

k is a real-valued number.

Let us denote by $u_0^\pm(x)$ and $u_1^\pm(x)$ the limiting values (traces) of functions $u(x, z)$ and its derivative $(\partial u/\partial z)(x, z)$ on the axis $z = 0$ from above and from lower semi-planes accordingly. We will search the positively-determined solution of Equation (1) in the above semi-plane $z > 0$ by given traces $u_0^+(x)$ and $u_1^+(x)$ on the line $z = 0$. This problem is over-determined.

Suppose that dependence of the field components on time has the form $e^{i\omega t}$. Denote

$$\gamma(\xi) = \sqrt{k^2 - \xi^2}, \quad \operatorname{Re} \gamma(\xi) \leq 0 \quad \text{or} \quad \operatorname{Im} \gamma(\xi) \geq 0.$$

It is proved [3] that $u_0^+(x)$ and $u_1^+(x)$ are traces on the line $z = 0$ of positively-oriented solution of Helmholtz equation in a semi-plane $z > 0$ if and only if, when

$$u_1^+(\xi) - i\gamma(\xi) u_0^+(\xi) = 0. \quad (2)$$

In case of the negatively-oriented solutions the solvability condition of the over-determined boundary value problem for Helmholtz equation in a lower-half-plane $z < 0$ has the form

$$u_1^-(\xi) + i\gamma(\xi) u_0^-(\xi) = 0. \quad (3)$$

Let the system of segments \mathcal{M} of axis x correspond to the metallic strips in the section $y = 0$ of opened space and \mathcal{N} is a supplement \mathcal{N} to whole axis. Let $u^0(x, z)$ be potential function of the external field.

The next conjugation problem corresponds to the diffraction problem. It is necessary to find the positively-oriented solution $u^+(x, z)$ of Helmholtz equation for $z > 0$ and the negatively-oriented solution $u^-(x, z)$ of Helmholtz equation for $z < 0$ satisfying for $z = 0$ conditions:

$$u_0^+(x) = -u_0^0(x), \quad u_0^-(x) = -u_0^0(x), \quad x \in \mathcal{M}, \quad (4)$$

$$u_0^+(x) = u_0^-(x), \quad u_1^+(x) = u_1^-(x), \quad x \in \mathcal{N}. \quad (5)$$

It is easy to get by Equalities (2) and (3) that *the diffraction problem is equivalent to the next one-sided boundary value problem: it is necessary to find the positively-oriented solution $u^+(x, z)$ of Helmholtz equation for $z > 0$ satisfying the conditions*

$$u_0^+(x) = -u_0^0(x), \quad x \in \mathcal{M}, \quad u_1^+(x) = 0, \quad x \in \mathcal{N}. \quad (6)$$

Two-dimensional diffraction problem of TE-polarized electromagnetic wave on thin conducting strips can have only one solution.

Proof. The difference of two solutions of problem should satisfy boundary conditions

$$u_0^+(x) = 0, \quad x \in \mathcal{M}, \quad u_1^+(x) = 0, \quad x \in \mathcal{N}.$$

Then

$$u_0^{+*}(x) u_1^+(x) = 0, \quad x \in (-\infty, +\infty).$$

We will integrate this pair equality by the whole axis and then we get

$$\int u_0^{+*}(x) u_1^+(x) dx = 0.$$

It follows by Parseval equality that

$$\int u_0^{+*}(\xi) u_1^+(\xi) d\xi = 0.$$

But by condition (2)

$$\int \gamma(\xi) |u_0^+(\xi)|^2 d\xi = 0.$$

Function $\gamma(\xi)$ has negative real part on one part of the axis ξ , and it has positive imaginary part on the other part of the axis. Therefore $u_0^+(\xi) \equiv 0$.

3. INTEGRAL EQUATION WITH A DIFFERENCE KERNEL

By boundary conditions (6) we have that *the diffraction problem is equivalent to integral equation with a difference kernel*

$$\int_{\mathcal{M}} u_1^+(t) K_{-1}(t-x) dt = -u_0^0(x), \quad x \in \mathcal{M}, \quad \text{where} \quad K_{-1}(t-x) = \frac{-i}{2\pi} \int \frac{1}{\gamma(\xi)} e^{i\xi(t-x)} d\xi \quad (7)$$

(integral can be easily expressed by Hankel function).

Really, let $u_1^+(x)$ on \mathcal{M} be the unknown function and, consequently,

$$u_1^+(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathcal{M}} u_1^+(t) e^{it\xi} dt.$$

Then

$$u_0^+(x) = \frac{1}{\sqrt{2\pi}} \int u_0^+(\xi) e^{-i\xi x} d\xi = \frac{1}{\sqrt{2\pi}} \int \frac{-i}{\gamma(\xi)} u_1^+(\xi) e^{-i\xi x} d\xi = \int_{\mathcal{M}} u_1^+(t) \left(\frac{-i}{2\pi} \int \frac{1}{\gamma(\xi)} e^{i\xi(t-x)} d\xi \right) dt.$$

Let us consider more general integral equation of form

$$\int_{\mathcal{M}} u(t) K(t-x) dt = f(x), \quad x \in \mathcal{M}, \quad \text{where} \quad K(t-x) = \frac{1}{2\pi} \int A(\xi) e^{i\xi(t-x)} d\xi \quad (8)$$

(function $K(x)$ is Fourier prototype of function $A(\xi)$ within the constant multiplier precision).

If real part of function $A(\xi)$ saves a sign on some part of axes ξ and on other part of axis imaginary part of this function saves a sign, then integral equation with a difference kernel can have one solution only.

Proof. Let $u_1(x)$ be a solution of integral equation redefined by zero to the whole axis, and

$$u_0(x) = \int u_1(t) K(t-x) dt.$$

Let us transform the right-hand side of this equality:

$$u_0(x) = \int u_1(t) \left(\frac{1}{2\pi} \int A(\xi) e^{i\xi(t-x)} d\xi \right) dt = \frac{1}{\sqrt{2\pi}} \int A(\xi) u_1(\xi) e^{-i\xi x} d\xi.$$

By this $u_0(\xi) = A(\xi) u_1(\xi)$.

For the difference $u(x)$ of two solutions of equation we have $u_0(x) = 0$ on \mathcal{M} and $u_1^*(x) = 0$ on \mathcal{N} . Then

$$\int A(\xi) |u_1(\xi)|^2 d\xi = 0.$$

4. QUASI-PERIODICAL SOLUTIONS OF INTEGRAL EQUATIONS AND BOUNDARY VALUE PROBLEMS

Let \mathcal{M} be l -periodical system of segments of real axis. Let $g(x)$ be l -periodical function. If integral equation with a difference kernel

$$\int_{\mathcal{M}} v(t) L(t-x) dt = g(x), \quad x \in \mathcal{M},$$

has only one solution then this solution is a l -periodical function.

For proof let us substitute the left part of equation $v(t+l)$ instead of $v(t)$:

$$\int_{\mathcal{M}} v(t+l) L(t-x) dt = \int_{\mathcal{M}} v(t+l) L(t+l-x-l) dt = \int_{\mathcal{M}} v(t_1) L(t_1-x-l) dt_1 = g(x+l) = g(x), \quad x \in \mathcal{M}.$$

As solution can be only one, then $v(t + l) = v(t)$.

Let us consider more general case. The function $u(x)$ is called (l, α) -quasi-periodical, if $u(x) = e^{i\alpha x}v(x)$, where $v(x)$ is l -periodical function.

Let $f(x)$ be (l, α) -quasi-periodical function. If integral equation with a difference kernel

$$\int_{\mathcal{M}} u(t) K(t - x) dt = f(x), \quad x \in \mathcal{M}, \quad (9)$$

has only one solution, then this solution is (l, α) -quasi-periodical function.

Proof. Let $f(x) = e^{i\alpha x}g(x)$ and $u(x) = e^{i\alpha x}v(x)$. By supposition equation

$$\int_{\mathcal{M}} e^{i\alpha t} v(t) K(t - x) dt = e^{i\alpha x} g(x), \quad x \in \mathcal{M},$$

has only one solution. Consequently, equation

$$\int_{\mathcal{M}} v(t) L(t - x) dt = g(x), \quad x \in \mathcal{M},$$

where $L(t) = e^{i\alpha t}K(t)$, has only one solution also. Then $v(t)$ is a periodical function.

Consequently, solution of diffraction problem of quasi-periodical wave on the periodic system of strips can be a quasi-periodical wave only.

5. THREE-DIMENSIONAL DIFFRACTION PROBLEMS ON THE SYSTEM OF SCREENS

Traces of difference of two solutions of diffraction problem in half-space in three-dimensional case should satisfy boundary conditions

$$e_x(x, y) = 0, \quad e_y(x, y) = 0, \quad (x, y) \in \mathcal{M}, \quad h_x(x, y) = 0, \quad h_y(x, y) = 0, \quad (x, y) \in \mathcal{N}. \quad (10)$$

Connections between their Fourier transforms were obtained in [4]:

$$\begin{aligned} \omega\mu_0\mu\gamma(\xi, \eta) h_x(\xi, \eta) &= -\xi\eta e_x(\xi, \eta) + (\xi^2 - k^2) e_y(\xi, \eta), \\ \omega\mu_0\mu\gamma(\xi, \eta) h_y(\xi, \eta) &= (k^2 - \eta^2) e_x(\xi, \eta) + \xi\eta e_y(\xi, \eta). \end{aligned} \quad (11)$$

Expression $-\xi\eta e_x(\xi, \eta) + (\xi^2 - k^2) e_y(\xi, \eta)$ is a Fourier transform of differential expression

$$\frac{\partial^2 e_x}{\partial x \partial y} - \left(\frac{\partial^2}{\partial x^2} + k^2 \right) e_y.$$

If $e_x(x, y) = 0, \quad e_y(x, y) = 0$ on \mathcal{M} , then

$$\frac{\partial^2 e_x}{\partial x \partial y} - \left(\frac{\partial^2}{\partial x^2} + k^2 \right) e_y = 0 \quad \text{and} \quad \left[\frac{\partial^2 e_x}{\partial x \partial y} - \left(\frac{\partial^2}{\partial x^2} + k^2 \right) e_y \right] \cdot h_x^*(x, y) = 0$$

both on \mathcal{M} and on \mathcal{N} .

Let us integrate by whole plane and pass to the Fourier transforms. Then we get

$$\int \int [-\xi\eta e_x(\xi, \eta) + (\xi^2 - k^2) e_y(\xi, \eta)] \cdot h_x^*(\xi, \eta) d\xi d\eta = 0$$

or

$$\int \int \gamma(\xi, \eta) |h_x(\xi, \eta)|^2 d\xi d\eta = 0.$$

By this $h_x(\xi, \eta) \equiv 0$. By analogy we can obtain that $h_y(\xi, \eta) \equiv 0$.

Let the system of thin conducting screens be periodic on coordinate x with the period p and periodic on coordinate y with period q . Let us denote by \mathcal{M} the screens placed in a rectangle of the periods and by \mathcal{N} another part of this rectangle.

Quasi-periodic functions with two periods have form

$$A(x, y, z) = \sum_m \sum_n A_{m,n}(z) e^{ip_mx+iq_ny}, \quad p_m = \frac{2\pi}{p}m + \alpha, \quad q_n = \frac{2\pi}{q}n + \beta,$$

α, β are Floquet parameters.

By the uniqueness theorem of solution of the diffraction problem we have that *if the quasi-periodic wave falls down on system of screens, then field appearing at its diffraction is quasi-periodic also.*

It is convenient to choose as potential functions of an electromagnetic field components of E_x and E_y of the electric vector. Wave diffraction problem on bi-periodic system of screens is reduced to the following problem: it is necessary to look for Floquet coefficients of these functions $a_{n,m}$ and $b_{n,m}$ by boundary conditions

$$E_x^+ = -E_x^0, \quad E_y^+ = -E_y^0, \quad E_x^- = -E_x^0, \quad E_y^- = -E_y^0 \quad \mathcal{M} \quad (12)$$

and by conjugation conditions on a plane $z = 0$

$$E_x^+ = E_x^-, \quad E_y^+ = E_y^-, \quad H_x^+ = H_x^-, \quad H_y^+ = H_y^- \quad \mathcal{N}. \quad (13)$$

It is easy to see that $a_{m,n}^+ = a_{m,n}^- = a_{m,n}$, $b_{m,n}^+ = b_{m,n}^- = b_{m,n}$. It is convenient to pass to unknown vectors-lines $c_{m,n} = (a_{m,n}, b_{m,n})$. Then the diffraction problem is reduced to the pair vector summatorial equation

$$\begin{aligned} \sum_m \sum_n c_{m,n} e^{i\frac{2\pi}{p}mx+i\frac{2\pi}{q}ny} &= - \sum_m \sum_n c_{m,n}^0 e^{i\frac{2\pi}{p}mx+i\frac{2\pi}{q}ny} \quad \text{on } \mathcal{M}, \\ \sum_m \sum_n c_{m,n} \Gamma_{m,n} e^{i\frac{2\pi}{p}mx+i\frac{2\pi}{q}ny} &= 0 \quad \text{on } \mathcal{N}, \\ \Gamma_{m,n} &= \begin{pmatrix} p_m q_n & k^2 - q_n^2 \\ k^2 - p_m^2 & p_m q_n \end{pmatrix}. \end{aligned} \quad (14)$$

Here $c_{m,n}^0$ are vector coefficients of the external field.

It is possible to pass to regular infinite set of the linear algebraic coefficients with respect to unknown Floquet coefficients by method of integral-summatorial identities (see, for example, [2])

The algorithm of the numerical solving of the wave diffraction problem on doubly-periodic system of screens is constructed based on a reduction method of the infinite linear set of equations. Numerical experiment is made on GPU with CUDA technology use.

ACKNOWLEDGMENT

Supported by RFBR 12-01-97012-r_povolzhé.a.

REFERENCES

1. Pleshchinskii, N. B., "The over-determined boundary value problems for the Maxwell equations set in the orthogonal coordinates and some applications for the electromagnetic wave diffraction problems," *this Volume*.
2. Pleshchinskii, N. B., I. E. Pleshchinskaya, and E. M. Karchevskiy, "The over-determined boundary value problem method in the electromagnetic waves propagation and diffraction theory," *PIERS Online*, Vol. 5, No. 5, 441–445, 2009.
3. Pleshchinskaya, I. E. and N. B. Pleshchinskii, "Over-determined boundary value problems for elliptic partial differential equations and their application to waves diffraction theory," *Uchenye zapiski Kazanskogo gosudarstvennogo universiteta*, Vol. 147, No. 3, 41–32, 2005 (in Russian).
4. Pleshchinskaya, I. E. and N. B. Pleshchinskii, "The over-determined Cauchy problems for the Maxwell equations set and electromagnetic waves diffraction on the metallic screens," *Conf. Proc. 11th Int. Conf. Mathematical Methods in Electromagnetic Theory MMET*06*, 255–257, Kharkov, Ukraine, June 2006.

The Over-determined Boundary Value Problems for the Maxwell Equations Set in the Orthogonal Coordinates and Some Applications for the Electromagnetic Wave Diffraction Problems

N. B. Pleshchinskii

Kazan State University, Russia

Abstract— The necessary and sufficient conditions of solvability of the over-determined problems for the Maxwell equations set are obtained in the cases of Cartesian, cylindrical and spherical coordinates. These conditions are used by reduction of the electromagnetic wave diffraction problems on the thin conducting screen placed on the coordinate surfaces to integral or summatorial equations. Moreover, these conditions can be used for regularization of the integral equations of the first kind obtained by solving the diffraction problems.

In the case on three-dimensional problems of electrodynamics the over-determined problems appear when both the tangential components of the electric vector and of the magnetic vector are given on the boundary of the partial domains.

It is shown that the conditions of solvability of the over-determined problems can be obtained in the different form. The common form is the dependence of the Fourier transforms or Fourier coefficients of the boundary functions. But the mixed form is possible when both expressions participate.

The set of the three-dimensional diffraction problems on the thin conducting screens in the wave-guided structures is considered as examples.

1. INTRODUCTION

The electromagnetic wave diffraction problem on ideally conducting infinitely thin screen in three-dimensional space is formulated as a conjugation problem for the Maxwell set of equations. Let screen \mathcal{M} be a part of surface \mathcal{S} , dividing space into two parts, and \mathcal{N} be a supplement of \mathcal{M} to \mathcal{S} . Traces of the field components on \mathcal{M} and on \mathcal{N} should satisfy the different conditions: sums of tangent components of vectors E of the external field and of the unknown field should be equal to zero from every side \mathcal{M} , and tangent components of vectors E and H should be continuous on \mathcal{N} .

In order to reduce the diffraction problem to the integral equation on \mathcal{M} or to some its analogues by the partial domain method it is necessary preliminary to build the solution of the Maxwell set of equations in each of two half-spaces with a common border \mathcal{S} . It is expedient to use the over-determined boundary value problems for the Maxwell set of equations as auxiliary boundary value problems, when tangent components of vector E and of vector H are given simultaneously on \mathcal{S} .

The solution of the over-determined problem exists if and only if, when boundary functions are connected with each other by some condition. This condition is substantially used for reducing the diffraction problem on the boundary of partial areas. Some aspects of method of the over-determined boundary value problem in the theory of propagation and diffraction of electromagnetic waves were discussed earlier in [1]. Review of publications, devoted to the method of the over-determined boundary value problem in the theory of boundary value problems for partial differential equations, is given in the article [2].

As at least one of partial areas is unbounded, it is necessary to take into account the radiation condition to formulate the conjugation problem. In this work a supposition that the unknown solution is oriented is used as a radiation condition. We will say that solution of the Maxwell set of equations is positively oriented (in relation to a normal to the surface \mathcal{S}), if corresponding to it wave transfers energy or (and) attenuates slowly in this direction. Three-dimensional space is a waveguide structure, because non-trivial solutions of the homogeneous Maxwell set of equations exist in it.

In orthogonal coordinates the coordinate surface \mathcal{S} can be considered as cross-section of waveguide structure, and normal to \mathcal{S} can be considered as an axis of structure.

2. CARTESIAN COORDINATES

Let us consider the Maxwell set of equations for complex amplitudes of components of the harmonic field

$$\operatorname{rot} H = i\omega\varepsilon_0\varepsilon E, \quad \operatorname{rot} E = -i\omega\mu_0\mu H, \quad (1)$$

dependence on time is chosen in a form $\exp(i\omega t)$. In the Cartesian system of coordinates (x, y, z) we will consider the plane $z = 0$ as a cross-section of three-dimensional space.

We apply the Fourier transformation by tangent variables $x \rightarrow \xi$, $y \rightarrow \eta$ and then exclude the unknown E_z , H_z from new equations. Then we get the set of two vector equations

$$i\omega\mu_0\mu \frac{\partial}{\partial z} \begin{pmatrix} H_x \\ H_y \end{pmatrix} = P(\xi, \eta) \begin{pmatrix} E_x \\ E_y \end{pmatrix}, \quad i\omega\varepsilon_0\varepsilon \frac{\partial}{\partial z} \begin{pmatrix} E_x \\ E_y \end{pmatrix} = P(\xi, \eta) \begin{pmatrix} H_x \\ H_y \end{pmatrix},$$

$$P(\xi, \eta) = \begin{pmatrix} -\xi\eta & \xi^2 - k^2 \\ k^2 - \eta^2 & \xi\eta \end{pmatrix}.$$

The fundamental set of solutions of these equations consists of functions $\exp[-i\gamma(\xi, \eta)z]$ and $\exp[i\gamma(\xi, \eta)z]$, where $\gamma(\xi, \eta) = \sqrt{k^2 - \xi^2 - \eta^2}$. We will calculate a radical so that the first summand determine the waves of positive orientation, and second summand determine the waves of negative orientation. If

$$H_x(\xi, \eta, z) = h_x(\xi, \eta) e^{-i\gamma(\xi, \eta)z}, \quad H_y(\xi, \eta, z) = h_y(\xi, \eta) e^{-i\gamma(\xi, \eta)z},$$

$$E_x(\xi, \eta, z) = e_x(\xi, \eta) e^{-i\gamma(\xi, \eta)z}, \quad E_y(\xi, \eta, z) = e_y(\xi, \eta) e^{-i\gamma(\xi, \eta)z},$$

then by vector equations it follows that vector-functions $e = (e_x, e_y)$ $h = (h_x, h_y)$ satisfy equations

$$\omega\mu_0\mu\gamma(\xi, \eta) h(\xi, \eta) = P(\xi, \eta) e(\xi, \eta) \quad \text{or} \quad \omega\varepsilon_0\varepsilon\gamma(\xi, \eta) e(\xi, \eta) = P(\xi, \eta) h(\xi, \eta). \quad (2)$$

These equalities establish a *connection between Fourier transforms of traces of tangents components of vectors E and H of the positively oriented field on a cross-section $z = 0$* . Equalities (2) are necessary and sufficient condition of solvability of the over-determined problem.

Consequently, functions (distributions) $e_x(x, y)$, $e_y(x, y)$, $h_x(x, y)$, $h_y(x, y)$ are traces of the positively oriented solution of the Maxwell set of equations in upper half-space (or in lower half-space) if and only if, when their Fourier images satisfy the conditions (2). In case of solutions of negative orientation it is necessary to put other sign before a function $\gamma(\xi, \eta)$. In paper [3] an analogical result was obtained by the Fourier integral transformation method by all three variables.

Conditions (2) are equipotent to vector integral equalities

$$e(x, y) = -\frac{1}{\omega\varepsilon_0\varepsilon} \iint K(x_1, y_1; x, y) h(x_1, y_1) dx_1 dy_1,$$

$$h(x, y) = \frac{1}{\omega\mu_0\mu} \iint K(x_1, y_1; x, y) e(x_1, y_1) dx_1 dy_1,$$
(3)

where

$$K(x_1, y_1; x, y) = \frac{1}{4\pi^2} \iint \frac{1}{\gamma(\xi, \eta)} P(\xi, \eta) e^{i(x_1-x)\xi + i(y_1-y)\eta} d\xi d\eta.$$

Let $e^0(x, y)$ be a trace on the plane $z = 0$ of tangent of component of vector E of electromagnetic wave falling down on a screen \mathcal{M} . The diffraction problem consists of the following. It is necessary to find positively oriented solution of the set of Equation (1) in upper half-space (sign $+$) and the negatively oriented decision in lower half-space (sign $-$), satisfying the boundary conditions

$$e^\pm(x, y) + e^0(x, y) = 0, \quad (x, y) \in \mathcal{M},$$

$$e^+(x, y) - e^-(x, y) = 0, \quad h^+(x, y) - h^-(x, y) = 0, \quad (x, y) \in \mathcal{N}. \quad (4)$$

We will rewrite equalities of the form (4) for traces of tangent components of vectors E^\pm , H^\pm . It follows by boundary conditions that $e^+ = e^- = -e^0$ on \mathcal{M} and $e^+ = e^-$ on \mathcal{N} . Denote $e = e^+ = e^-$. Then $h^+ + h^- = 0$ everywhere on a plane $z = 0$, from here $h^+ = h^- = 0$ on \mathcal{N} . Therefore

$$e(x, y) = \mp \frac{1}{\omega\varepsilon_0\varepsilon} \iint_{\mathcal{M}} K(x_1, y_1; x, y) h_\pm(x_1, y_1) dx_1 dy_1.$$

It is easy to see that *diffraction problem on the plane screen \mathcal{M} is equivalent to vector integral equation*

$$\mp \frac{1}{\omega\varepsilon_0\varepsilon} \iint_M K(x_1, y_1; x, y) h^\pm(x_1, y_1) dx_1 dy_1 = -e_0(x, y), \quad (x, y) \in \mathcal{M}. \quad (5)$$

In [3] (see also [2]) it is shown, how integral equation of 1st kind (5) can be transformed into integral equation of 2nd kind by means of conditions of solvability of the over-determined problems.

3. CYLINDRICAL COORDINATES

Let us consider cylindrical surface $r = R$ as a section of waveguide structure in the cylindrical system of coordinates (r, α, z) . Every component of vectors E and H we will decompose into the Fourier series by a variable α . Then the Fourier coefficients with number n of the unknown functions should satisfy the set of equations

$$\begin{aligned} \frac{in}{r} H_{z,n} - \frac{\partial H_{\alpha,n}}{\partial z} &= i\omega\varepsilon_0\varepsilon E_{r,n}, & \frac{in}{r} E_{z,n} - \frac{\partial E_{\alpha,n}}{\partial z} &= -i\omega\mu_0\mu H_{r,n}, \\ \frac{\partial H_{r,n}}{\partial z} - \frac{\partial H_{z,n}}{\partial r} &= i\omega\varepsilon_0\varepsilon E_{\alpha,n}, & \frac{\partial E_{r,n}}{\partial z} - \frac{\partial E_{z,n}}{\partial r} &= -i\omega\mu_0\mu H_{\alpha,n}, \\ \frac{1}{r} \frac{\partial}{\partial r} (rH_{\alpha,n}) - \frac{in}{r} H_{r,n} &= i\omega\varepsilon_0\varepsilon E_{z,n}, & \frac{1}{r} \frac{\partial}{\partial r} (rE_{\alpha,n}) - \frac{in}{r} E_{r,n} &= -i\omega\mu_0\mu H_{z,n}. \end{aligned} \quad (6)$$

In particular case, when the field components do not depend on the coordinate z , set of Equation (6) splits on independent subsystems. Any solution of the Maxwell set of equations is a sum of two solutions being the waves of parallel and perpendicular polarization.

Solutions of parallel polarization have the form $E_n = (0, 0, E_{z,n})$, $H_n = (H_{r,n}, H_{\alpha,n}, 0)$. The potential function $u_n(r) = E_{z,n}z$, n should satisfy the Bessel equation for every n

$$u_n'' + \frac{1}{r} u_n' + \left(k^2 - \frac{n^2}{r^2}\right) u_n = 0.$$

Therefore

$$E_z(r, \alpha) = \sum_{n=-\infty}^{+\infty} [a_n H_n^{(1)}(kr) + b_n H_n^{(2)}(kr)] e^{in\alpha}, \quad (7)$$

a_n, b_n are arbitrary constants.

Terms with the coefficients a_n determine the waves of negative orientation relatively the external normal to the section, and terms with the coefficients b_n determine the waves of positive orientation. General representation of form (7) gives the non-oriented wave.

The Maxwell set of equations loses sense in the cylindrical system of coordinates by $r = 0$. Therefore the additional conditions should be given to compensate this loss (they determine the behavior of the unknown field by $r \rightarrow 0$). These conditions have the same nature that the radiation conditions. If there are no sources on the axis $z = 0$, then all energy leaving from any surface surrounding the axis $z = 0$ returns on this surface. Therefore it should be $a_n = b_n$ for all n in representation (7). Then the sums of Hankel functions are equal to the Bessel functions.

The field in a waveguide structure can be determined by traces of tangent components of vectors E and H on its cross-section, i.e., by conditions

$$E_z(R, \alpha, z) = e_z(\alpha, z), \quad H_\alpha(R, \alpha, z) = h_\alpha(\alpha, z), \quad (8)$$

or in particular case by conditions

$$E_z(R, \alpha) = e_z(\alpha), \quad H_\alpha(R, \alpha) = h_\alpha(\alpha). \quad (9)$$

It is easy to see that the oriented wave is determined in a unique fashion by one trace, and by two traces the non-oriented wave is determined.

For the positively oriented wave conditions (9) are reduced to equalities

$$\sum_{n=-\infty}^{+\infty} b_n H_n^{(2)}(kR) e^{in\alpha} = e_z(\alpha), \quad -i\sqrt{\frac{\varepsilon_0\varepsilon}{\mu_0\mu}} \sum_{n=-\infty}^{+\infty} b_n H_n^{(2)'}(kR) e^{in\alpha} = h_\alpha(\alpha).$$

It follows from this that *functions $e_z(\alpha)$ and $h_\alpha(\alpha)$ are traces on the cross-section $r = R$ of waveguide structure of the positively oriented wave of parallel polarization if and only if, when either (in language of Fourier coefficients)*

$$\frac{k}{i\omega\mu_0\mu} H_n^{(2)'}(kR) e_{z,n} = H_n^{(2)}(kR) h_{\alpha,n}, \quad n = 0, \pm 1, \dots \quad (10)$$

or (in language of functions)

$$e_z(\alpha) = \frac{i\omega\mu_0\mu}{k} \frac{1}{2\pi} \int_0^{2\pi} h_\alpha(\alpha_1) \sum_{m=-\infty}^{+\infty} \frac{H_n^{(2)}(kR)}{H_n^{(2)\prime}(kR)} e^{im(\alpha-\alpha_1)} d\alpha_1, \quad 0 < \alpha < 2\pi. \quad (11)$$

If the field components depend on the coordinate z , then we will suppose that this dependence has a slow growth at infinity for any fixed r and α .

Let us apply the Fourier transformation on the variable $z \rightarrow \zeta$ to Equation (6). All unknown functions depend on the variable r and on a parameter ζ .

Potential functions $E_z = u$ and $H_z = v$ should satisfy equations of the form

$$\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \left[k^2 - \zeta^2 - \frac{n^2}{r^2} \right] u = 0.$$

Particular solutions of this equation are cylindrical functions $Z_n(\sqrt{k^2 - \zeta^2} r)$. Here it is also possible to use the Hankel functions. But if a multiplier at r is imaginary, then it is expedient to pass to the MacDonald functions. In general case solution contains two terms of such form, and if solution is oriented then it contains only one element.

Thus, for the oriented decisions we have

$$E_{z,n}(r, z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} c_n(\zeta) Z_n(\sqrt{k^2 - \zeta^2} r) e^{-i\zeta z} d\zeta,$$

where $c_n(\zeta)$ are some distributions.

As well as in particular case, the non-oriented solution of the Maxwell set of equations is determined in a unique fashion by conditions (8) in space without the line $r = 0$. For the oriented equation it is sufficiently to set only one trace of the field on a cylindrical surface, for example, $e_z(\alpha, z)$.

If both traces of the field are given on a cylindrical surface, then oriented solution exists if and only if, when distributions $e_z(\alpha, z)$ and $h_\alpha(\alpha, z)$ satisfy the solvability conditions of the over-determined problem.

It is easy to obtain integral equations to which the electromagnetic wave diffraction problem on the screen located on a cylindrical surface $r = R$ can be reduced. In particular case, when the field does not depend on the coordinate z , integral equation of the diffraction problem can be written down as a summatorial equation for Fourier coefficients b_n . In general case two-dimensional integral equation can be transformed into summatorial-integral equation for functions $c_n(\zeta)$.

4. SPHERICAL COORDINATES

It is natural to consider the sphere $r = R$ as a cross-section in the spherical coordinates system (r, α, θ) . Maxwell equations set in the spherical coordinates is a set of six partial differential equations with respect to six functions $E_r, E_\theta, E_\alpha, H_r, H_\theta, H_\alpha$ by three variables. Each of the unknown functions can be expanded into Fourier series by variable α . The Fourier coefficients with number m (this number is not shown) satisfy the set of equations

$$\begin{aligned} \frac{\partial(\sin \theta H_\alpha)}{\partial \theta} - im H_\theta &= i\omega\varepsilon_0\varepsilon r \sin \theta \cdot E_r, & \frac{\partial(\sin \theta E_\alpha)}{\partial \theta} - im E_\theta &= -i\omega\mu_0\mu r \sin \theta \cdot H_r, \\ im H_r - \sin \theta \frac{\partial(r H_\alpha)}{\partial r} &= i\omega\varepsilon_0\varepsilon r \sin \theta \cdot E_\theta, & im E_r - \sin \theta \frac{\partial(r E_\alpha)}{\partial r} &= -i\omega\mu_0\mu r \sin \theta \cdot H_\theta, \\ \frac{\partial(r H_\theta)}{\partial r} - \frac{\partial H_r}{\partial \theta} &= i\omega\varepsilon_0\varepsilon r E_\alpha, & \frac{\partial(r E_\theta)}{\partial r} - \frac{\partial E_r}{\partial \theta} &= -i\omega\mu_0\mu r H_\alpha. \end{aligned}$$

Any solution of this set of equations can be represented as a sum of two particular solutions: of magnetic type for $E_r = 0$ and of electric type for $H_r = 0$.

In the axis-symmetrical case ($m = 0$) the potential function u of the electric field and the potential function v of the magnetic field should satisfy equation

$$r \frac{\partial^2(ru)}{\partial r^2} + \frac{\partial}{\partial \theta} \left(\frac{1}{\sin \theta} \frac{\partial(\sin \theta \cdot u)}{\partial \theta} \right) + k^2 r^2 u = 0. \quad (12)$$

It is known that the particular solutions of this equation have the form

$$u_n(r, \theta) = \frac{1}{\sqrt{kr}} Z_{n+1/2}(kr) \Theta_n(\theta), \quad \Theta_n(\theta) = \text{const } P_n^{(1)}(\cos \theta), \quad n = 1, 2, \dots$$

here $Z_{n+1/2}(\cdot)$ are cylindrical functions and $P_n^{(1)}(\cdot)$ are joint Legendre functions. The functions $\Theta_n(\theta)$ are orthogonal-normalized with weight $\sin \theta$ on segment $[0, \pi]$.

Non-oriented solution of the set of Equation (1) in spherical coordinates can be written down by the potential functions in the following way

$$u(r, \theta) = \sum_{n=1}^{+\infty} \left[\vec{u}_n \frac{1}{\sqrt{kr}} H_{n+1/2}^{(2)}(kr) + \overleftarrow{u}_n \frac{1}{\sqrt{kr}} H_{n+1/2}^{(1)}(kr) \right] \Theta_n(\theta). \quad (13)$$

First summands in the sums determine the positively oriented waves and the second summand determine the negatively oriented ones.

It is easy to obtain the solvability conditions of the over-determined problems in the case of spherical coordinates system. It is necessary to decompose by the functions $\Theta_n(\theta)$ the traces on the sphere $r = R$ of the components E_θ , E_α , H_θ , H_α of the oriented solution. The constants u_n and v_n are determined in two ways by these coefficients of the decomposition. Compare the corresponding expressions and obtain the dependence between the traces of tangential components of the vectors E and H on the sphere in the terms of their Fourier coefficients.

The dual summatorial equation of the electromagnetic wave diffraction on the spherical screen was constructed in the paper [4].

In the case when $m \neq 0$ the electric and magnetic potential functions are represented by the cylindrical functions and joint Legendre functions also. The unknown coefficients of the field decomposition have two indexes: m and n .

ACKNOWLEDGMENT

Supported by RFBR 12-01-97012-r_povolzh'e.a.

REFERENCES

1. Pleshchinskii, N. B., I. E. Pleshchinskaya, and E. M. Karchevskiy, "The over-determined boundary value problem method in the electromagnetic waves propagation and diffraction theory," *PIERS Online*, Vol. 5, No. 5, 441–445, 2009.
2. Pleshchinskaya, I. E. and N. B. Pleshchinskii, "Over-determined boundary value problems for elliptic partial differential equations and their application to waves diffraction theory," *Uchenye zapiski Kazanskogo gosudarstvennogo universiteta*, Vol. 147, No. 3, 41–32, 2005 (in Russian).
3. Pleshchinskaya, I. E. and N. B. Pleshchinskii, "The over-determined Cauchy problems for the Maxwell equations set and electromagnetic waves diffraction on the metallic screens," *Conf. Proc. 11th Int. Conf. Mathematical Methods in Electromagnetic Theory, MMET' 06*, 255–257, Kharkov, Ukraine, June 2006.
4. Karchevskiy, E. M. and N. B. Pleshchinskii, "Parallel algorithm of solving the electromagnetic wave diffraction problem on the spherical screen," *PIERS Proceedings*, 131–134, Moscow, Russia, August 19–23, 2012.

Plasma Treatment as a Versatile Technique for Preparation of Plasmonic Nanoantenna Arrays

L. Štolcová¹, M. Domonkos¹, T. Ižák², J. Proška¹,
M. Procházka³, and A. Kromka²

¹Department of Physical Electronics, Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University in Prague, Břehová 7, 115 19 Prague 1, Czech Republic

²Institute of Physics, Academy of Sciences of the Czech Republic
Cukrovarnická 10, 162 00 Prague 6, Czech Republic

³Faculty of Mathematics and Physics, Institute of Physics
Charles University, Ke Karlovu 5, 121 16 Prague 2, Czech Republic

Abstract— Metal nanostructures supporting localized surface-plasmon resonance (LSPR) have attracted growing research interest in the last decades due to their ability to confine light and produce high enhancement of the local electromagnetic field in their vicinity. These effects gave rise to novel scientific methods, such as surface-enhanced spectroscopies.

In this contribution, we report on a simple, relatively low-cost, and high-throughput technique to produce meso-scale structures with certain features as small as a few nanometres. This property enables to control comfortably the optical response of the metal nanostructures, and create large periodic arrays of rationally designed plasmonic substrates. By adjusting the plasma etching conditions and character (i.e., reactive ion etching or pulsed linear-antenna microwave plasma), self-assembled monodisperse beads either decrease their size to form non-close packed spheres, create long necks toward adjacent spheres, or transform to corrugated anisotropic structures known as nanocorals.

Selected types of nanoantenna arrays were successfully tested as substrates for surface-enhanced Raman scattering (SERS) by measuring SERS spectra of ρ -aminothiophenol, a prototypical SERS probe, at submicromolar concentrations.

1. INTRODUCTION

Plasmonic nanoantennas can be defined as nanostructures converting free-propagating optical radiation to localized energy, and vice versa, via the localized surface-plasmon resonance (LSPR) [1, 2]. This phenomenon occurs when the incident light excites resonant oscillations of electron gas contained in noble metal nanostructures and causes deep-subwavelength confinement and thus also significant local enhancement of the electromagnetic field in the particle vicinity. These properties can be utilized in a broad field of applications, such as surface-enhanced spectroscopies, sensing nanoscale imaging, or photovoltaics.

The spectroscopy based on surface-enhanced Raman scattering (SERS) particularly benefits from the field enhancement, since the inherently weak signal from molecules is the major disadvantage of Raman spectroscopy. The detection limit of SERS is several orders of magnitude lower, and in specific cases, single-molecule SERS spectra can be collected [3]. However, some fundamental aspects of SERS remain to be understood, e.g., the role of the so-called chemical enhancement. Moreover, there are technical challenges regarding SERS substrates to be overcome, namely SERS signal reproducibility and homogeneity over larger areas which only rarely coincide with prominent field enhancements.

The optical response of a metal nanostructure, and hence the LSPR frequency, is determined by the material, shape and size of the metal nanostructure, and also by the coupling with adjacent plasmonic objects. Theoretical calculations show that a small change in the geometry can lead to dramatic LSPR shifts. Therefore, the rational design [4] of metallic nanostructures requires the control of geometrical parameters on the nanoscale. Top-down lithographic techniques (EBL and FIB) are used to obtain nanoantennas of arbitrary shape and size down to the order of 10 nm, but their low throughput and high cost present serious drawbacks. Alternatively, self-assembly of metallic particles or utilizing dielectric colloidal particles as templates and/or masks for metal deposition can be employed to produce large areas of both 2D and 3D periodic nanoantenna arrays. Although the spectrum of structure shapes obtained by this technique is limited, it can be significantly broadened by using template-assisted self-assembly, angled deposition, or etching [5]. Dry etching is a particularly popular technique because it does not contaminate the etched samples.

In this paper, we present three diverse periodic nanoantenna array types produced by combining self-assembly of polystyrene (PS) microspheres, gold sputter coating, and plasma treatment. The prepared structures illustrate the versatility of the presented technique and its capability to control the nanoantenna dimensions with the precision comparable to (or even better than) the costly top-down approaches.

2. EXPERIMENTAL

The preparation procedures and their final products are schematically depicted in Figure 1.

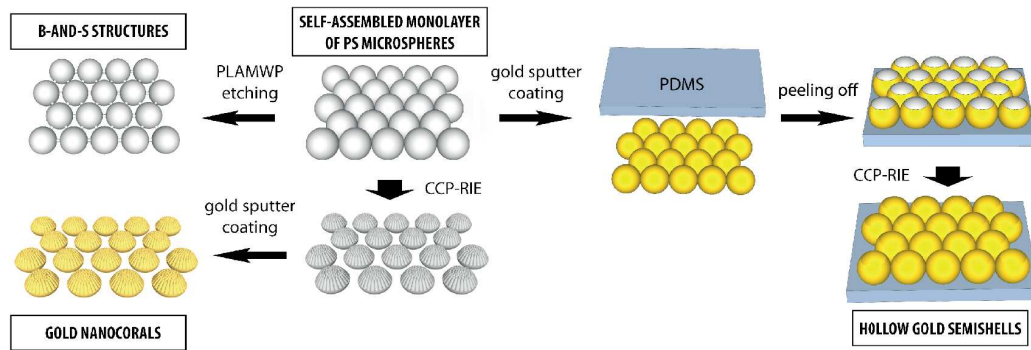


Figure 1: Schematics of the nanoantenna arrays preparation.

PS microspheres of diameters (253 ± 8) nm and (471 ± 13) nm in aqueous dispersion (microParticles GmbH Germany) were diluted with ethanol (1:1 v/v) and deposited onto water surface using a glass pipette as described in [6]. There they formed monolayers with hexagonal close-packed ordering, which were then transferred onto cleaned silicon or glass substrates (cca 1×1 cm²) and left to dry.

Corrugated non-close-packed anisotropic structures called nanocorals [7] were produced by etching the monolayers of PS spheres 471 nm and 253 nm in diameter in a capacitively coupled reactive ion etching (CCP-RIE) plasma system (Phantom III/Phantom LT RIE System, Trion Technology) with total gas pressure 90 m Torr, RF power of 100 W, O₂ flow rate of 50 sccm, for 30 s and 20 s, respectively. To improve the adhesion to the substrate, the arrays of bigger (originally 471 nm) nanocorals were thermally treated at 150°C for 30 s.

'Ball and stick' (b-and-s) structures were formed by etching monolayers of PS spheres in pulsed linear-antenna microwave plasma (PLAMWP) [8] system (AK 400, Roth and Rau, AG), in which two linear MW antennas (two quartz tubes) parallel to one another were arranged above the substrate holder. The process was performed with a MW and RF glow discharges simultaneously, total gas pressure 1 mbar, RF power of 600 W, pulsed microwave power of 2×1700 W, O₂ flow rate of 198 sccm, and process duration 1 min.

20 nm of gold was deposited onto both close- and non-close-packed PS templates by magnetron sputtering in a high-resolution sputter coater (Cressington 208HR).

Gold semishells on a PDMS substrate were prepared as follows: PDMS (Sylgard Elastomer Kit 184 Dow Corning) was mixed with curing agent at a ratio 10:1, stirred and degassed by centrifugation, and cured in an oven at 75°C for 24 h. The gold-coated close-packed PS microspheres were attached to PDMS and peeled off. PS was removed from the gold semishells by CCP-RIE at gas pressure 90 m Torr, RF power of 100 W, O₂ flow rate of 50 sccm, and process duration 150 s.

All samples were characterized using field emission scanning electron microscope (FE-SEM) JEOL JSM-7500F.

Gold nanocorals were immersed in a water solution of p-ATP (Sigma Aldrich) of concentration 10^{-7} M for 11 hours. Then they were thoroughly rinsed in cold water and dried in nitrogen. SERS spectra were taken by the confocal Raman microspectrometer LabRam HR800 (Horiba Jobin-Yvon) with a nitrogen cooled CCD detector, using a 785 nm excitation line, objective 100 \times , accumulation time 30×1 s laser power 40 μ W, and approx 2 μ m spot size of the laser beam.

3. RESULTS AND DISCUSSION

High-quality monolayers of PS microspheres on Si and glass substrates were obtained. The defect-free close-packed arrays of the spheres 471 nm and 253 nm in diameter were almost 50×50 μ m²

and $10 \times 10 \mu\text{m}^2$, respectively. Subjecting them to plasma treatment affected only their size and shape, and did not disrupt their hexagonal ordering. Moreover, the etching process was performed homogeneously over the whole sample area.

During the PLAMWP etching of microspheres, thin necks appeared between them, creating a b-and-s geometry (Figure 2). Similar structures have been achieved by different techniques, such as hyperthermal neutral beam etching [9] or RIE of PS spheres [10, 11], and the neck formation has been ascribed to PS melting prior to or during the etching process. The b-and-s structure has been used as a mask to obtain hexagonal array of gold bow-tie nanoantennas with controllable gap sizes [11], but to our knowledge, never in this size range. The width of the formed necks between the 471 nm and 253 nm spheres was estimated from SEM micrographs as approx. 60 nm and 20 nm, respectively, while their length as 30 nm for both of the sphere sizes. Using b-and-s with very long necks (Figure 2(b)) as a mask results in planar (2D) truncated triangle nanoantenna arrays. It should be noted that the truncation size can be finely tuned by the neck length. More experiments to tailor metallic nanoantenna shapes using b-and-s structures as masks are currently being developed.

The RIE process resulted in anisotropic shrinking of PS microspheres and formation of prominent corrugations at their upper part. Figure 3(a) shows the side view of the ellipsoidal corrugated structure obtained from a PS sphere 471 nm in diameter. Its height and width was estimated from as 270 nm and 320 nm, respectively. The prepared gold-nanocorals are shown in Figures 3(b), (c). SERS spectra of p-ATP measured on this substrate displayed a high signal-to noise ratio and an excellent homogeneity over the whole well-ordered domains. Figure 4 shows raw SERS spectra of p-ATP collected from 22×22 points separated by $2 \mu\text{m}$ (a square grid) at the nanocoral array. The signal enhancement displayed by gold nanocorals can be attributed to plasmon resonances at the corrugations and their coupling, to the effect of corrugation-contributed roughness and coupling between neighbouring particles [12].

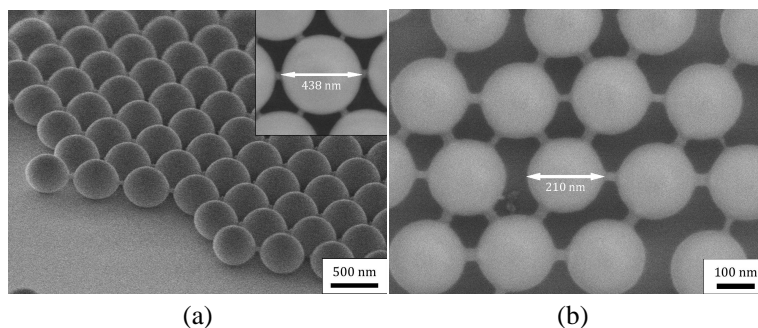


Figure 2: SEM micrographs of b-and-s structures produced by PLAMWP etching of PS spheres (a) 471 nm and (b) 253 nm in diameter.

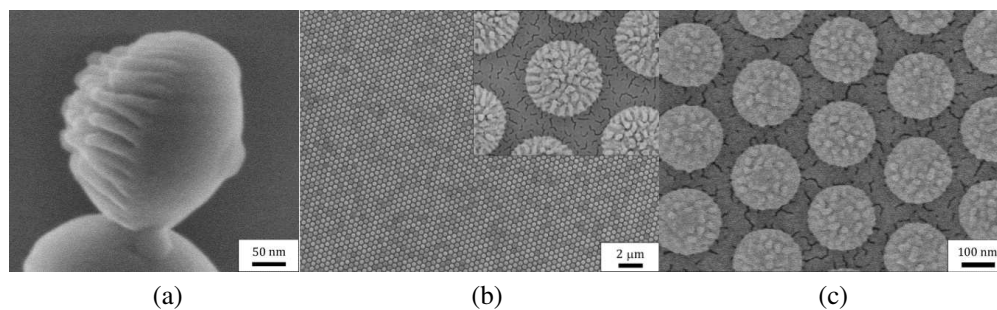


Figure 3: SEM micrographs: (a) side view of a PS nanocoral, gold nanocorals produced by RIE of PS spheres (b) 471 nm and (c) 253 nm in diameter.

Next, we optimized the process of removing PS templates from 20 nm thick gold semishells. Metallic semishells are well-known for the high degree of LSPR tunability (via adjusting their size, thickness, and fractional height) and significant field enhancement around the sharp rims of their apertures [13]. Hollow gold semishell array, prepared by selectively etching out the PS in oxygen

plasma, is shown in Figure 5(a). The gold grains visible at the bottom of the individual semishell (see Figure 4(b)) imply that plasma etching is able to remove the PS almost perfectly. This is a great advantage over wet chemical methods (dissolution), which leave a thin layer of PS at the inner surface of the semishell [14].

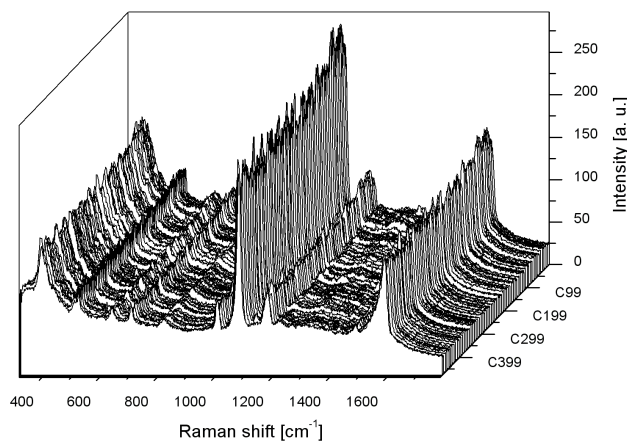


Figure 4: Raw SERS spectra of p-ATP measured at 22×22 points separated by $2 \mu\text{m}$ at the nanocoral array (C1–C484 denote 22×22 spots where the individual spectra were taken).

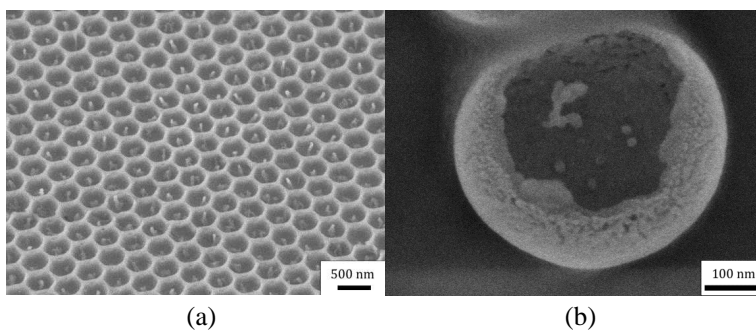


Figure 5: SEM micrographs of (a) a hollow gold semishell array and (b) an individual semishell.

4. CONCLUSION

In this contribution, it was demonstrated that colloidal self-assembly combined with plasma treatment and metal sputter coating is a versatile technique for the preparation of both 2D and 3D plasmonic nanoantenna arrays. Plasma treatment did not disturb the hexagonal ordering of the primary template and enabled a very fine control over the size and shape of the etched microspheres. Using PLAMWP etching, b-and-s structures with 30 nm wide and 20 nm long necks were prepared. These structures can be utilized as masks for the preparation of metallic truncated triangle nanoantennas with tunable truncation sizes. Using RIE, distinct surface corrugations occurred on the microspheres, which contribute to the enhancement displayed by the ‘nanocorals’. SERS spectra of p-ATP measured on nanocoral arrays at submicromolar concentrations revealed not only a high signal-to-noise ratio, but also an excellent homogeneity of the enhancement over the whole ordered domains. Plasma etching of PS spheres from thin gold semishells resulted in an almost perfect removal of the PS. This presents a great advantage over dissolution, particularly for the possible application in SERS.

ACKNOWLEDGMENT

This work was supported by the Czech Science Foundation grants No. P205/13/20110S (JP, LS) and P108/12/G108 (AK, TI).

REFERENCES

1. Giannini, V., A. I. Fernandez-Dominguez, S. C. Heck, and S. A. Maier, “Plasmonic nanoantennas: Fundamentals and their use in controlling the radiative properties of nanoemitters,” *Chem. Rev.*, Vol. 111, No. 6, 3888–3912, 2011.
2. Novotny, L. and N. van Hulst, “Antennas for light,” *Nature Photon.*, Vol. 5, No. 2, 83–90, 2011.
3. Le Ru, E. C. and P. G. Etchegoin, “Single-molecule surface-enhanced Raman spectroscopy,” *Annu. Rev. Phys. Chem.*, Vol. 63, 65–87, 2012.
4. Banholzer, M. J., J. E. Millstone, L. Qina, and C. A. Mirkin, *Chem. Soc. Rev.*, Vol. 37, No. 5, 885–897, 2008.
5. Zhang, J., Y. Li, X. Zhang, and B. Yang, “Colloidal self-assembly meets nanofabrication: From two-dimensional colloidal crystals to nanostructure arrays,” *Adv. Mater.*, Vol. 22, No. 38, 4249–4269, 2010.
6. Kosiorok, A., W. Kandulski, P. Chudzinski, K. Kempa, and M. Giersig, “Shadow nanosphere lithography: Simulation and experiment,” *Nano Lett.*, Vol. 4, No. 7, 1359–1363, 2004.
7. Wu, L. Y., B. M. Ross, S. Hong, and L. P. Lee, “Bioinspired nanocorals with decoupled cellular targeting and sensing functionality,” *Small*, Vol. 6, No. 4, 503–507, 2010.
8. Kromka, A., O. Babchenko, T. Izak, S. Potocky, M. Davydova, N. Neykova, H. Kozak, Z. Remes, K. Hruska, and B. Rezek, “Pulsed linear antenna microwave plasma — A step ahead in large material depositions and surface functionalization,” *Proceedings of Nanocon 2011*, 271–279, Brno, Czech Republic, Sep. 2011.
9. Cho, Y. S., G. R. Yi, J. H. Moon, D. C. Kim, B. J. Lee, and S. M. Yang, “Connected open structures from close-packed colloidal crystals by hyperthermal neutral beam etching,” *Langmuir*, Vol. 21, No. 23, 10770–10775, 2005.
10. Cong, C. X., W. C. Junus, Z. Shen, and T. Yu, “New colloidal lithographic nanopatterns fabricated by combining pre-heating and reactive ion etching,” *Nanoscale Res. Lett.*, Vol. 4, No. 11, 1324–1328, 2009.
11. Lohmuller, T., L. Iversen, M. Schmidt, C. Rhodes, H. L. Tu, W. C. Lin, and J. T. Groves, “Single molecule tracking on supported membranes with arrays of optical nanoantennas,” *Nano Lett.*, Vol. 12, No. 3, 1717–1721, 2012.
12. Hsieh, H. Y., J. L. Xiao, C. H. Lee, T. W. Huang, C. S. Yang, P. C. Wang, and F. G. Tseng, “Au-coated polystyrene nanoparticles with high-aspect-ratio nanocorrugations via surface-carboxylation-shielded anisotropic etching for significant SERS signal enhancement,” *J. Phys. Chem. C*, Vol. 115, No. 33, 16258–16267, 2011.
13. Van Dorpe, P. and J. Ye, “Semishells: Versatile plasmonic nanoparticles,” *ACS Nano*, Vol. 5, No. 9, 6774–6778, 2011.
14. Stolcova, L., J. Proska, and M. Prochazka, “Hexagonally ordered gold semishells as tunable SERS substrates,” *Proceedings of Nanocon 2012*, 225–229, Brno, Czech Republic, Oct. 2012.

The Dispersive Properties of the Three-dimensional Photonic Crystals with Diamond Lattices Containing the Epsilon-negative Materials

Hai-Feng Zhang^{1,2}, Wen-Ping He², Shao-Bin Liu¹, and Yu-Qing Chen²

¹College of Electronic and Information Engineering
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

²Nanjing Artillery Academy, Nanjing 211132, China

Abstract— In this paper, dispersive properties of three-dimensional (3D) photonic crystals (PCs) with diamond lattices composed of the isotropic positive-index materials and epsilon-negative materials are theoretically investigated based on a modified plane wave expansion (PWE) method. The eigenvalue equations of such structure (spheres with epsilon-negative materials inserted in the dielectric background) are deduced. It can be obviously seen that a photonic band gap (PBG), a flatbands region, and two stop band gaps (SBGs) in the Γ - X and Γ - L directions appear, respectively. The results show that the upper edges of flatbands region can not be tuned by any parameters except for the electronic plasma frequency. The first PBG and the first SBGs above the flatbands region in the Γ - X and Γ - L directions for the 3D PCs can be modulated by the filling factor, relative dielectric constant and electronic plasma frequency, respectively. However, the damping factor has no effect on the locations of the first PBG and the first SBGs above the flatbands region in the Γ - X and Γ - L directions.

1. INTRODUCTION

Since the initial works of Yablonovitch [1] and John [2], the photonic crystals (PCs) have been attracted increasing attentions of researches for their rich physics and potential applications. In recent years, the metamaterials or left-handed materials have been introduced to PCs to form the tunable PBGs since firstly proposed by Veselago in 1967 [3], so the PCs containing the metamaterials become a new active research area. Similar to the conventional PCs, the PCs containing the metamaterials display strong spatial dispersion, resulting in the appearance of electromagnetic band gap structure [4]. The metamaterials always can be divided into two categories [5]. One configuration, in which the permeability is negative but the permittivity is positive, gives rise to so-called epsilon-negative (ENG) materials. The other is that the permeability is positive but the permittivity is negative. In this case, the mu-negative (MNG) materials can be obtained. Moreover, the double-negative metamaterials hardly can be found in nature materials but the ENG materials always can be obtained, such as metal [6], plasma [7] and superconductor [8], etc.. Due to the unusual properties of the ENG materials, the electromagnetic waves (EM waves) propagating in the PCs composed of the ENG materials and right-hand materials have different dispersive properties. Up to now, the PCs composed of the ENG materials have been studied in detail [9–11]. To our knowledge, there are few reports about the dispersive properties of 3D PCs containing the isotropic dielectric and ENG material which can be regarded as a frequency-dependence medium. Thus, such kind of 3D PCs becomes a new research focus. Diamond lattices can be found in the most nature real-space structure for the optical medium. Sometimes, the researchers are more interesting in the SBG in some special applications, which is the gap in the spectrum for a fixed direction of the incident light [12]. In this paper, we intend to study the dispersive properties of 3D PCs containing the isotropic dielectric and ENG materials (spheres with the ENG materials are arranged in the dielectric background) with diamond lattices by PWE method. We use a Drude-like model to describe the effective dielectric function of the ENG material, and the damping factor also is considered. The influences of relative dielectric constant, electronic plasma frequency, damping factor and filling factor on the flatbands, first (1st) PBG and 1st SBGs above the flatbands region in the Γ - X and Γ - L directions are discussed, respectively.

2. THEORETICAL MODEL AND NUMERICAL METHOD

We assume the relative dielectric function for the dielectric and ENG materials are ε_a and ε_p , respectively. The radius of the spheres and lattice constant are R and a , respectively. The ε_p is

given by [13]

$$\varepsilon_p(\omega) = \varepsilon_b - \frac{\omega_p^2}{\omega(\omega + j\gamma)} \quad (1)$$

where ε_b , ω_p and γ are the dielectric constant of ENG material, the electronic plasma frequency and the damping factor that contribute to the absorption and losses, respectively. The PWE method [14] is used to computing the band structures of such 3D PCs.

3. NUMERICAL RESULTS AND DISCUSSION

As we know, a symmetric set of primitive vectors for the diamond lattice is $\mathbf{a}_1 = (0.5a, 0.5a, 0)$, $\mathbf{a}_2 = (0, 0.5a, 0.5a)$, $\mathbf{a}_3 = (0.5a, 0, 0.5a)$. The reciprocal lattice vector basis can be defined as $\mathbf{b}_1 = (2\pi/a, 2\pi/a, -2\pi/a)$, $\mathbf{b}_2 = (-2\pi/a, 2\pi/a, 2\pi/a)$, $\mathbf{b}_3 = (2\pi/a, -2\pi/a, 2\pi/a)$. The high symmetry points have the coordinate as $\Gamma = (0, 0, 0)$, $X = (2\pi/a, 0, 0)$, $W = (2\pi/a, \pi/a, 0)$, $K = (1.5\pi/a, 1.5\pi/a, 0)$, $L = (\pi/a, \pi/a, \pi/a)$, and $U = (2\pi/a, 0.5\pi/a, 0.5\pi/a)$. The convergence accuracy is better than 1% for the lower energy bands, when a total number of 729 plane waves can be used [14]. Without loss of generality, we plot $\omega a/2\pi c$ with the normalization convention $\omega_{p0}a/2\pi c = 1$. Thus, we can define the electronic plasma frequency as $\omega_p = 0.3\omega_{p0}$ to make the problem scale-invariant, and we also choose the damping factor as $\gamma = \gamma_o = 0.02\omega_p$, $\varepsilon_a = 13.9$, $\mu_a = 1$, $\varepsilon_b = 1$ and $\mu_p = 1$, respectively. Here, we only focus on the 1st SBGs above the flatbands region in the Γ - X and Γ - L directions and 1st PBG for such 3D PCs.

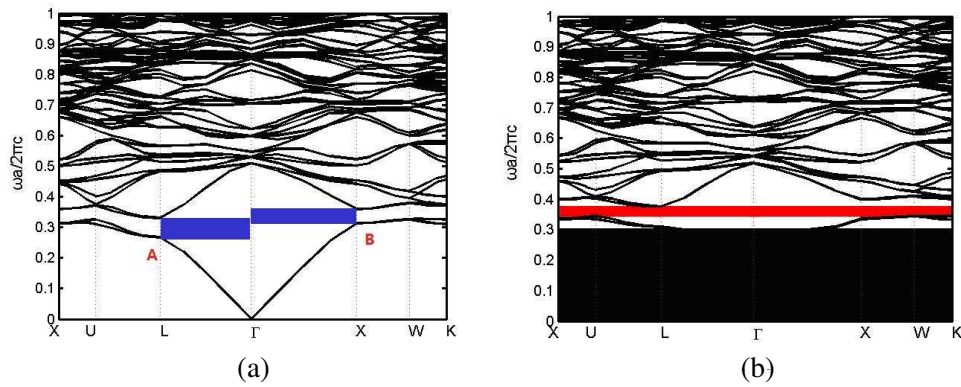


Figure 1: The band structure with $\varepsilon_a = 13.9$ and $f = 0.35$ but with different ω_p and γ . (a) $\omega_p = 0$, $\gamma = 0$, and (b) $\omega_p = 0.3\omega_{p0}$, $\gamma = 0.02\omega_p$. The red shaded region indicates the PBG, and the blue shaded regions indicate the SBGs.

The dispersion curves with ω_p , γ , $\varepsilon_a = 13.9$ and $f = 0.35$, respectively, are plotted in Fig. 1. The red and blue shaded regions indicate the PBG and SBGs, respectively. As shown in Fig. 1(a), there are two SBGs in the Γ - X and Γ - L directions can be observed as ω_p and γ are null, which are indicated by A and B. They present themselves at 0.2669–0.3313 ($2\pi c/a$), and 0.3135–0.3591 ($2\pi c/a$), respectively. The 1st PBG also can be obtained, which covers 0.3276–0.3305 ($2\pi c/a$). As the ENG material is introduced in the PCs, the edges of such two SBGs and PBG shift to higher frequencies, and a new flatbands region appears. The region of flatbands runs 0–0.3 ($2\pi c/a$). There exist the flatbands because of the existence of surface plasmon modes [14]. In the flatbands, the group velocity is slow. The positions of two SBGs are located at 0.3092–0.3761 ($2\pi c/a$) and 0.3371–0.398 ($2\pi c/a$), respectively. The 1st PBG runs from 0.3450 to 0.3759 ($2\pi c/a$). Obviously, the bandwidths of two SBGs and PBG can be enlarged by inclusion of the ENG materials in the 3D PCs.

In Fig. 2, we plot the effects of different parameters on 1st PBG and SBGs for such 3D PCs. The shaded regions indicate the PBGs. As shown in Figs. 2(a) and (d), the edges of 1st PBG and SBGs are downward to lower frequencies, and the bandwidths increase first and then decrease with increasing ε_a . The 1st PBG and SBGs will be closed at $\varepsilon_a = 40$. The maximum relative bandwidths of 1st PBG is 0.099, which can be found at $\varepsilon_a = 18$. However, The maximum relative bandwidth of 1st SBGs in the Γ - X and Γ - L directions for such 3D PCs are 0.168 and 0.212, which can be found at $\varepsilon_a = 15$ and 10, respectively. Figs. 2(b) and (e) reveal that the edges of 1st PBG and SBGs are upward to higher frequencies, and the bandwidths increase first and then decrease with increasing

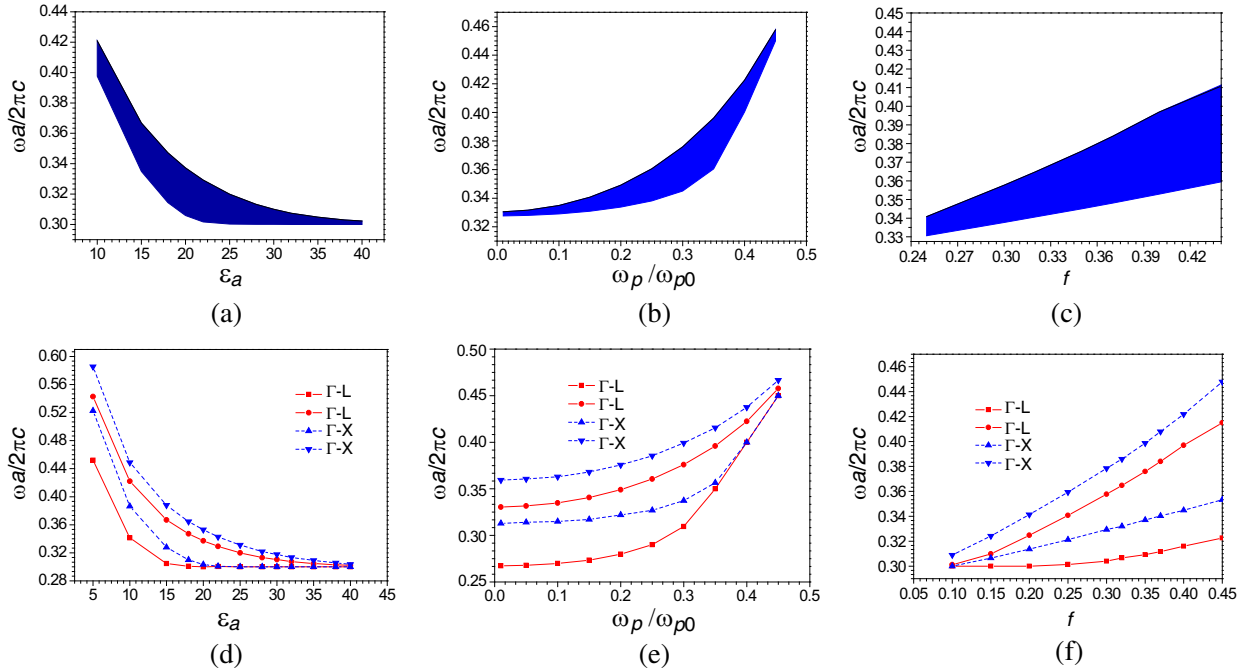


Figure 2: The effects of different parameters on the PBGs and SBGs. (a) ε_a on PBG, (b) ω_p on PBG, (c) f on PBG, (d) ε_a on SBGs, (e) ω_p on SBGs, and (f) f on SBGs, respectively. The red shaded regions indicate PBGs.

ω_p/ω_{p0} . If $\omega_p/\omega_{p0} > 0.45$, 1st PBG and SBGs will disappear. The maximum relative bandwidth is 0.094, which can be found at $\omega_p/\omega_{p0} = 0.35$. Similarly, the maximum relative bandwidths of 1st SBGs in the Γ -L and Γ -X directions for such PCs are 0.220 and 0.168, which can be found at $\omega_p/\omega_{p0} = 0.2$ and 0.3, respectively. It can be seen from Figs. 2(d) and (f) that the edges of 1st PBG and SBGs are upward to higher frequencies, and the general trends for 1st PBG and SBGs are the bandwidths almost linearly increase with increasing the filling factor of the ENG materials. If the f is larger than 0.25, the PBG appears obviously. The 1st SBGs will never appear until f is larger than 0.1. The maximum relative bandwidths of 1st PBG and SBGs in the Γ -X and Γ -L directions are 0.138, 0.237 and 0.229, which can be found at the cases of $f = 0.45$, respectively. As mentioned above, the 1st PBG and SBGs can be tuned by the ε_a , ω_p and f , respectively. This can be explained in physics that changing ε_a , ω_p and f mean that the refractive contrast and refractive contrast for such 3D PCs also are changed. It is notice that the 1st SBG in Γ -X direction has the largest relative bandwidth compared to those for 1st SBG in the Γ -L direction and PBG. It also notice that if filling factor is small enough and close to null, the 3D PCs can be looked as a dielectric block. The flatbands region will disappear. It is worth mentioning that the damping factor has no effects on the dispersive properties of such 3D PCs since the damping factor only contribute to the absorption and losses [15].

Finally, we investigate the influences of the parameters for such 3D PCs on the upper edge of flatbands region. The upper edge of flatbands region is around the ω_p because of the existence of surface plasmon region. Obviously, changing the ε_a , f and γ have no effects on the cutoff frequencies of the ENG materials, respectively. Thus, the cutoff frequencies of the ENG materials can not be changed, and the coupling between the ENG material spheres are also unchanged. The upper edge of flatbands region will linearly increase with increasing ω_p . This is because the higher surface plasmon modes can be widely extended through the interface between the ENG material spheres with increasing ω_p .

4. CONCLUSIONS

In summary, the dispersive properties of 3D PCs with diamond lattices containing the ENG materials for EM waves have been theoretically studied by the PWE method. The results show that the edge frequencies of 1st PBG and SBGs above the flatbands region in the Γ -L and Γ -X directions for such PCs shift downward to lower frequencies with increasing the relative dielectric constant, and the bandwidths are enlarged first then narrowed. The maximum relative bandwidths for the

1st PBG and SBGs always appear at low- ε_a regions. It is also can seen that the edges of such PBG and SBGs shift upward to higher frequencies with increasing the electronic plasma frequency, and bandwidths are enlarged first then narrowed. The larger filling factor means the larger bandwidths of 1st PBG and SBGs in the Γ - L and Γ - X directions, and their relative bandwidths almost linearly increase with increasing the filling factor. However, changing damping factor has no effects on the dispersive properties of proposed 3D PCs. The upper edge of flatbands region can not be tuned except for the electronic plasma frequency, and will increase linearly with increasing the electronic plasma frequency.

ACKNOWLEDGMENT

This work was supported by the Fundamental Research Funds for the Central Universities and the Funding of Jiangsu Innovation Program for Graduate Education (Grant No. CXZZ11.0211).

REFERENCES

1. Yablonovitch, E., "Inhibited spontaneous emission of photons in solid state physics and electronics," *Phys. Rev. Lett.*, Vol. 58, 059–2061, 1987.
2. John, S., "Strong localization of photons in certain disorder ed dielectric superlattices," *Phys. Rev. Lett.*, Vol. 58, 2486–2489, 1987.
3. Veselago, V. G., "The electrodynamics substances with simultaneously negative value of ε and μ ," *Sov. Phys. Uspekhi*, Vol. 10, No. 4, 509–514, 1968.
4. Zhang, H. F., S. B. Liu, and X. K. Kong, "Photonic band gaps in one-dimensional magnetized plasma photonic crystals with arbitrary magnetic declination," *Phys. Plasma*, Vol. 19, No. 12, 122103, 2012.
5. Chen, Y., "Broadband wave plate: Approach from one-dimensional photonic crystals containing metamaterials," *Phys. Lett. A*, Vol. 375, No. 7, 1156–1159, 2011.
6. Sievenpiper, D. F. and E. Yablonovitch, "3D metallo-dielectric photonic crystals with strong capacitive coupling between metallic islands," *Phys. Rev. Lett.*, Vol. 80, 2829–2832, 1998.
7. Zhang, H. F., S. B. Liu, and X. K. Kong, "Dispersion properties of three-dimensional plasma photonic crystals in Diamond lattice arrangement," *J. Lightwave Technol.*, Vol. 17, No. 11, 1694–1702, 2013.
8. Zhang, H. F., S. B. Liu, X. K. Kong, B. R. Bian, and Y. Dai, "Omnidirectional photonic band gaps enlarged by Fibonacci quasi-periodic one-dimensional ternary superconductor photonic crystals," *Solid State Commun.*, Vol. 152, No. 7, 2113–2119, 2012.
9. Li, P. and Y. Liu, "Mutichannel filtering properties of photonic crystals consisting of single-negative material," *Phys. Lett. A*, Vol. 373, No. 21, 1870–1873, 2009.
10. Feng, S., Y. Li, B. Lei, J. Hu, Y. Wang, and W. Wang, "Imaging properties and negative refraction of the metallic photonic crystal at near-infrared frequency," *Solid State Commun.*, Vol. 152, No. 11, 929–932, 2012.
11. Zhang, H. F., S. B. Liu, X. K. Kong, C. Chen, and B. R. Bian, "The characteristics of photonic band gaps for three-dimensional unmagnetized dielectric plasma photonic crystals with simple-cubic lattice," *Optic Commun.*, Vol. 288, 82–90, 2013.
12. Zhang, H. F., S. B. Liu, X. K. Kong, and B. X. Li, "Investigation on the dispersion of three-dimensional nomagetized plasma photonic crystals with face-centered-cubic lattices," *Solid State Commun.*, Vol. 162, 34–42, 2013.
13. Deng, X. H., J. T. Liu, J. H. Huang, L. Zou, and N. H. Liu, "Omnidirectional bandgaps in Fibonacci quasicrystals containing single-negative materials," *J. Phys.: Condens. Matter*, Vol. 22, No. 5, 055403, 2010.
14. Zhang, H. F., S. B. Liu, H. Yang, and X. K. Kong, "Analysis of band gap in dispersive properties of tunable three-dimensional photonic crystals doped by magnetized plasma," *Phys. Plasma*, Vol. 20, No. 3, 032118, 2013.
15. Zhang, H. F., S. B. Liu, and B. X. Li, "The properties of photonic band gaps for three-dimensional tunable photonic crystals with simple-cubic lattices doped by magnetized plasma," *Optics & Laser Technol.*, Vol. 50, 93–102, 2013.

A New Triple-band Polarization-insensitive Wide-angle Microwave Metamaterial Absorber

Bo-Rui Bian¹, Shao-Bin Liu^{1,2}, Hai-Feng Zhang¹, Bin-Xiang Li¹, and Ben Ma¹

¹College of Electronic and Information Engineering

Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

²State Key Laboratory of Millimeter Waves, Southeast University, Nanjing 210000, China

Abstract— We propose a new triple-band absorber in microwave frequencies in which the design, analysis, fabrication and measurement of the absorber is presented. The absorber is constructed of a periodic array of tetra-lashing triangle resonators printed on a dielectric material backed by metal ground. By adjusting the geometry parameters of the structure, we can realize a triple-band polarization-insensitive wide-angle absorber in three different resonance frequencies. Experiment results demonstrate excellent absorption rates in the designed frequency bands over wide angles of incident waves for both transverse electric and magnetic polarizations and have good correspondence with the required results.

1. INTRODUCTION

There has been considerable interest in using artificial electromagnetic metamaterial in the design of an absorber with nearly perfect absorption since Landy et al. [1] proposed a perfect absorber composed of electric resonators and cut wires. To date, large efforts and works have been devoted to this area, aiming to design metamaterial absorbers (MMAs) exhibiting excellent performance within single-band [1–4], dual-band [5, 6], multi-band [7–9], as well as broadband [10–12] operations. In addition, the progress of the design has been demonstrated in almost every spectral range, extending from the microwave region to the visible frequencies. Concurrently, MMAs have a host of potential applications including the radar imaging [13, 14], thermal emission [15], detection or sensing [16, 17], and solar cells [18].

In this paper, a new triple-band wide-angle polarization-insensitive microwave metamaterial absorber with tetra-lashing triangle resonators (TLTR) structure has been designed and fabricated. The absorber shows three distinctive absorption peaks at frequencies of 3.07 GHz, 5.65 GHz and 8.11 GHz with absorption rates 99.87%, 99.98% and 99.99%, respectively. The experimental results show excellent absorption rates and the characteristic of polarization-insensitive for a wide range of incidence angles in the designed frequencies, which are in good correspondence with the required results.

2. STRUCTURES DESIGN AND SIMULATIONS

The proposed MMA based on tetra-lashing triangle resonators (TLTR) is presented in Figure 1(a). The top layer consists of tetra-triangle resonators with shorted triangles set in a periodic pattern and the bottom layer is a solid metal. Two metallic layers are separated by a dielectric substrate FR4 with only 1.5 mm thickness, whose relative permittivity is 4.3 and tangent loss is 0.025. All metals in the absorber are made of copper, which has a frequency independent conductivity $\sigma = 5.8 \times 10^7$ S/m. The complete MA is the periodic extension of the unit cell in both x and y directions, as given in Figure 1(b).

Through optimizing the TLTR geometry and the thickness of the dielectric substrate, we can adjust the effective capacitance and inductance of the structure to realize strong resonance. In order to maximize the absorptive efficiency characterized as $A(\omega) = 1 - T(\omega) - R(\omega)$, both the transmission, $T(\omega) = |S_{21}|^2$, and reflection, $R(\omega) = |S_{11}|^2$, should be minimized. Due to the background of structure, the transmission is equal to zero.

A numerical simulation is performed for the MMA in Figure 1(a) with a commercial program, CST MICROWAVE STUDIO 2010. Periodic boundary conditions are used in the x and y directions, and a plane wave is incident downward on the MMA with the electric field polarized along the y -direction (TE) as the excitation source. The simulation result of the absorption curve is depicted in Figure 2(a). It can be seen that there are three nearly unity absorption peaks at $f_1 = 3.07$ GHz, $f_2 = 5.65$ GHz, and $f_3 = 8.11$ GHz, with absorption rates of 99.87%, 99.98%, and 99.99%, respectively.

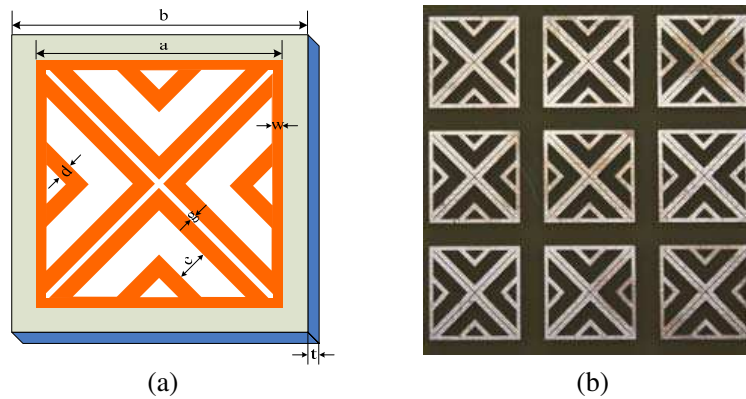


Figure 1: (a) Perspective view of the proposed MMA structure unit cell and geometric dimension parameters: $a = 13$, $b = 18$, $c = 2$, $d = 0.7$, $g = 0.14$, $t = 1.5$, $w = 0.6$. (b) Photograph of fabricated MMA sample. All numbers denote size in millimeters.

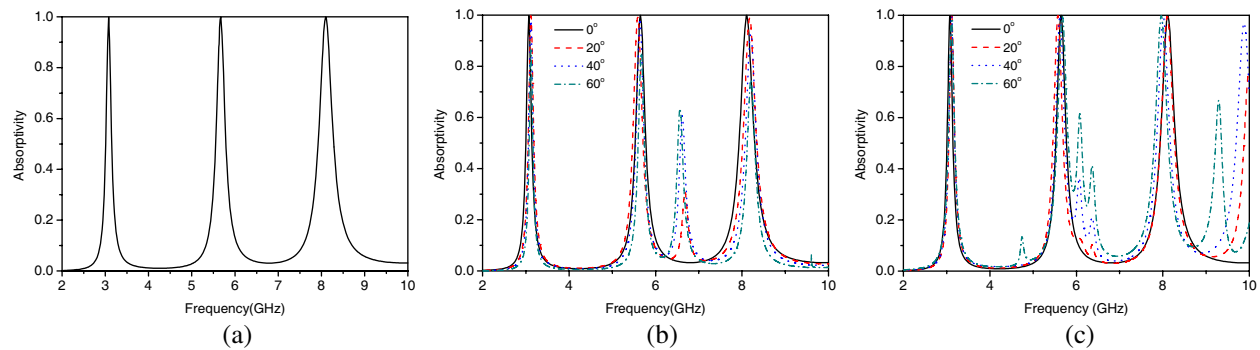


Figure 2: (a) Simulated absorptivity of the TLTR absorber with triple-band configuration. (b) The absorptivity as function of frequency at various angles of incidence for TE incident radiation. (c) TM incident radiation.

The simulated absorption curves of the proposed triple-band MA against frequency for different oblique incidence angles (θ , defined as the angle between the wave vector and the normal) for both TE and TM polarizations are showed in Figure 2(b) and Figure 2(c), respectively.

Following the Figures 2(b) and 2(c), it is obvious that the lowest peak absorption maintains at more than 99% and undergoes a negligible variation when θ varies from 0° to 60° both TE and TM polarization cases. Nevertheless, the condition is slightly different at f_2 and f_3 because the absorption curves vary continuously and a series of ripples are exhibited around f_2 and f_3 . This is induced mainly because of the parasitic resonances of structure which enhance sharply as θ increases, and is partially due to the calculation errors of the simulation engine in the analysis of such a refined structure. To be specific, for the TE case shown in Figure 2(b), at normal incidence the three peak absorption values of 99.87%, 99.98%, and 99.99% are obtained. With increasing angles of oblique incidence, the absorptions of three bands remain quite large at 97.91%, 97.35%, and 92.71% at 40° . Beyond this there is a decrease in the triple-band absorptions at 60° as the incident magnetic field can no longer efficiently drive circulating currents between the two metallic layers. For the case of TM polarization shown in Figure 2(c), the three peak absorption values at normal incidence are close to 1 and remain great than 99% for all angles of oblique incidence. In this case, the magnetic field can efficiently drive the circulating currents at all angles of incidence which is important to maintain impedance matching. In addition, there is also a slight frequency blue-shift of 110 MHz from 0° to 60° . As these simulations reveal, this triple-band MA absorber operates quite well for both TE and TM polarizations over a large range of angles of incidence.

3. EXPERIMENT RESULTS

To experimentally investigate the absorptive properties of the proposed triple-band MA absorber, we have fabricated a MA sample, as shown in Figure 1(a) and Figure 3(a). The unit cells are fabricated using the print circuit board technique on the FR4 substrate, with total footprint of $216 \times 216 \text{ mm}^2$ and thickness of 1.5 mm as that used in numerical simulations. Figure 3 illustrates

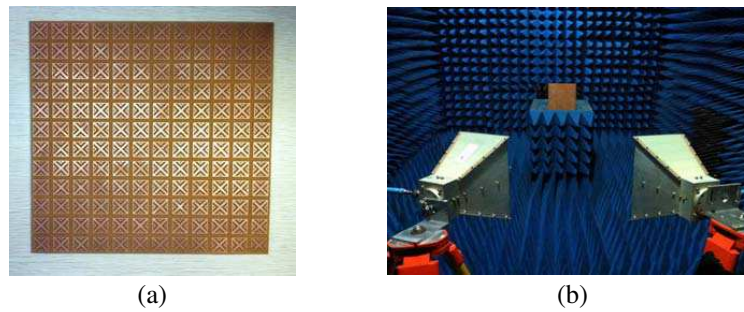


Figure 3: (a) Photographs of the fabricated MA sample. (b) The experimental setup in which the MA sample is placed in the center of flatbase.

the fabricated sample and the experimental simulation setup in which two pairs of horn antennas connecting to the vector network analyzer (Agilent N5230A), We place the MA sample in the center, and moving the transmitting and receiving antennas along the parallel line to measure the reflectance at different angles of incidence.

We first measure the reflection response of a copper sheet to set as a standard reflection response, and then replace the copper sheet with the MMA samples. The disparity of the reflection responses between the copper sheet and the MA sample can be seen as the absorptivity. Figure 4 depicts the measured absorptive spectrum as a function of frequency for TE and TM radiation, respectively. In both TE and TM cases, the EM wave is obliquely incident with different angles ranging from 0° to 60° in steps of 30° .

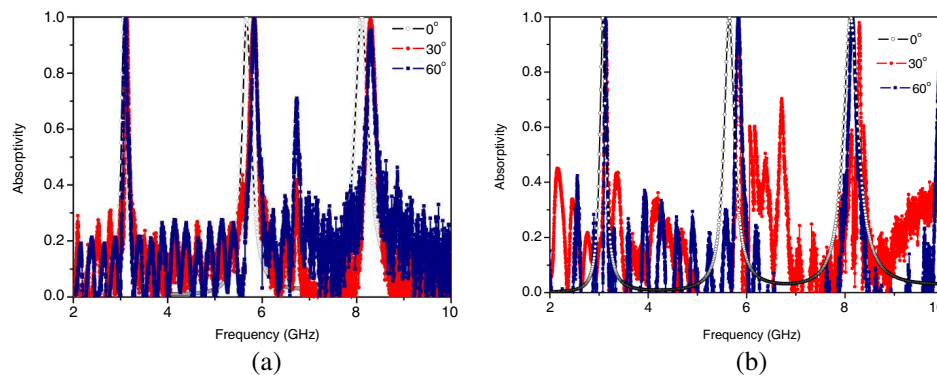


Figure 4: (a) Measured absorption as a function of frequency for (a) TE and (b) TM radiation, respectively, at different angles of incidence.

Through the comparison, we can observe that the experimental results are in excellent agreement with simulations. The absorption peaks gradually decrease as the incident angle increases, yet are still greater than 90%. The slight frequency shift upwards in the measurement case is attributed to the tolerances that are inherent in the fabrication process.

4. CONCLUSIONS

In summary, we proposed a new metamaterial absorber that can operate three different absorptive peak values in microwave region. By manipulating the periodic patterned structures, we can realize a triple-band wide-angle polarization-insensitive absorber at three specific resonance frequencies. The experiments also demonstrated that MMA perform perfectly under different oblique TE and TM incidences. The new MMA can be used to reduce RCS in the aircraft stealth technology. EM spatial filter or bolometer are also promising applications of MMA.

ACKNOWLEDGMENT

This work was supported by the Aviation Science Foundation (Key project 20121852030), Jiangsu Province Natural Science Foundation (Grant No. BK2011727), the Open Research Program in Jiangsu Key Laboratory of Meteorological Observation and Information Processing (Grant No. KDXS1207) and Funding for Outstanding Doctoral Dissertations in NUAU (Grant No. BCXJ11-05).

REFERENCES

1. Landy, N. I., S. Sajuyigbe, J. J. Mock, D. R. Smith, and W. J. Padilla, “Perfect metamaterial absorber,” *Phys. Rev. Lett.*, Vol. 100, 207402-1–4, 2008.
2. Avitzour, Y., Y. A. Urzhumov, and G. Shvets, “Wide-angle infrared absorber based on a negative-index plasmonic metamaterial,” *Phys. Rev. B*, Vol. 79, 045131-1–5, 2009.
3. Cheng, Y., H. Yang, Z. Cheng, and N. Wu, “Perfect metamaterial absorber based on a split-ring-cross resonator,” *Appl. Phys. A*, Vol. 102, 99–103, 2011.
4. Cheng, Y., H. Yang, Z. Cheng, and B. Xiao, “A planar polarization-insensitive metamaterial absorber,” *Photon. Nanostr. Fundam. Appl.*, Vol. 9, 8–14, 2011.
5. Wen, Q. Y., H. W. Zhang, Y. S. Xie, Q. H. Yang, and Y. L. Liu, “Dual band terahertz metamaterial absorber: Design, fabrication, and characterization,” *Appl. Phys. Lett.*, Vol. 95, 241111-1–3, 2009.
6. Singh, P. K., K. A. Korolev, M. N. Afsar, and S. Sonkusale, “Single and dual band 77/95/110 GHz metamaterial absorbers on flexible polyimide substrate,” *Appl. Phys. Lett.*, Vol. 99, 264101-1–4, 2011.
7. Shen, X., Y. Yang, Y. Zang, J. Gu, J. Han, W. Zhang, and T. J. Cui, “Triple-band terahertz metamaterial absorber: Design, experiment, and physical interpretation,” *Appl. Phys. Lett.*, Vol. 101, 154102-1–4, 2012.
8. Zhou, J., L. Jin, and E. Y. B. Pun, “Tunable multi-channel nonreciprocal perfect absorber based on resonant absorption,” *Opt. Lett.*, Vol. 37, No. 13, 2613–2615, 2012.
9. Xu, H. X., G. M. Wang, M. Q. Qi, J. G. Liang, and J. Q. Gong, “Triple band polarization insensitive wide angle ultra-miniature metamaterial transmission line absorber,” *Phys. Rev. B*, Vol. 86, 205104-1–10, 2012.
10. Hendrickson, J., J. P. Guo, B. Y. Zhang, W. Buchwald, and R. Soref, “Wideband perfect light absorber at midwave infrared using multiplexed metal structures,” *Opt. Lett.*, Vol. 37, No. 3, 371–373, 2012.
11. Ding, F., Y. Cui, X. Ge, Y. Jin, and S. He, “Ultra-broadband microwave metamaterial absorber,” *Appl. Phys. Lett.*, Vol. 100, 103506-1–4, 2012.
12. Fallahi, A., A. Yahaghi, H. Benedickter, H. Abiri, M. Shahabadi, and C. Hafne, “Thin wide-band radar absorbers,” *IEEE Trans. on Antennas and Propag.*, Vol. 58, No. 12, 4051–4058, 2010.
13. Noor, A. and Z. Hu, “Metamaterial dual polarised resistive Hilbert curve array radar absorber,” *IET Microw. Antennas Propag.*, Vol. 4, No. 6, 667–673, 2010.
14. Costa, F. and A. Monorchio, “A frequency selective radome with wideband absorbing properties,” *IEEE Trans. on Antennas and Propag.*, Vol. 60, No. 6, 2740–2747, 2012.
15. Chang, Y. C., C. M. Wang, M. N. Abbas, M. H. Shih, and D. P. Tsai, “T-shaped plasmonic array as a narrow band thermal emitter or biosensor,” *Opt. Exp.*, Vol. 17, 13526–13531, 2009.
16. Maier, T. and H. Brückl, “Wavelength-tunable micro-bolometers with metamaterial absorbers,” *Opt. Lett.*, Vol. 34, No. 19, 3012–3014, 2009.
17. Talghader, J. J., A. S. Gawarikar, and R. P. Shea, “Spectral selectivity in infrared thermal detection,” *Light: Science & Applications*, Vol. 1, No. 8, e24, 2012.
18. Ferry, V. E., A. Polman, and H. A. Atwater, “Modeling light trapping in nanostructured solar cells,” *ACS Nano*, Vol. 5, No. 12, 10055–10064, 2011.

Studies on Fano Resonances in Subwavelength Plasmonic Nanostructures

J. Fiala, P. Kwiecien, and I. Richter

Department of Physical Electronics, Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University in Prague, Brehova, 11519 Prague 1, Czech Republic

Abstract— We use an interacting dipole approximation in order to reveal an interplay between resonances localized in individual particles and lattice resonances originating in the array. In this contribution, our first analysis is performed on perfectly conducting array of thin discs. Asymmetrical shapes of extraordinary reflections are achieved, and given into the context with Fano resonances fitted with Fano formula.

1. INTRODUCTION

Recently, by utilizing our experience with the analysis of interactions at the nanoscale of the periodical subwavelength hole array structures and fishnet metamaterials [1], our interest has been turned into characterization of resonances for the case of ‘inversed’ structures, being thus e.g., periodical arrays of metallic nanoparticles. Such subwavelength-scaled plasmonic nanostructures have newly gained much interest for their applications in surface plasmon resonance (SPR) biosensing as well as surface-enhanced Raman scattering (SERS) spectroscopy.

Indeed, it has been observed that metallic nanoparticles can sustain localized plasmon excitations which travel across neighbors. Also in the context of plasmonic resonances, Fano interferences are a powerful tool to exhibit very sharp and strong features due to what is known as lattice resonances, where the interaction between a discrete state and a continuum gives rise to resonant suppression and enhancement of scattering, resulting in a typical asymmetric resonance shape, as was also observed in case of extraordinary optical transmission [2–7].

With our aim being the understanding and revelation of an interplay between resonances localized in individual particles and lattice resonances originating in the array, a simple model that characterizes an isolated building blocks in terms of equivalent induced dipoles of both magnetic and electric character, and a lattice resonances originating in coupling of such dipoles is being developed based on the previous work of Garcia de Abajo [8, 9].

2. MODEL OF COUPLED DIPOLE APPROXIMATION

After an external electric field impinges on a particle at position R_n , an electric dipole moment is induced:

$$\vec{p}_n = \alpha E(\vec{R}_n), \quad (1)$$

where α is the electric polarizability of the particle and $E(\vec{R}_n)$ the total electric field acting on it. The dipole produces an electric field at point r that can be written using the Green function

$$G(\vec{r}) = (k^2 + \nabla\nabla) \frac{e^{ikr}}{r}, \quad (2)$$

as follows

$$G(\vec{r})\vec{p}_n = \frac{e^{ikr}}{r^3} \left[(k^2 r^2 + ikr - 1) \vec{p}_n - (k^2 r^2 + 3ikr - 3) \frac{(\vec{r} \cdot \vec{p}_n) \vec{r}}{r^2} \right], \quad (3)$$

where k is the light momentum in free space. Assuming an periodic array of particles, the dipole moment of each particle is influenced by the electric field of other particles. The total electric field is thus [10]:

$$\vec{p}_n = \alpha \left[\vec{E}_{inc}(\vec{R}_n) + \sum_{n' \neq n} G(\vec{R}_n - \vec{R}_{n'}) \vec{p}_{n'} \right]. \quad (4)$$

Since a plane wave is impinging on the periodic array of particles, the solution can be assumed in the form of $\vec{p}_n = \vec{p} e^{i\vec{k}_{\parallel} \cdot \vec{R}_n}$ where k_{\parallel} is the light momentum parallel to the array of particles.

Consequently, the self-consistent dipole moment reads [11–13]:

$$\vec{p} = \frac{1}{\frac{1}{\alpha} - \sum_{n \neq 0} G(\vec{R}_0 - \vec{R}_n) e^{i\vec{k}_{\parallel} \cdot \vec{R}_n}}. \quad (5)$$

From Bohren & Huffman [14], analytical expressions for the electric polarizability of several particle shapes can, in fact, be obtained and adopted into our model. Finally, for normal incidence, the zero order reflection coefficient in the far field region can be obtained as:

$$r = \frac{2\pi i k / \Lambda^2}{\frac{1}{\alpha} - G_{\text{sum}}}. \quad (6)$$

A limitation of this simple model is that the particles must be small compared to the wavelength and their spacing. The model becomes exact in the limit of narrow holes or small particles in perfect-conductor films. The main advantage of the model lies in its capability of distinguishing between particle and lattice properties.

3. EXAMPLE OF APPLICATION OF THE MODEL

The purpose of application of the coupled dipole model described above is, mainly, to understand rich physics of nanoplasmonic resonances, to help designing and controlling of rigorous calculations, and to achieve optimized efficient platform for sensing [15]. As an example, for the purpose of this contribution, one of the representative structures is formed by identical thin disc nanoparticles of radius ρ which form a periodical array with a lattice constant Λ , as depicted in Fig. 1. Specifically, the lattice constant was chosen to be $\Lambda = 800$ nm, with two radius values for comparison: $\rho = 100$ nm and 160 nm. The fully numerical calculations were performed by using our in-house developed software tool based on the Fourier modal method (rigorous coupled wave analysis [16]).

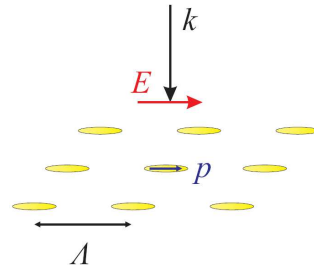


Figure 1: A proposed resonant structure based on surface plasmon resonance, consisting of a periodical array of perfectly conducting thin discs placed in air.

The reflection spectra of the representative structure demonstrate a dramatic change, see Figs. 2(a), 3(a): an extraordinary reflection occurs at the wavelength slightly above the lattice constant; however, the reflectance completely vanishes at the wavelength equals the lattice constant. The latter mentioned behaviour — the presence of narrow bright and dark bands in the spectrum of an reflection grating — was originally observed by Wood and theoretically described by Rayleigh, who assigned the loci of such spectral behaviour to the wavelength where a diffraction order becomes grazing. Here we are, however, focusing on the Babinet complementary problem. The origin of the former is usually interpreted in terms of coupling of the incident light to a surface resonance. In our case of a periodic array of small particles, the coupling of the scattered dipolar field with diffraction modes induces a geometric resonance slightly above the Rayleigh wavelength.

In the spectrum, an asymmetric shape of the resonance curve can be distinguished, signifying a Fano resonance. For a Fano resonance to occur, spectral overlap of broad state or continuum and a narrow discrete state (resonance) is necessary where destructive and constructive interferences take place in a very narrow frequency range. A distinctly asymmetric shape of a Fano resonance then exhibits the following functional form:

$$I_R \propto \frac{(F\gamma + \omega - \omega_0)^2}{(\omega - \omega_0)^2 + \gamma^2}, \quad (7)$$

where ω_0 and γ denote the position and width of the resonance in the frequency scale, respectively, and F is the so-called Fano parameter, describing the degree of asymmetry.

The collective response of dipoles — as discussed above — can be therefore regarded as a lattice surface resonance that is strongly coupled to the propagating light for external plane-wave illumination, being the narrow resonance in the language of Fano; the transmitted light actually represents the continuum. Coupling of these two states results in the resonance with an asymmetric line shape described by the Fano formula (see Eq. (7)). The Fano fits of the reflection spectra, calculated rigorously, are shown in Fig. 4, where the Fano parameters were chosen as $F = 0.10$ and $F = 0.19$, respectively.

The slight difference in resonance lines between the rigorously calculated data and the model

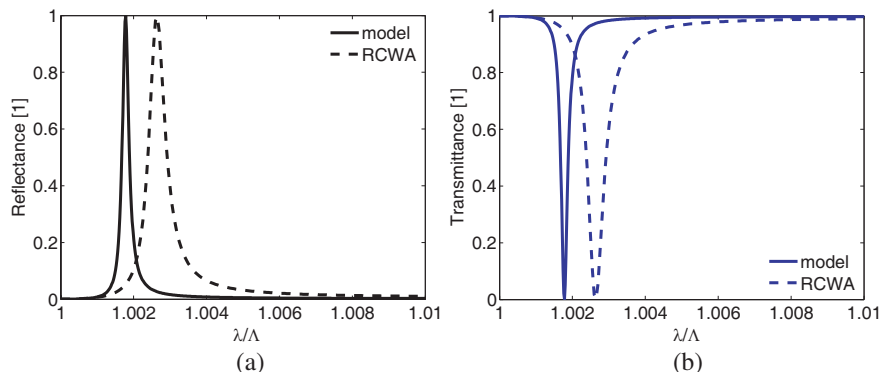


Figure 2: (a) Spectral reflection and (b) transmission characteristic of a proposed nanoarray introduced in Fig. 1. The structure parameters are: perfect conducting thin discs of diameter $\rho = 100$ nm and the lattice constant $\Lambda = 800$ nm.

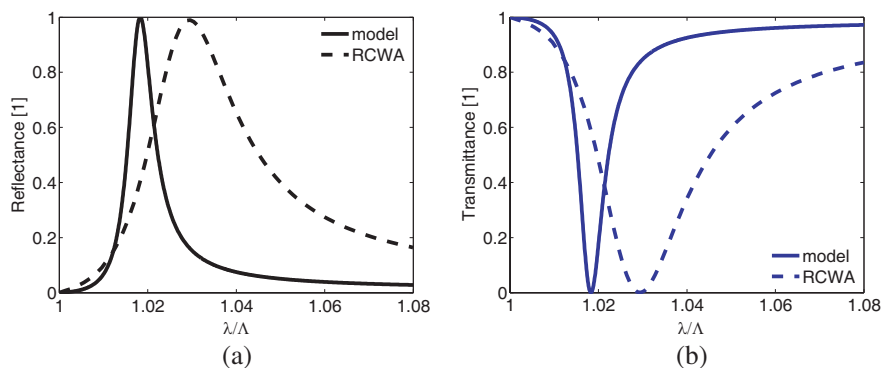


Figure 3: (a) Spectral reflection and (b) transmission characteristic of a proposed nanoarray introduced in Fig. 1. The structure parameters are: perfect conducting thin discs of diameter $\rho = 160$ nm and the lattice constant $\Lambda = 800$ nm.

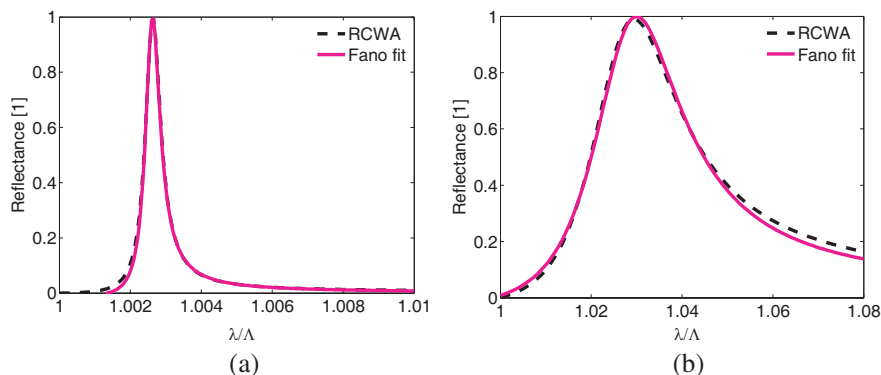


Figure 4: Fit of the resonance lines in the reflection spectra according to Fano formula for (a) $\rho = 100$ nm, and (b) $\rho = 160$ nm disc radius, as compared to the rigorous RCWA calculation.

predictions in Figs. 2 and 3 is due to the definition of the metal material in our numerical tool where an approach has to be chosen in order to simulate a perfect conducting material. The difference therefore lies, strictly speaking, in comparing the non-dispersive and dispersive case, at which the imperfect conductivity of the metal affects the resonance.

4. CONCLUSION

An approximation of coupled dipoles with a clear physical insight was presented in order to reveal fundamental dependences of structure resonances as a starting point for subsequent rigorous calculations. It was demonstrated that the coupled dipole approximation is capable of a good prediction of the resonance positions. The future work will be focused on extending the dipole model to structures with real metals, and on capturing the influence of realistic structure thickness's. In addition, the relationship of the Fano fitting parameters to the actual optical/geometrical properties will be the subject of our studies, too.

ACKNOWLEDGMENT

This work was financially supported by the Czech Science Foundation (projects P205/10/0046 and P205/12/G118).

REFERENCES

1. Fiala, J. and I. Richter, "Interaction of light with subwavelength apertures: A comparison of approximate and rigorous approaches," *Opt. and Quantum Electron.*, Vol. 41, 409–427, 2009.
2. Miroshnichenko, A., S. Flach, and Y. Kivshar, "Fano resonances in nanoscale structures," *Rev. Mod. Phys.*, Vol. 82, 2257–2289, 2010.
3. Luk'yanchuk, B., N. I. Zheludev, S. A. Maier, N. J. Halas, P. Nordlander, H. Giessen, and C. T. Chong, "The Fano resonance in plasmonic nanostructures and metamaterials," *Nat. Mater.*, Vol. 9, 707–715, 2010.
4. Christ, A., Y. Ekinici, H. H. Solak, N. A. Gippius, S. G. Tikhodeev, and O. J. F. Martin, "Controlling the Fano interference in a plasmonic lattice," *Phys. Rev. B*, Vol. 76, 201405, 2007.
5. Augui, B. and W. L. Barnes, "Collective resonances in gold nanoparticle arrays," *Phys. Rev. Lett.*, Vol. 101, 143902, 2008.
6. Francescato, Y., V. Giannini, and S. A. Maier, "Plasmonic systems unveiled by fano resonances," *ACS Nano*, Vol. 6, 1830–1838, 2012.
7. Gallinet, B. and O. J. F. Martin, "Ab initio theory of Fano resonances in plasmonic nanostructures and metamaterials," *Phys. Rev. B*, Vol. 83, 235427, 2011.
8. Garcia de Abajo, F. J., "Site and lattice resonances in metallic hole arrays," *Opt. Exp.*, Vol. 14, No. 1, 7–18, 2006.
9. Garcia de Abajo, F. J., "Colloquium: Light scattering by particle and hole arrays," *Rev. Mod. Phys.*, Vol. 79, 1267–1290, 2007.
10. Collin, R. E. and W. H. Eggimann, "Dynamic interaction fields in a two-dimensional lattice," *IRE Trans. Microwave Theory Tech.*, Vol. 9, 110–115, 1961.
11. Draine, B. T. and P. J. Flatau, "Discrete-dipole approximation for scattering calculations," *J. Opt. Soc. Am. A*, Vol. 11, 1491–1499, 1994.
12. Zhou, F., Y. Liu, and Z. Y. Li, "Surface-plasmon-polariton-assisted dipole-dipole interaction near metal surfaces," *Opt. Lett.*, Vol. 36, 1969–1971, 2011.
13. Yurkin, M. A. and A. G. Hoekstra, "The discrete dipole approximation: An overview and recent developments," *J. Quant. Spectr. Radiat. Transf.*, Vol. 106, 558–589, 2007.
14. Bohren, C. F. and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley-Interscience, New York, 1983.
15. Homola, J., *Surface Plasmon Resonance Based Sensors*, Springer-Verlag, 2006.
16. Ctyroky, J., P. Kwiecien, and I. Richter, "Fourier series-based bidirectional propagation algorithm with adaptive spatial resolution," *J. Lightwave Technol.*, Vol. 28, 2969–2976, 2010.

Spectral Density Analysis of Thin Gold Films: Thickness and Structure Dependence of the Optical Properties

P. C. Lansåker¹, E. Tuncer², I. Valyukh³, H. Arwin³,
C. G. Granqvist¹, and G. A. Niklasson¹

¹The Ångström Laboratory, Department of Engineering Sciences
Uppsala University, P. O. Box 534, SE-751 21 Uppsala, Sweden

²3M Austin Center, EEBG/CRML, Austin, TX 78726, USA

³Laboratory of Applied Optics, Department of Physics, Chemistry and Biology
Linköping University, SE-58183 Linköping, Sweden

Abstract— In this paper we study the feasibility of representing the optical properties of ultrathin gold films by effective medium theories. Gold films with mass thicknesses in the range of 1.4 to 9.2 nm were deposited by DC magnetron sputtering onto non-heated glass substrates. Optical measurements in the range 0.25 to 2 μm were carried out by spectroscopic ellipsometry, and the effective complex dielectric function of each film was determined. The gold films were modelled as a mixture of gold and air, and a general effective medium description using the spectral density function (SDF) was used to describe their optical properties. Numerical inversion of the experimental dielectric function gave a broad and rather featureless SDF, with a few superimposed peaks, both for island structures and percolating films. The broad background is qualitatively similar to predictions of the Bruggeman model [14].

1. INTRODUCTION

Gold-based nanoparticles and thin films are of large interest for applications in green nanotechnology [1]. Their optical properties have been described either by the island film theory involving surface susceptibilities [2] or by treating them as a mixture of gold and air using effective medium theories [3]. Recent studies of island and network films have shown that such films can be represented by an effective dielectric function together with an effective thickness of the same order of magnitude as the nominal thickness [4, 5]. The island film theory seems to have some advantages regarding the comparison between theory and experiments for films consisting of nanoparticles [5], while effective medium theories (EMTs) are commonly used to model the optical properties of nanocomposites. However, EMTs are very sensitive to the actual nanostructure of the film. The structure of the films can be rigorously represented by a spectral density function (SDF) [6, 7], which can be obtained by numerical inversion of experimental data [8–10]. In the present paper we determine the dielectric function of gold nanoparticle films and network gold films with mass thickness d_{mass} ranging from 1.4 nm to 9.2 nm. We obtain the SDF from the experimental data and discuss the physical implications of the shape of the SDF.

2. THEORY — SPECTRAL DENSITY ANALYSIS

Non-homogenous gold films can be considered as nanocomposites consisting of gold and air [3, 11]. If the dimensions of the structures are much smaller than the wavelength of the incident radiation, one can describe the optical properties of these nanocomposites by an effective dielectric function ε^{eff} which depends on the dielectric functions of the constituents as well as on the details of the structure [3, 12]. Nanocomposites can exhibit a number of different structures that affect the optical properties of the composites, and SDF analysis is an excellent analytic method for studies of these structural impacts in two-phase composites [6, 7, 10]. For a mixture with inclusions, the effective dielectric function can be written in terms of the spectral density function $g(x)$ as [8–10]

$$\varepsilon^{eff} = \varepsilon_h \left(1 + \int_0^1 \frac{g(x)}{1 + \left(\frac{\varepsilon_p}{\varepsilon_h} - 1\right) x} dx \right), \quad (1)$$

where ε_h and ε_p are the dielectric functions of the host material and the “particles”, respectively. In the SDF analysis, the resonance with spectral parameter $x = 0$ depends on the DC conductivity and is called the percolation strength ξ_{perc} . In our case we obtained ε_{eff} from ellipsometric measurements and used Eq. (1), together with known or independently measured dielectric functions

of the constituents, to derive $g(x)$. The inversion of the experimental ε^{eff} , in order to obtain $g(x)$, is difficult and needs advanced computational techniques. As in earlier papers [8, 10], we employ a Monte-Carlo technique based on Bayesian numerical statistics, where the values of x are varied in each set.

The volume fraction of the “particle” component, f , can be obtained by the sum rule

$$f = \xi_{perc} + \xi_{nonperc} = \xi_{perc} + \int_{0+}^1 g(x) dx. \quad (2)$$

In order to discuss the shape of the SDF, we note that the commonly used Maxwell-Garnett (MG) [13] and Bruggeman (Br) [14] effective medium theories can be represented by $g(x)$ as given by [8]

$$g_{MG}(x) = f\delta(x - L(1 - f)), \quad (3)$$

and [15]

$$g_{Br} = [-9x^2 + 6(1 + f)x - (3f - 1)^2]^{1/2} / 4\pi x, \quad (4)$$

respectively. It is realized that MG theory exhibits a resonant peak in the SDF, while the Br theory has a very broad and featureless peak [16].

3. SAMPLES

3.1. Thin Film Deposition

Nine gold films were produced by DC magnetron sputtering onto non-heated glass substrates in a versatile deposition system [17]. Results from three of them — with island structure, a meandering network, and a homogenous configuration — will be discussed in detail. Rutherford backscattering spectroscopy was performed on four of the gold films, and the mass thickness was determined from the results. The remaining five values of d_{mass} were obtained by linear scaling with deposition time.

3.2. Structural and Electrical Characterization

The structure and porosity of the gold films were studied by scanning electron microscopy (SEM), employing a LEO 1550 FEG instrument with in-lens detection. The experimental area fraction covered by gold, f_{SEM} , of the films was estimated from a MATLAB-based analysis of SEM images, and the relation $d_{geom} = d_{mass}/f_{SEM}$ gave the geometric thickness d_{geom} for all films. Sheet resistance R was recorded between sputtered gold contacts, and the electrical conductivity σ_{DC} of the films was computed. This allows us to obtain experimental values of ξ_{perc} from the relation

$$\xi_{perc} = \frac{\sigma_{DC}}{\sigma_{bulk,DC}}. \quad (5)$$

The value of $\sigma_{bulk,DC}$ was approximated by the conductivity of the gold film with $d_{mass} \approx 6.6$ nm, which was the most conducting film. This procedure was necessary in order to take into account the

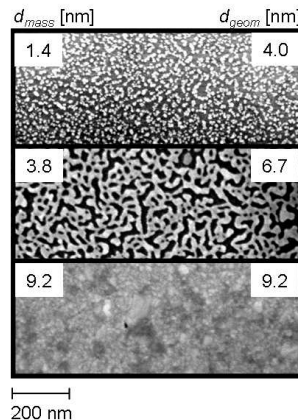


Figure 1: Scanning electron micrographs of thin gold films. The values of d_{mass} , d_{geom} and f_{SEM} were 1.4 nm, 4.0 nm and 0.363 for the upper image; 3.8 nm, 6.7 nm and 0.563 for the middle image, and 9.2 nm, 9.2 nm and 0.999 for the lower image, respectively.

increased grain boundary scattering in thin films, which leads to “bulk” conductivities considerably lower than values tabulated for polycrystalline gold.

SEM images of three gold films are illustrated in Figure 1, showing an island structure for $d_{mass} \approx 1.4$ nm, a meandering film for $d_{mass} \approx 3.8$ nm, and a homogenous film for $d_{mass} \approx 9.2$ nm. Image analysis of the SEM pictures gave values of f_{SEM} for the films, from which d_{geom} was calculated and presented in Figure 1.

4. OPTICAL CHARACTERIZATION

4.1. Ellipsometry Data

Ellipsometric parameters Ψ and Δ in the 0.25 to 2.0 μm range were measured in reflection mode at several angles of incidence between 50° and 75° using a variable-angle rotating analyzer ellipsometer (VASE from J. A. Woollam Co., Inc.) [4, 5, 18]. In addition, spectral transmittance was measured at normal incidence. The dielectric functions of the non-homogenous gold films were represented by a few (two or three) Lorentz oscillators as well as by a Drude oscillator for network and continuous films [4, 5, 18, 19]. The thicknesses of the films were fixed and equal to d_{geom} . The effective dielectric function of the films was obtained from the best fitting parameters.

Figure 2 shows dispersions of the dielectric functions for nanoparticle-type, meandering and homogenous films, as obtained by spectroscopic ellipsometry. The underlying values of the parameters of Lorentz and Drude oscillators, as well as confidence limits, will be presented elsewhere.

4.2. Spectral Density Function Analysis

Figure 2 shows that the experimental data could be accurately fitted by the SDF equation for the effective dielectric function in Eq. (1) for all films, with only minor discrepancies for nanoparticle films. Figure 3 shows $g(x)$ for films with d_{mass} of 1.4, 3.8 and 9.2 nm. The figure displays similar features for all three $g(x)$, which is surprising in view of the large differences in the nanostructure of the films.

The most notable feature in Figure 3 is the broad structure extending from -2.5 to -0.5 on a logarithmic x -scale. Such a broad response is a characteristic feature of the Bruggeman theory, Eq. (4), and we conclude, in agreement with previous work [3], that the Bruggeman theory provides a good approximation of the optical properties of nanostructured gold films. Secondly, there are one or two peaks superimposed on the broad background. These peaks can be interpreted as broadened localized plasmon resonances, such as those occurring in the MG theory, and they become less pronounced for the thicker films that are more continuous. Thirdly, there is an increase of $g(x)$ at large values of x , being close to unity. However, the main features of resonances at $x > 0.5$ fall outside of our experimental wavelength range. A similar feature is observed in evaluating distributions of relaxation times for dielectric spectroscopy, where the data outside the experimental window result

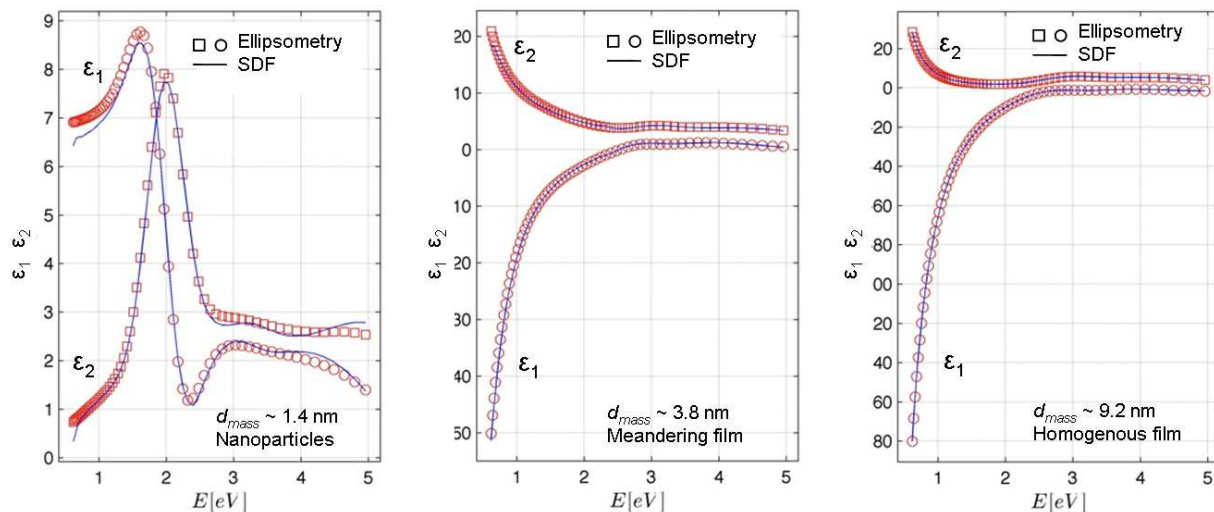


Figure 2: Real (ϵ_1) and imaginary (ϵ_2) parts of the dielectric function for a nanoparticle film with $d_{mass} \approx 1.4$ nm, a meandering film with $d_{mass} \approx 3.8$ nm, and a homogenous film with $d_{mass} \approx 9.2$ nm. Symbols denote experimental data and curves are fits to the SDF-based EMT in Eq. (1).

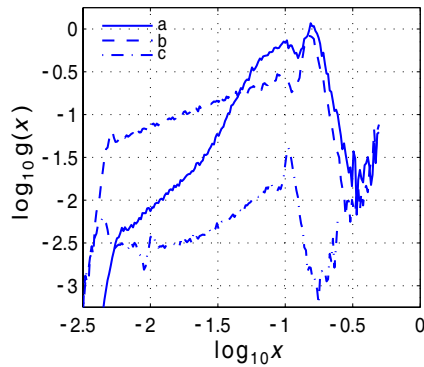


Figure 3: Spectral density functions $g(x)$ for gold films with (a) $d_{mass} \approx 1.4$ nm, (b) $d_{mass} \approx 3.8$ nm and (c) $d_{mass} \approx 9.2$ nm.

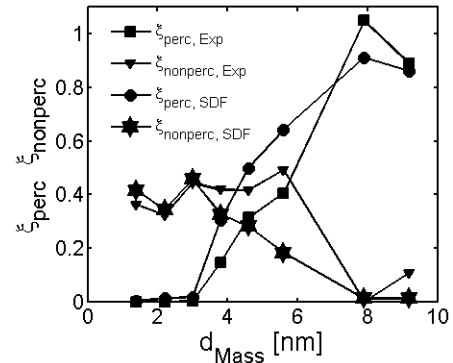


Figure 4: Parameters ξ_{perc} and $\xi_{nonperc}$ from experimental results and SDF analysis, as a function of mass thickness.

in spurious high dielectric relaxation strengths because of lack of data at those frequencies. In our case, spectral frequency values over 0.5 could not model the data within the measured experimental window and yielded high spectral density values which should be disregarded (the spectral frequency is defined as $(\epsilon_p - \epsilon_h)/\epsilon_h$).

Figure 4 compares the experimental percolation and non-percolation strengths, Eqs. (5) and (6), of gold in the films with those inferred from the SDF analysis, Eq. (2). The experimental values of $\xi_{nonperc}$ were obtained from

$$\xi_{nonperc} = f_{SEM} - \xi_{perc}. \quad (6)$$

Figure 4 shows that the percolation contribution to f is zero for the nanoparticle films, then increases steeply around the percolation threshold, and becomes high for the continuous films. The non-percolating component decreases with film thickness and is close to zero for the continuous films. These trends are as expected, and the SDF estimates are reasonably close to the experimental values derived from Eqs. (5) and (6).

5. DISCUSSION AND CONCLUSION

We studied three gold films — with island, network and homogenous film structure — and derived their effective dielectric functions from ellipsometry. We performed a spectral density function analysis, which showed similar features in $g(x)$ despite the various nanostructures of the films. A common broad feature indicates that the Bruggeman effective medium model is suitable as a first approximation to the optical properties of nanostructured gold films. In order to employ SDF analysis as a standard characterization tool, a number of theoretical problems have to be solved, and most importantly the features in the $g(x)$ function must be interpreted in terms of specific structural features in the samples.

ACKNOWLEDGMENT

Financial support was received from the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS). The RBS measurements were performed at the Tandem Laboratory at Uppsala University. We gratefully acknowledge Daniel Primetzhofer for advice on RBS and Pär Lansåker for guidance regarding the MATLAB-based image-analysis program.

REFERENCES

1. Smith, G. B. and C. G. Granqvist, *Green Nanotechnology: Solutions for Sustainability and Energy in the Built Environment*, CRC Press, Boca Raton, FL, USA, 2011.
2. Bedeaux, D. and J. Vligter, *Optical Properties of Surfaces*, Imperial College Press, London, UK, 2004.
3. Smith, G. B., G. A. Niklasson, J. S. E. M. Svensson, and C. G. Granqvist, *J. Appl. Phys.*, Vol. 59, 571, 1986.
4. Loncaric, M., J. Sancho-Parramon, and H. Zorc, *Thin Solid Films*, Vol. 519, 2946, 2011.
5. Mendoza-Galvan, A., K. Järrendahl, A. Dmitriev, T. Pakzieh, M. Käll, and H. Arwin, *Opt. Express*, Vol. 19, 12093, 2011.

6. Bergman, D. J., *Phys. Rep.*, Vol. 43, 377, 1978.
7. Bergman, D. J., *Ann. Phys.*, Vol. 138, 78, 1982.
8. Tuncer, E., *Phys. Rev. B*, Vol. 71, 012101, 2005.
9. Tuncer, E., *Materials*, Vol. 3, 585, 2010.
10. Niklasson, G. A., T. K. Boström, and E. Tuncer, *Proc. SPIE*, Vol. 8168, 81680S, 2011.
11. Milton, G. W., *The Theory of Composites*, Cambridge University Press, Cambridge, UK, 2002.
12. Sancho-Parramon, J., S. Bosch, A. Abdolvand, A. Podlipensky, G. Seifert, and H. Graener, *Proc. SPIE*, Vol. 5963, 596320, 2005.
13. Garnett, J. C. M., *Philos. Trans. Roy. Soc.*, Ser. A, Vol. 203, 237, London, 1904.
14. Bruggeman, D. A. G., *Ann. Phys.*, 5th Series, Vol. 24, 636–679, Leipzig, 1935.
15. Ghosh, K. and R. Fuchs, *Phys. Rev. B*, Vol. 38, 5222, 1988.
16. Theiß, W., *Adv. Solid State Phys.*, Vol. 33, 149, 1993.
17. Lansåker, P. C., J. Backholm, G. A. Niklasson, and C. G. Granqvist, *Thin Solid Films*, Vol. 518, 1225, 2009.
18. Tompkins, H. G. and E. A. Irene, Editors, *Handbook of Ellipsometry*, William Andrew Publishing, Norwich, NY, 2005.
19. *Guide to Using WVASE32*, J. A. Woollam Co., Inc., Lincoln, NE, USA, 1995.

Variable Forbidden Regions in Metamaterial Planar Waveguide with Nonlinear Cladding

Yaw-Dong Wu, Ming-Hsiung Cheng, and Tien-Tsorng Shih

Department of Electronic Engineering, National Kaohsiung University of Applied Sciences
Kaohsiung, Taiwan, R.O.C.

Abstract— In this paper, we used the modal theory to analyze TE waves in the three-layer planar waveguide with the metamaterial guiding film and the nonlinear cladding layer, and derive a formula for the guiding power density. The power density is very helpful to analyze the forbidden regions of the proposed waveguide structure. In recent years, the metamaterial in the planar waveguide has been widely studied in many researches. The feature of the metamaterial includes the permittivity ϵ and the permeability μ less than zero. The forbidden regions always exist in the three-layer planar waveguide with the metamaterial guiding film and the nonlinear cladding layer. In this study, the guiding power density can help to provide a clear picture of forbidden regions in the metamaterial film with different permeability values. The numerical results show that the forbidden regions are different in metamaterial film with the same width but the permeability values are different. The greater permeability value is, the narrower the forbidden region is. The study shows that the value of guiding power density is less than zero at metamaterial and is greater than zero at right-handed medium. Therefore, the guiding power density is at the opposite direction because the values of permeability are positive and negative. It agrees with the phenomenon of negative refraction in metamaterial. The simulation result is helpful to find out the propagation situation of the TE polarized waves

1. INTRODUCTION

In 1968, the left-handed material (LHM) had been presented by Veselago and fabricated in experiment by Shelby et al. [1, 2]. The LHM possesses simultaneously negative dielectric permittivity and permeability [3, 4] and has a negative refractive index. Some analyses of the three-layer planar waveguide with LHM film have been proposed [5]. In the symmetric slab waveguide structure with a core of LHM was investigated [6] and the forbidden region of the three-layer and multilayer planar waveguide with LHM film had been analyzed [7, 8].

Here we investigate the forbidden region of the three-layer planar waveguide with the LHM film and the nonlinear cladding by using the modal theory [9–11]. The model theory is used to find out the forbidden region in the planar waveguide with LHM and the nonlinear cladding. By solving Maxwell's equation, matching the boundary, and poynting vector, the field of the transverse electric polarized wave and guiding power density will be obtained. We used the power density curves to obtain the forbidden regions of the three-layer planar waveguide with LHM film and the nonlinear cladding. The forbidden regions are usually used with respect to discuss the relation of propagation constant and power density

2. ANALYSIS

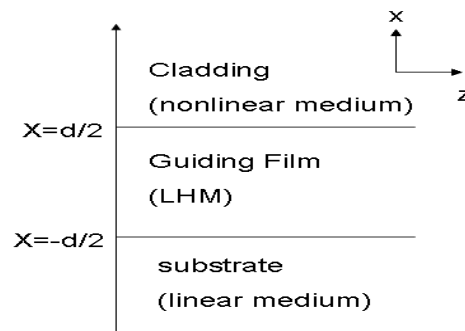


Figure 1: The structure of metamaterial planar waveguide with the nonlinear cladding.

In this section, we use the model theory to drive the TE-polarized field for the three-layer planar waveguide with metamaterial guiding film and the nonlinear cladding layer. The structure is shown

in Fig. 1. We consider the transverse electric polarized waves propagating along z direction. The wave equation can be reduced to as follows.

$$\nabla^2 E_{yj} = \frac{n_j}{c^2} \frac{\partial^2 E_{yj}}{\partial t^2}, \quad j = f, c, s \quad (1)$$

The solution of the Eq. (1) can be expressed as:

$$E_{yj}(x, z, t) = E_j(x) \exp[j(\omega t - \beta k_o z)], \quad j = f, c, s \quad (2)$$

The subscripts f , c and s in Eqs. (1)–(2) are used to denote the guiding film with LHM, nonlinear cladding, and linear substrate, respectively. In Eq. (2), β is the effective refractive index, ω is the angular frequency, and k_0 is the wave number in the free space. For the Kerr-type nonlinear medium, the square of the refractive index n_c^2 can be expressed as:

$$n_c^2 = n_{co}^2 + \alpha |E(x)|^2 \quad (3)$$

where n_{co} in the linear refractive index of the nonlinear medium and α is the nonlinear coefficient. By substituting Eq. (2) into the wave equation Eq. (1), the transverse electric fields in each layer has the general form.

$$E_c(x) = \sqrt{\frac{2}{\alpha}} A_c q_c / \cosh[k_o q_c (x - x_c)] \quad x \geq d/2, \text{ in the nonlinear cladding} \quad (4)$$

$$E_f(x) = A_f \cos[k_o q_f (x - x_f)] \quad -d/2 \leq x \leq d/2, \quad \beta < n_f, \text{ in the guiding film} \quad (5a)$$

$$E_f(x) = A_f \cosh[k_o Q_f (x - x_f)] \quad -d/2 \leq x \leq d/2, \quad \beta > n_f, \text{ in the guiding film} \quad (5b)$$

$$E_s(x) = A_s \exp(k_o q_s x) \quad x \leq -d/2, \text{ in the substrate} \quad (6)$$

The q_c , q_f , Q_f and q_s can be expressed as

$$q_j = \sqrt{\beta^2 - n_j^2} \quad j = i, c, s$$

$$q_f = \sqrt{n_f^2 - \beta^2}$$

$$Q_f = \sqrt{\beta^2 - n_f^2}$$

The A_c , A_f , A_s , q_c , q_f , Q_f , q_s , x_f , and x_c are all constants, these constants can be obtained by matching the boundary conditions. The constant A_c is arbitrary. The amplitude parameters A_f and A_s are proportional to A_c , it's simplicity to assume $A_c = 1$ for numerical method. By matching the boundary conditions, we obtained the following equations:

$$\tan(k_o q_f d_f) = \frac{q_f q_s \left(\frac{\mu_f}{\mu_s}\right) + q_f q_c \left(\frac{\mu_f}{\mu_c}\right) \tanh\left[k_o q_c \left(\frac{d_f}{2} - x_c\right)\right]}{q_f^2 - q_c q_s \left(\frac{\mu_f}{\mu_s} \frac{\mu_f}{\mu_c}\right) \tanh\left[k_o q_c \left(\frac{d_f}{2} - x_c\right)\right]} \quad (7)$$

$$\tanh(k_o Q_f d_f) = \frac{Q_f q_s \left(\frac{\mu_f}{\mu_s}\right) - Q_f q_c \left(\frac{\mu_f}{\mu_c}\right) \tanh\left[k_o q_c \left(\frac{d_f}{2} - x_c\right)\right]}{Q_f^2 - q_c q_s \left(\frac{\mu_f}{\mu_s}\right)^2 \tanh\left[k_o q_c \left(\frac{d_f}{2} - x_c\right)\right]} \quad (8)$$

The transcendental Eqs. (7) and (8) can be solved by numerical method on a computer. When the constant β is determined, the constants q_c , q_f , Q_f , q_s are also determined.

The energy density of the guiding film can be calculated by $S = 1/2 \text{Re}[E_f(x) \times H_f(x)^*]$ and the energy flux is along the z direction. Substitute Eq. (5) into pointing vector averaged, we can obtained guiding power density P_{av} as follows.

$$P_{av} = \frac{\beta}{\mu_f} \int_{-\infty}^{\infty} \frac{c \varepsilon_0 E_f^2(x)}{2} \quad (9)$$

From the Eq. (9) we find out the guiding power density depends on β and μ_f parameter.

3. NUMERICAL RESULTS

In this section, we used the guiding power density formula derived in Eq. (9) to find the distribution of the forbidden regions. The numerical results are shown in Figs. 2–4. In the Fig. 2, the guiding power density curve versus β is plotted with the wavelength in free space $\lambda = 1.3 \mu\text{m}$, $n_{co} = n_s = 1.55$, $n_f = 1.57$, $d = 2 \mu\text{m}$, $u_f = -1$, $u_f = -5$, $\alpha = 6.3786 \times 10^{-12} \text{ m}^2/\text{V}^2$. In the Fig. 3, the guiding power density curve versus β is plotted with the wavelength in free space $\lambda = 1.3 \mu\text{m}$, $n_{co} = n_s = 1.55$, $n_f = 1.57$, $d = 4 \mu\text{m}$, $u_f = -1$, $u_f = -5$, $\alpha = 6.3786 \times 10^{-12} \text{ m}^2/\text{V}^2$. In the Fig. 4, the guiding power density curve versus β is plotted with the wavelength in free space $\lambda = 1.3 \mu\text{m}$, $n_{co} = n_s = 1.55$, $n_f = 1.57$, $d = 5 \mu\text{m}$, $u_f = -1$, $u_f = -5$, $\alpha = 6.3786 \times 10^{-12} \text{ m}^2/\text{V}^2$. In Fig. 2, the curve of the guiding power density is less zero at RHM medium, is greater than zero at LHM medium, and is opposite direction. In Figs. 3–4, the curve of the guiding power density is plotted, the greater permeability value is, the narrower the forbidden region is, and the forbidden regions are variable from permeability $u_f = -1$ to $u_f = -5$.

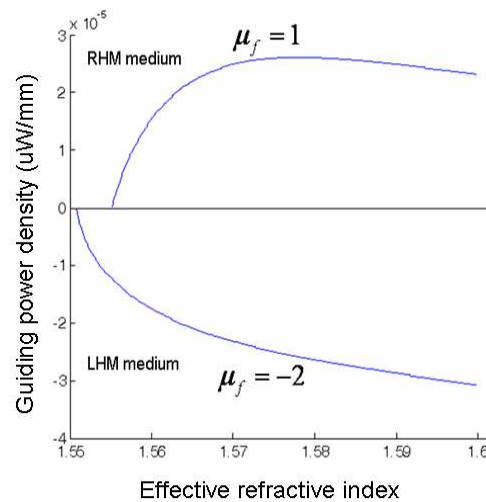


Figure 2: The power density is opposite direction at LHM medium and RHM medium.

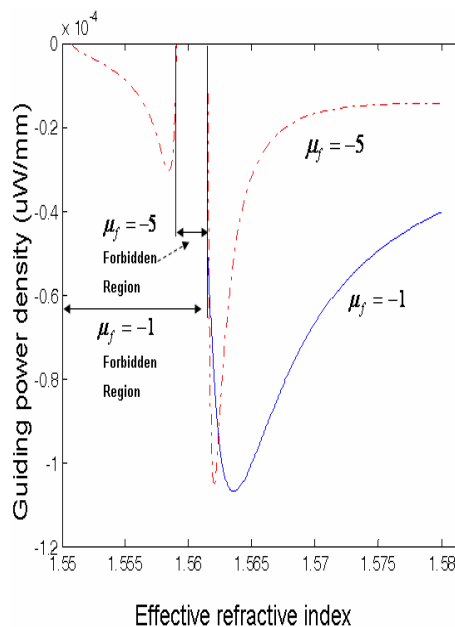


Figure 3: Variable forbidden region from $u_f = -1$ to $u_f = -5$ at $d = 4 \mu\text{m}$.

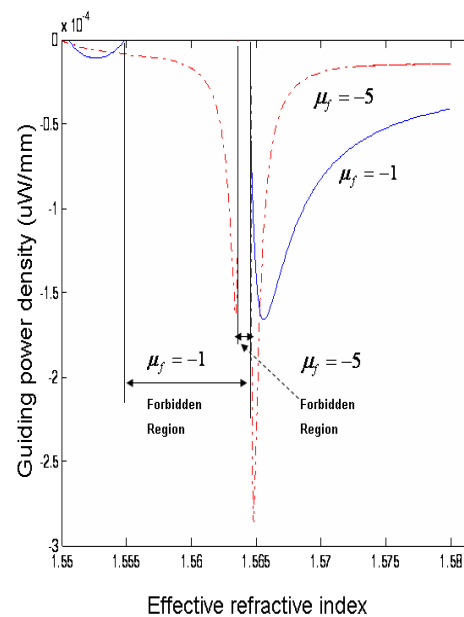


Figure 4: Variable forbidden region from $u_f = -1$ to $u_f = -5$ at $d = 5 \mu\text{m}$.

4. CONCLUSION

We have analyzed the transverse electric field of each layer in three-layer metamaterial planar waveguide with nonlinear cladding. We also derived the formula of the guiding power density to analyze the forbidden region. The numerical results show that the forbidden regions always exit in the three-layer planar waveguide with the LHM guiding film and the nonlinear cladding. The greater permeability value is, the narrower the forbidden region is. The numerical results show that the forbidden regions are different in metamaterial film with the same width but the permeability values are different. The greater permeability value is, the narrower the forbidden region is. The study shows that the value of guiding power density is less than zero at metamaterial and is greater than zero at right-handed medium. Therefore, the guiding power density is at the opposite direction because the values of permeability are positive and negative. It agrees with the phenomenon of negative refraction in metamaterial. The simulation result is helpful to find out the propagation situation of the TE polarized waves.

ACKNOWLEDGMENT

This work was partly supported by National Science Council R.O.C. and under Grant No. 101-2221-E-151-077.

REFERENCES

1. Shelby, R. A., D. R. Smith, and S. Schultz, "Experimental verification of a negative index of refraction," *Science*, Vol. 292, 77, 2001.
2. Veselago, V. G., "The electrodynamics of substances with simultaneously negative values of ϵ and μ ," *Sov. Phys. Usp.*, Vol. 10, 509, 1968.
3. Ziolkowski, R. W. and E. Heyman, "Wave propagation in media having negative permittivity and permeability," *Phys. Rev. E*, Vol. 64, 056625, 2001.
4. Milonni, P. W. and G. J. Maclay, "Quantized-field description of light negative-index media," *Opt. Commun.*, Vol. 228, 161–165, 2003.
5. Peacock, A. C. and N. G. R. Broderick, "Guided modes in channel waveguides with a negative index of refraction," *Opt. Express*, Vol. 11, 2502–2510, 2003.
6. He, Y., Z. Cao, and Q. Shen, "Guided optical modes in asymmetric left-hand waveguides," *Opt. Commun.*, Vol. 245, 125–135, 2005.
7. Kuo, C. W., S. Y. Chen, Y. D. Wu, and M. H. Chen, "Analyzing the multilayer optical planar waveguides with double-negative metamaterial," *Progress In Electromagnetics Research*, Vol. 110, 163–178, 2010.
8. Wu, Y. D., M. H. Cheng, and T. T. Shih, "Analysis of the forbidden region for multilayer planar waveguide with LHM," *PIERS Proceeding*, 313–316, Moscow, Russia, Aug. 19–23, 2012.
9. Doer, C. R. and H. Kogelnik, "Dielectric waveguide theory," *J. Lightwave Technol.*, Vol. 26, 1176–1187, 2008.
10. Wu, Y. D. and M. H. Chen, "Method for analyzing multilayer nonlinear optical waveguide," *Opt. Express*, Vol. 13, 7982–7995, 2005.
11. Kuo, C. W., S. Y. Chen, M. H. Chen, C. F. Chang, and Y. Wu, "Analyzing multilayer optical waveguide with all nonlinear layers," *Opt. Express*, Vol. 15, 2499–2516, 2007.
12. Wu, Y. D., M. H. Chen, C. K. Kuo, S. Y. Chen, and C. F. Chang, "The study of multilayer planar optical waveguide structure with nonlinear cladding," *Optical and Quantum Electronics*, Vol. 40, 495–512, 2008.

New All-optical Switch Based on the Local Nonlinear Plasmonic Mach-Zehnder Interferometer Waveguides

Yaw-Dong Wu, Sheng-Rong Hong, and Tien-Tsornng Shih

Department of Electronic Engineering

National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan, R.O.C.

Abstract— In this paper, we proposed a new all-optical switch based on the local nonlinear metal-insulator-metal (MIM) plasmonic Mach-Zehnder (MZ) interferometer waveguides. The proposed structure is composed of the local Kerr-type nonlinear MIM plasmonic MZ interferometer waveguide and the straight MIM plasmonic control waveguide. In recent year, the surface plasmon has attracted much attention due to the potential of application. Surface plasmon polariton (SPP) is an electromagnetic excitation at the interface of dielectrics and metals, which can be used to control light in subwavelength. The local Kerr-type nonlinear MIM plasmonic MZ interferometer waveguide is used in this switching structure for its large third-order nonlinear susceptibility and ultrafast response properties. Some important optical properties of the switching structure have been studied numerically by the finite-difference time-domain method (FDTD), including the transmission characteristics and the switching effect. The proposed new all-optical switching structure shows great advantages of the ultracompact device size and the ultrafast switching effect. By properly fixing the input signal power and changing the control power, the output signal beam will show the switching effect. In this structure we can found that the proposed all-optical switch exhibits an excellent switching effect. It would be a potential key component in the application of the ultra-high-speed data processing system.

1. INTRODUCTION

In recent years, the development of optical components is growing prosperity. Accordingly, all-optical signal processing and control also becomes important. All-optical signal processing in integrated photonic circuits and its applications in optical computing and communications require the ability to control light with light [1]. Therefore, all-optical switching devices and waveguide with nonlinear Kerr effect logic gate has begun to attract attention. Several all-optical switching and logic devices using optical nonlinearity have ever been proposed and implemented [2, 3]. Most of the conventional all-optical devices are based on uniformly nonlinear structure. In optical waveguide structure of uniform nonlinearity, several interesting optical properties have shown, however, there are still more attractive propagation characteristics in waveguide structures combined a nonlinear material with a linear one. Mach-Zehnder waveguide interferometer device has been developed for use of modulating, switching, and logic gates, etc.. Most of them are operated by the principles of electric-optic effect where the change of refractive index in the arms of mach-zehnder interferometer is produced by applying the external electric field.

In this paper, we propose and numerically investigate a novel nanoscale all-optical switching based on the local nonlinear metal-insulator-metal (MIM) plasmonic Mach-Zehnder (MZ) interferometer waveguides with the straight pumping waveguide. The Kerr nonlinear medium in the structure is changed by varying the pump light and the signal transmission can thus be controlled.

2. ANALYSES AND NUMERICAL RESULTS

Figure 1 shows the structure of the proposed all-optical switch, which consists the metal claddings, the local nonlinear metal-insulator-metal plasmonic Mach-Zehnder interferometer waveguides, and the straight pumping waveguide. The metal is assumed as silver, whose relative permittivity can be described by the well-known Drude model with $(\varepsilon_\infty, \omega_p, \gamma) = (3.7, 1.38 \times 10^{16} \text{ Hz}, 2.73 \times 10^{13} \text{ Hz})$ [4], where ε_∞ is the dielectric constant at the infinite frequency, γ and ω_p represent the electron collision frequency and bulk plasma frequency, respectively. The increased loss of silver in the nanostructure [5] is neglected and not involved in our simulations. The Kerr nonlinear material in one arm of the Mach-Zehnder interferometer whose dielectric constant ε_f depends on the intensity of electric field $|E|^2$: $\varepsilon_f = \varepsilon_0 + \chi^{(3)}|E|^2$. The value of linear dielectric constant ε_0 is set as 2.0. The Kerr nonlinear material is assumed to be Ag-BaO, and its third-order nonlinear susceptibility at wavelength of 820 nm is $\chi^{(3)} = 4.8 \times 10^{-10}$ esu [6]. When TM-polarized lights are injected to the MIM structure, the incident light will be coupled into the waveguide and the SPP waves are formed

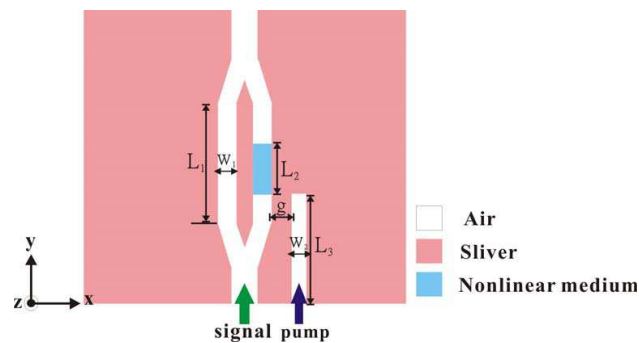


Figure 1: Schematic diagram of the proposed all-optical switch based on the local nonlinear plasmonic MZ interferometer waveguides.

on metal interfaces. Meanwhile, the incident light is coupled into the mach-zehnder interferometer through the coupling aisle. The FDTD method is utilized to calculate the linear and nonlinear responses of our structure [7, 8].

As shown in Fig. 1, the width of the straight pumping waveguide $w_2 = 50$ nm, the length of $L_3 = 400$ nm, and the gap between the straight pumping waveguide and the MZ interferometer $g = 20$ nm. The geometric parameters of Mach-Zehnder interferometer is assumed as $w_1 = 50$ nm, $L_1 = 400$ nm, and $L_2 = 200$ nm. Fig. 2(a) shows the transmission spectra of MIM waveguide with the Mach-Zehnder interferometer. The transmitted dip corresponding to wavelengths of 795 nm is observed. Therefore, the signal light wavelength is set to be 795 nm. As can shown in Fig. 2(b), the transmission efficiency can be effectively dips at a specific frequency point by add nonlinear material in one arm of the MZ interferometer. Fig. 3 shows the transmission responses with pump light off and on. It's shows a significant switching effect. To illustrate more clearly the performance of signal light under conditions of switching off and on, the magnetic field distributions of signal light with the pumping light off and on are plotted in Fig. 4. When the state of the pumping

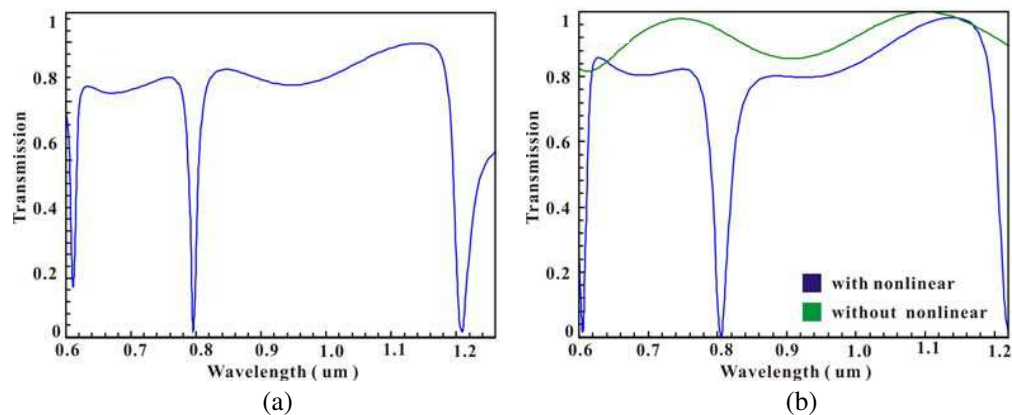


Figure 2: (a) Transmission spectra. (b) Transmission spectra without nonlinear medium (only Mach-Zehnder interferometer) and with nonlinear medium (whole waveguide).

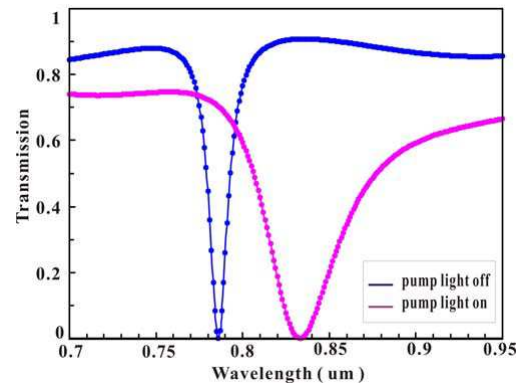


Figure 3: Signal transmission with pump light off and on. The signal intensity is 1.9×10^4 V²/m², pump intensity is 1.9×10^6 V²/m².

light is OFF, the state of the output port is ON and when the state of the pumping light is ON, the state of the output port is OFF. The numerical results show good agreement with the signal transmission response with switching ON and OFF.

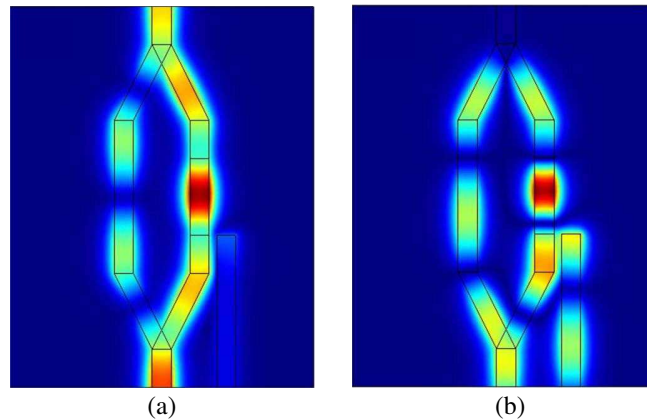


Figure 4: Magnetic field distributions of signal light with switching (a) off and (b) on.

3. CONCLUSIONS

In this paper, we proposed a new all-optical switch based on the local nonlinear metal-insulator-metal plasmonic Mach-Zehnder interferometer waveguides. The transmission characteristics and switching effect in this structure have been investigated numerically by the FDTD simulations. By properly fixing the input signal power and changing the control power, the output signal beam will show the switching effect. In this structure we can find that the proposed all-optical switch exhibits an excellent switching effect. It would be a potential key component in the application of the ultra-high-speed data processing system.

ACKNOWLEDGMENT

This work was partly supported by National Science Council R.O.C. and under Grant No. 101-2221-E-151-077.

REFERENCES

1. Gibbs, H. M., *Optical Bistability: Controlling Light with Light*, Academic, New York, 1985.
2. Moloney, J. V., J. E. Stich, and G. I. Stegeman, "Stability of nonlinear stationary waves guided by a thin film bounded by nonlinear media," *Appl. Phys. Lett.*, Vol. 48, 826–828, 1986.
3. Moloney, J. V., J. E. Stich, and G. I. Stegeman, "Numerical evidence for nonstationary, nonlinear, slab-guided," *Opt. Lett.*, Vol. 11, 315–317, 1986.
4. Tao, J., X. G. Huang, X. S. Lin, J. H. Chen, Q. Zhang, and X. P. Jin, "Systematical research on characteristics of double-sided teeth-shaped nanoplasmonic waveguide filters," *JOSA B*, Vol. 27, No. 2, 323–327, 2010.
5. Drachev, V. P., U. K. Chettiar, A. V. Kildishev, H. K. Yuan, W. Cai, and V. M. Shalaev, "The Ag dielectric function in plasmonic metamaterials," *Opt. Express*, Vol. 16, No. 2, 1186–1195, 2008.
6. Zhang, Q., W. Liu, Z. Xue, J. Wu, S. Wang, D. Wang, and Q. Gong, "Ultrafast optical Kerr effect of Ag-BaO composite thin films," *Appl. Phys. Lett.*, Vol. 82, No. 6, 958–960, 2003.
7. Li, Q., T. Wang, Y. K. Su, M. Yan, and M. Qiu, "Coupled mode theory analysis of mode-splitting in coupled cavity system," *Opt. Express*, Vol. 18, No. 8, 8367–8382, 2010.
8. Taflove, A. and S. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, 2nd Edition, Artech House, Boston, MA, 2000.
9. Lu, H., X. Liu, L. Wang, Y. Gong, and D. Mao, "Ultrafast all-optical switching in nanoplasmonic waveguide with Kerr nonlinear resonator," *Opt. Express*, Vol. 19, No. 4, 2910–2915, 2011.

Magnetolectric Microwave Module for Phased Array

R. V. Petrov, M. I. Bichurin, and A. S. Tatarenko
Novgorod State University, Russia

Abstract— Design of receiving and transmitting microwave modules based on magnetolectric devices is considered. The microwave devices for use in the module are discussed. The main characteristics of devices are given. The calculated characteristics of the module are: a gain of 25 dB, noise figure of 0.9 dB at 10 GHz frequency range. The main advantages of the proposed module is the possibility of precise electronic adjustment in a given frequency range, fine tuning and low phase noise.

1. INTRODUCTION

The article is devoted to research of design the receiving and transmitting microwave modules using devices based on magnetolectric (ME) composites. Progress which has been made recently in the field of ME materials and microwave devices [1–4], allows us to offer a microwave module, which would have certain advantages over traditional modules based on semiconductor and ferrite devices. Advantages of the proposed module are the electronic tuning of any of ME microwave devices, which allow to create the necessary amplitude-frequency response at any time, as well as low level of intrinsic noise, high radiation resistance and low power consumption at the control circuit.

Proposed module can be used for design of receiving and transmitting devices, including phased arrays system [5, 6]. The advantages of our module are: small size and weight, the electronic control method, high radiation resistance, good thermal performance, the possibility of miniaturization and integration on a single substrate, the possibility of creating micro antenna system. All this performances allows the use of the proposed modules to create radar, including the cosmos system, as well as use the device to local computer networks and other communication systems.

2. MICROWAVE MAGNETOELECTRIC DEVICES

ME device such as an antenna, isolator, filter, attenuator, phase shifter, gyrator, amplifier, modulator, detector devices have been created by ME technology over the past twenty years, and the specifications are not inferior to similar competitive models [7–15]. Consider the basic technical and technological possibilities of the proposed devices by the example of ME attenuator.

The main characteristics of studied microwave attenuators [7–9] are: continuously tuned by electric field of attenuation is up to 35 dB at a frequency of 4.478 GHz, VSWR is 1.1, adjusting potential is 400 V. Characteristics of amplitude versus frequency is shown in Fig. 1.

Attenuator control characteristic is shown in Fig. 2.

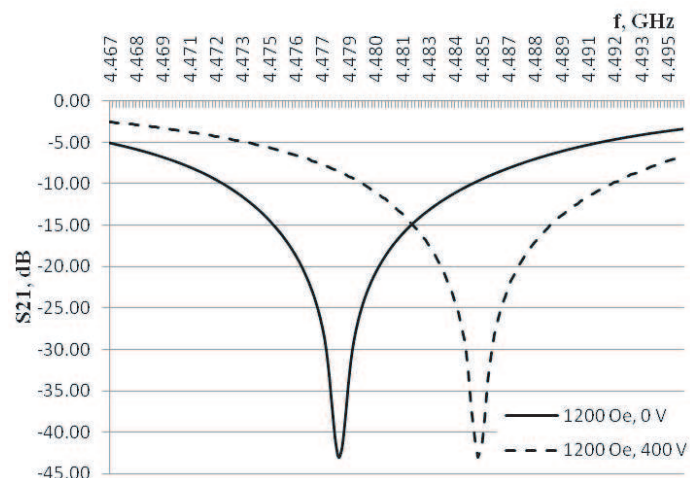


Figure 1: Amplitude vs. frequency of ME microwave attenuator.

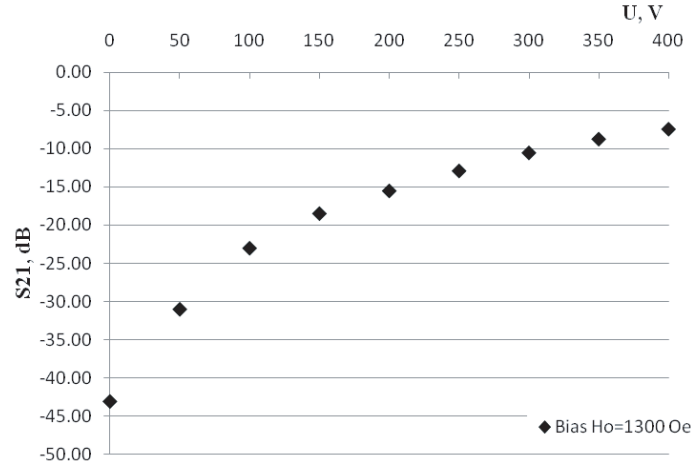


Figure 2: The control characteristic of ME attenuator.

The shift of the FMR line in Fig. 1 may be determined by the following formula

$$\delta f = \frac{\gamma AU}{2\pi d} \quad (1)$$

where γ is the gyromagnetic ratio, A is magnetoelectric coefficient, in our case $A = 3.5 \cdot 10^{-6}$ m/V, U is the control voltage, d is the thickness of the piezoelectric material.

A topology of ME microwave attenuator is shown in Fig. 3. The base of the attenuator is a microstrip transmission line with two stubs on dielectric substrate; ME resonator and two electrodes for applying control voltage. ME resonator can be made from single-phase or layered materials. The layered materials, which are mainly is ferrite YIG film in conjunction with the piezoelectric PZT are the optimal use today. YIG film thickness of 60 microns; GGG substrate thickness of 400 microns and a diameter of 4 mm. The thickness of PZT of 0.5 mm and a diameter of 5 mm. The formula (1) shows that than is thinner the piezoelectric layer, than the bigger the shift of FMR line and therefore we have the more efficient control of device. On the other side, from this formula, it follows that a decrease of the piezoelectric thickness can reduce the control voltage proportionately.

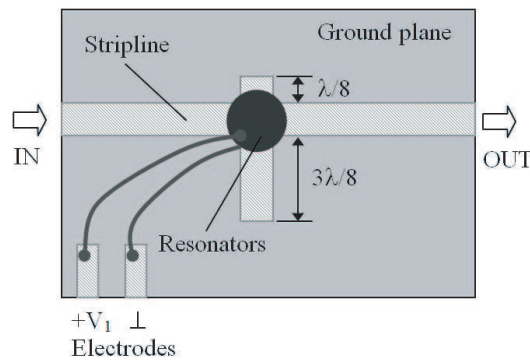


Figure 3: The topology of ME microwave attenuator.

Characteristic of attenuator calculated by the formula for the transmission coefficient by the matrix method as

$$S = e^{-L+i\phi} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \quad (2)$$

where $L = P_{in}/P_{out}$ is the ratio of the input and output power, ϕ is the phase shift.

P_{out} can be defined by the formula

$$P_{out} = P_{in} \cdot e^{-0.5\left(\frac{AU}{d\sigma}\right)^2}, \quad (3)$$

where σ is half-width of the FMR line.

Similarly, can be calculate S -parameters for all devices which used in the module. Then, by the matrix method can be defined a general formula for the transmission coefficient of the module as a whole.

3. MICROWAVE MAGNETOELECTRIC MODULE

Ferrite-piezoelectric structures are ideal for the design of ME modules. ME interaction in these structures allows to implement most of the functions which are necessary for work of receiving and transmitting modules. Hybrid technology allows to combine separately manufactured devices into a single module.

The basic ME elements of the modules for microwave phased arrays (Fig. 4) are devices: antenna, isolator, filter, attenuator, phase shifter, gyrator, modulator, amplifier, detector, a circulator [5, 10–16]. ME effect can be implemented in these devices for a precise setting, or adjustment, etc..

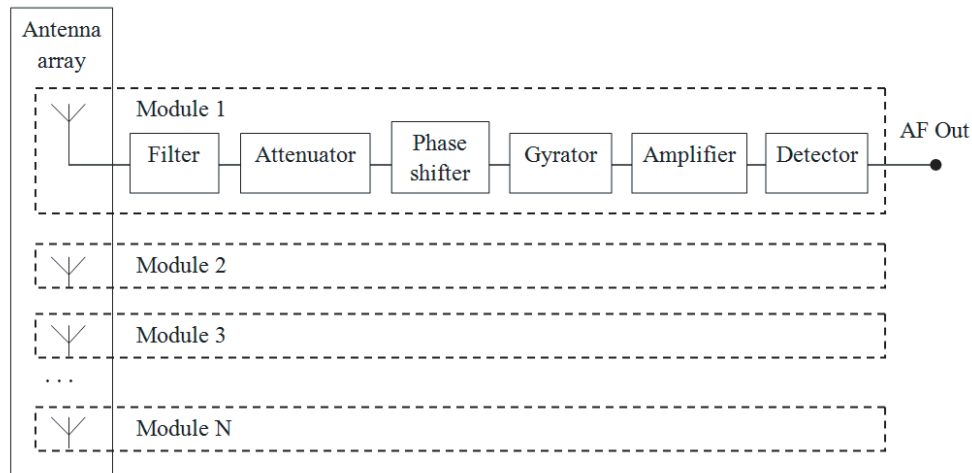


Figure 4: ME receiving module of phased array antenna.

For example, a module developed by Company “Micran” (Russia) based on semiconductor devices has the following features: dynamic range gain is 23 dB, phase adjustment range is 360 degrees with steps of 5.6 degrees and a noise figure of the receiver is not more than 5 dB. We propose the module with comparable characteristics and the following parameters in the range of 10 GHz: a dynamic gain range is 25 dB, phase adjustment range is 360 degrees (continuous), and a noise figure of the receiver is not more than 0.9 dB, which exceeds the characteristics of “Micran” module in several times. The proposed solution is based on the ME devices, which improves sensitivity of the system, and provides advantages for the module as a whole.

The general formula, in which the characteristics of the module is calculated at the center frequency by the matrix method as

$$S = e^{-L+K+i\Phi} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \quad (4)$$

where L is adjustable attenuation, K is adjustable amplification coefficient, Φ is adjustable phase shift.

4. CONCLUSIONS

Advances in technology allow to develop and produce ME devices with competitive performance compared to devices based on other physical principles. ME microwave modules have advantages such as small size and light weight, electronic tuning, continuous phase adjustment, low noise of receiver, high radiation resistance and miniaturization.

Application of ME film structures will significantly reduce the control voltage. Piezoelectric and ferrite materials with better properties than used now, will further significantly improve the parameters of devices, and application of microwave materials, such as, for example, hexaferrites will significantly expand the frequency range of modules up to 100 GHz.

ACKNOWLEDGMENT

The reported study was partially supported by the grant of the Federal Target Program “Scientific and pedagogical staff of innovative Russia” on 2009–2013 and by RFBR research project #13-02-98801.

REFERENCES

1. Bichurin, M. I. and D. Viehland, Eds., *Magnetolectricity in Composites*, 273, Pan Stanford Publishing, Singapore, 2012.
2. Nan, C.-W., M. I. Bichurin, S. Dong, D. Viehland, and G. Srinivasan, “Multiferroic magnetolectric composites: Historical perspective, status, and future directions,” *J. Appl. Phys.*, Vol. 103, 031101, 2008.
3. Bichurin, M. I. and V. M. Petrov, “Magnetic resonance in layered ferrite-ferroelectric structures,” *Sov. Phys. Tech. Phys.*, Vol. 33, 1389, 1989.
4. Bichurin, M. I., R. V. Petrov, and A. S. Tatarenko, “Magnetolectric microwave devices for PAS,” *Modern Problems in Science and Education*, Vol. 4, 2011, URL: www.science-education.ru/98-4727.
5. Bichurin, M. I. and R. V. Petrov, “Magnetolectric phasers for PAS,” *Proceedings of the 2nd International Conference and Exhibition on Satellite Communications (ICSC'96)*, 172–176, Moscow, Russia, 1996.
6. Bichurin, M. I., V. M. Petrov, R. V. Petrov, G. N. Kapralov, Yu. V. Kiliba, F. I. Bukashev, A. Yu. Smirnov, and A. S. Tatarenko, “Magnetolectric microwave devices,” *Ferroelectrics*, Vol. 280, 211, 2002.
7. Bichurin, M. I., R. V. Petrov, and Yu. V. Kiliba, “Magnetolectric microwave phase shifters,” *Ferroelectrics*, Vol. 204, 311, 1997.
8. Tatarenko, A. S. and M. I. Bichurin, “Electrically tunable resonator for microwave applications based on hexaferrite-piezoelectric layered structure,” *American Journal of Condensed Matter Physics*, Vol. 2, No. 5, 135–139, 2012, DOI: 10.5923/j.ajcmp.20120205.05.
9. Bichurin, M. I., R. V. Petrov, and A. V. Filippov, “Magnetolectric microwave gyrator,” Patent RU 2357356 C1 (11.03.2008).
10. Petrov, R. V., M. I. Bichurin, Yu. V. Kiliba, and S. N. Ivanov, “Magnetolectric microwave amplifier,” Patent RU 2439751 C1 (28.07.2010).
11. Bichurin, M. I. and S. N. Ivanov, “Microwave power selective detector,” Patent RU 2451942 C1 (11.01.2011).
12. Tatarenko, A. S., M. I. Bichurin, and D. V. Lavrentieva, “Magnetolectric microwave isolator,” Patent RU 119940 U1 (05.03.2012).
13. Bichurin, M. I., A. S. Tatarenko, G. A. Semenov, D. V. Lavrentieva, and D. A. Semenov, “Magnetolectric attenuator,” Patent RU 115125 U1 (26.10.2011).
14. Petrov, V. M. and K. V. Lavrentyeva, “Adjustable inductive element,” Patent RU 123575 U1 (27.02.2012).
15. Bichurin, M. I., A. S. Tatarenko, and R. V. Petrov, “Active phased-array antenna receiving module,” Patent RU 123586 U1 (03.05.2012).
16. Bichurin, M. I., V. M. Petrov, and G. A. Semenov, “Magnetolectric material for components of radio-electronic devices,” Patent RU 2363074C1 (11.03.2008).

High Tunable and Low Loss Capacitor Using a BZN/BST/BZN Multi-layer Thin Film Dielectric for Reconfigurable RF Circuit Applications

Young Chul Lee

Department of Electronics Engineering, Mokpo National Maritime University (MMU), Korea

Abstract— In this paper, a high-tunability and low-loss capacitor using a multi-layer dielectric of BZN/BST/BZN is designed and fabricated for reconfigurable RF applications. By alternatively depositing a high tunable BST ferroelectric and low-loss BZN paraelectric thin film dielectric, the multi-layer dielectric thin film of BZN/BST/BZN obtained a tunability of 47%, $\tan\delta$ of 0.005, and relative dielectric constant of 225 at 10 MHz. The fabricated tunable capacitor on a quartz wafer using this multi-layer dielectric achieved a Q factor of 10 and a tunability of 63% at 1 GHz and 20 V.

1. INTRODUCTION

In the recent wireless communication market, lately, several networks such as cellular networks (3G and 4G), personal area network (PAN), wireless local area network (WLAN), etc. coexist. These various wireless communication systems are integrated into the single mobile terminal by using switches and some reconfigurable components [1]. In order to intelligently and automatically support multi-band, multi-standard, multi-function, and multi-service communication system, reconfigurable RF components and circuits are key technologies to be developed. Therefore, various tunable components have been developed constantly for several years using tunable dielectrics [2, 3] as well as microelectromechanical systems (MEMS). Ferro- or para-electrics technologies based on barium-strontium titanate (BST), bismuth zinc niobate (BZN) or lead zinc niobate (PZN) has pursued for low loss, low voltage and high operation frequency. In order to increase the tunability and reduce the DC bias voltage, an inter-digital capacitor (IDC) [4] embedding its electrodes into the tunable dielectrics. Ferro-electric thin-films based on BST have shown high-tunability characteristics in spite of relatively lossy feature, compared to the low-loss para-electric ones. In order to compensate or optimize opposite features between ferr- and para-electric dielectrics, a para/ferro/para-electric thin-film multilayer (BZN/BST/BZN) has been proposed for low-loss and high-tunability applications [5].

In this paper, in order to make up for the weak points of the high-loss BST thin film, low-loss and high-tunability MIM capacitor using a multi-layer dielectric thin film inserting a BST thin film between low-loss and -tunability BZN thin films have been fabricated and characterized.

2. FABRICATION OF THE MIM TUNABLE CAPACITORS

A vertical structure of the MIM tunable capacitor using a multi-layer thin-film dielectric is shown in Fig. 1(a). The multi-layer (BZN/BST/BZN) dielectric is utilized for low-loss and high-tunability purpose. The tunable capacitors were fabricated on a quartz substrate in coplanar waveguide (CPW) configuration. The first metal of Ti/Pt ($= 100/1,000\text{\AA}$) was deposited and patterned as a bottom electrode of the MIM tunable capacitor by using lift-off process. For the multi-layer thin-film dielectric, the first BZN pyrochlore thin-film of 700\AA was deposited by RF-magnetron sputtering and a stoichiometric $\text{Bi}_2(\text{Zn}_{1/3}\text{Nb}_{2/3})_2\text{O}_7$ ceramic target was used in an high purity O_2/Ar mixture atmosphere. The detailed process conditions were described in detail at the previous work [3]. By using the Inductive Coupled Plasma (ICP) dry etcher, the BZN film was defined and then the second thin-film (BST) of $1,500\text{\AA}$ was deposited by using a $\text{B}_6\text{S}_4\text{T}$ target and then etched. The final thin-film BZN dielectric of 700\AA was deposited and patterned. The photolithography and etching process were carried out by using the same photo mask and dry etcher, respectively, for the multi-layer dielectric. After patterning of each layer, post-annealing processes were carried out at 600°C for 60 minutes in air to crystallize the film. The lift-off pattern as the second metal (Cr/Au= $100\text{\AA}/1,000\text{\AA}$) was defined on the top of the multi-layer dielectric as the top electrode. Fig. 1(b) shows the fabricated tunable capacitor and its whole size is $327 \times 642 \mu\text{m}^2$. The MIM capacitor is located in the size of $10 \times 20 \mu\text{m}^2$.

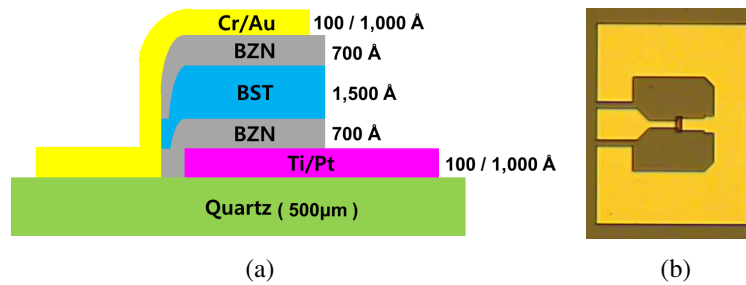


Figure 1: (a) Vertical structure of a MIM tunable capacitor using a multi-layer (BZN/BST/BZN: 700Å/1,500Å/700Å) dielectric. The thickness of a bottom (Ti/Pt) and top (Cr/Au) electrode is 100/1,000Å thick and (b) the fabricated MIM tunable capacitor.

Table 1 presents characteristics of each thin-film and multi-layer dielectric used for fabrication of the MIM tunable capacitors for this work. Their characteristics were measured at 10 MHz and 20V by using a metal-insulator-metal (MIM) (Pt/dielectric/Ag) dot (250 μm diameter) capacitor on a Si/SiO₂(3,000Å)/TiO₂(200Å)/Pt wafer. The multi-layer (BZN/BST/BZN) [5] thin-film dielectric shows a low-loss and high-tunability feature, compared with the lossy BST and low tunable BZN thin-film, respectively.

Table 1: Characteristics of the thin-film dielectric analyzed by using dot capacitors.

Dielectric Thin Films	ϵ_r	$\tan\delta$	Tunability
BST (4,000 Å)	250	0.028	48%
BZN (4,000 Å)	165	0.003	18%
BZN/BST/BZN (1,000Å/2,000 Å/1,000Å)	225	0.005	47%

3. MEASURED RESULTS

By using a vector network analyzer (HP 8510C) and probing method, the fabricated tunable capacitor was characterized on a probe station. The effective capacitance (C_{eff}) and percentage tunability (T) of the fabricated capacitors were analyzed by measuring complex reflection coefficients (S_{11}). Using the measured S_{11} data, the C_{eff} , and tunability are analyzed by using following equations.

$$C_{eff} = -\frac{1}{2\pi \cdot freq. \cdot \text{Im}(Z_{11})} \quad [\text{F}]$$

$$T = \frac{C_{\max} - C_{\min}}{C_{\max}} \quad [\%]$$

Z_{11} is the total impedance of the device under test calculated from the measured S_{11} . C_{\min} and C_{\max} are the measured minimum and maximum capacitance, respectively, within the applied bias voltage range.

Figure 2(a) shows measured S_{11} data on the Smith Chart of the fabricated MIM tunable capacitors. The test frequency and DC bias voltage are from 100 MHz to 3 GHz and 0 V, respectively. The self-resonant frequency (SRF) is over 3 GHz. Fig. 2(b) shows C_{eff} as a function of applied DC bias voltage from DC to 3.0 GHz. As the DC bias voltage increases, C_{eff} reduced. Especially, as the frequency increases, the reduction rate is decreased. That means the tunability is low at high frequency.

Figure 3(a) shows the tunability (T) characteristics with different DC bias voltages of the fabricated tunable capacitor. At 1 GHz, tunability of 45, 58, and 63% is achieved at the applied voltage of 10, 15, and 20 V, respectively. Its Q-factor characteristics with different DC bias voltages are shown in Fig. 3(b). At low frequency around 0.5 GHz, high Q-factor over 20 is obtained. However, over 1 GHz its characteristic is rapidly reduced. That results from a ohmic loss of the thin conductor of 1,000Å.

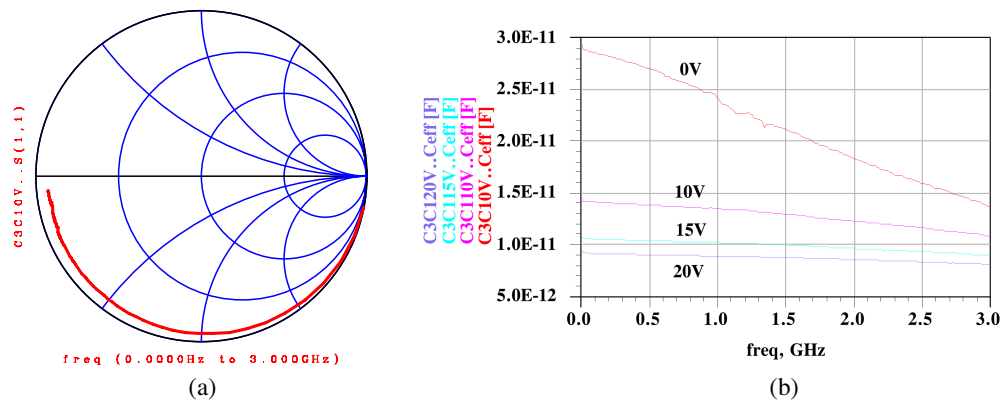


Figure 2: (a) Measured S_{11} data and (b) analyzed C_{eff} characteristics with different DC bias conditions.

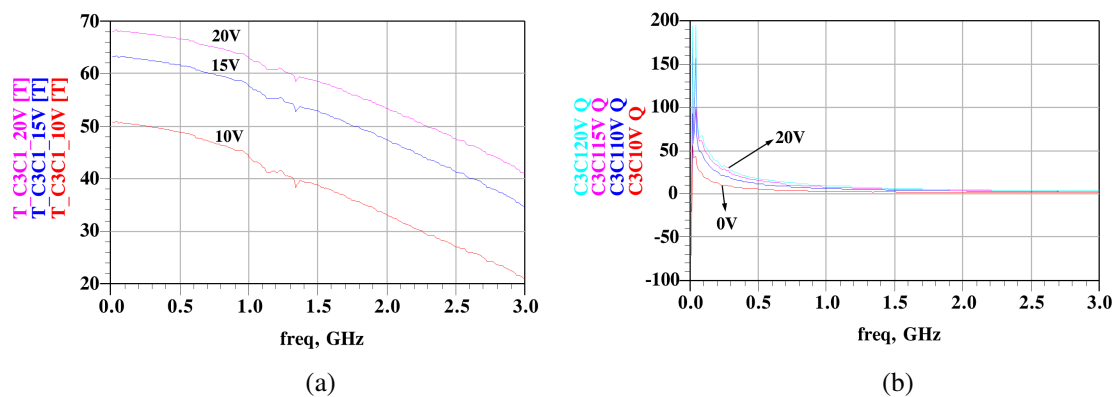


Figure 3: (a) Tunability (T) and (b) Q-factor characteristics with different DC bias voltages of the fabricated tunable capacitor.

4. CONCLUSION

In this work, a high tunable capacitor using a multi-layer dielectric of BZN/BST/BZN is designed and characterized for reconfigurable RF applications. The fabricated tunable capacitor on a quartz wafer using this multi-layer dielectric achieved a Q-factor of 10 and tunability of 60% at 800 MHz and 15 V. Its size is $327 \times 642 \mu\text{m}^2$.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2011-0015031).

REFERENCES

1. Boeck, G., D. Pienkowski, R. Circa, M. Otte, B. Heyne, P. Rykaczewski, R. Wittmann, and R. Kakerow, "RF front-end technology for reconfigurable mobile systems," *IEEE Microwave and Optoelectronics Conference*, Vol. 2, 863–868, 2003.
2. Tuttlebee, W. H. W., "Software-defined radio: Facets of a developing technology," *IEEE Personal Communications Magazine*, Vol. 6, No. 2, 38–44, April 1999.
3. Hong, Y. P., S. Ha, H. Y. Lee, Y. C. Lee, K. H. Ko, D. W. Kim, H. B. Hong, and K. S. Hong, "Voltage tunable dielectric properties of rf sputtered $\text{Bi}_2\text{O}_3\text{-ZnO-Nb}_2\text{O}_5$ pyrochlore thin films," *Thin Solid Films*, Vol. 419, 183–188, 2002.
4. Lee, Y. C., Y. P. Hong, D. M. Kim, and K. H. Ko, "Very high tunable inter-digital capacitor using bismuth zinc niobate thin-film dielectrics for microwave applications," *IEE Electronics Letters*, Vol. 42, 851–853, 2006.
5. Lee, B. J. and K. H. Ko, "Enhancement of BST tunability using Para-Ferro electric multi-layer thin film," *Journal of the European Ceramic Society*, submitted, 2010.

Large-signal Properties of 3C-SiC/Si Heterojunction DDR IMPATT Devices at Terahertz Frequencies

Suranjana Banerjee¹, Aritra Acharyya², Monojit Mitra³, and J. P. Banerjee²

¹Academy of Technology, West Bengal University of Technology
Adisaptagram, Hooghly, West Bengal 712121, India

²Institute of Radio Physics and Electronics, University of Calcutta
92, APC Road, Kolkata, West Bengal 700009, India

³Bengal Engineering and Science University, India

Abstract— The authors have carried out large-signal modeling and simulation of anisotype heterojunction Double-Drift Region (DDR) 3C-SiC/Si Impact Avalanche Transit Time (IMPATT) devices designed to operate at 0.3 and 0.5 THz. A full scale large-signal simulation software based on non-sinusoidal voltage excitation model is used for the present study. Two different heterojunction structures such as *N*-3C-SiC/*p*-Si and *n*-Si/*P*-3C-SiC are considered and their large-signal properties are compared with Si homojunction DDR IMPATTs. The results show that *n*-Si/*P*-3C-SiC heterojunction DDR structure of IMPATTs not only excels its complementary counterpart but also its homojunction counterpart with respect to large-signal conversion efficiency and power output at 0.3 and 0.5 THz frequencies.

1. INTRODUCTION

The terahertz (THz) frequency band popularly called ‘terahertz-gap’ which lies between the millimeter-wave and infrared regions of the electromagnetic spectrum, is defined as the frequency range of 0.1–10 THz, i.e., between the wavelengths 3.0 and 0.03 mm. This particular frequency band is in great demand for various applications such as THz imaging [1], spectroscopy [2], bio-sensing [3], quality inspection in various industrial branches [4–6], medical and pharmaceutical applications [7, 8], THz astronomy [9], etc. The wide bandwidth of THz spectrum can be useful for broadband wireless communication providing data rate of tens of Gbps (gigabytes per second). Notwithstanding the tremendous application potential of THz frequency band, this field has not yet been fully exploited since it is difficult to develop reliable solid-state sources capable of generating, detecting and processing the terahertz signal with sufficient power required for system application. SiC is an excellent material as compared to conventional Si and GaAs for high power and high frequency applications owing to its favorable material parameters like wide bandgap, high breakdown voltage, high carrier saturation velocity and higher thermal conductivity. SiC exists in different polytype forms such as 4H-, 6H-, and 3C-SiC. Among these polytypes, 4H-SiC has already been reported as a base material for IMPATT diodes. The RF power output from 4H-SiC IMPATTs is reported to be very low in contrast to that obtained from simulation studies [10–12]. This is due to the problem of high series resistance associated with undepleted epitaxy and contact layers. It may be noted that 3C-SiC has the lowest bandgap among other poly-type forms. Replacement of 4H-SiC by 3C-SiC as base material for DDR IMPATTs may reduce the parasitic series resistance of the device since both the electron and hole mobility of 3C-SiC are higher than those of 4H-SiC. Thus 3C-SiC IMPATTs are expected to be more suitable for fabrication of high power THz IMPATT devices.

Further, heterojunction IMPATTs have some advantages over their homojunction counterparts as regards lower noise figure and reduced tunnel component of current [13]. The heterojunction between Si and 3C-SiC is technologically more viable and attractive than the heterojunction between Si and other poly-types of SiC due to moderate conduction band discontinuity. This has prompted the authors to study the large-signal properties of anisotype heterojunction DDR Si/3C-SiC IMPATTs designed to operate at 0.3 and 0.5 THz by using a full scale large-signal simulation software based on non-sinusoidal voltage excitation model [14]. Two types of DDR heterojunction structures, i.e., *N*-3C-SiC/*p*-Si and *n*-Si/*P*-3C-SiC are considered for the present study. The results show that the THz performance of *n*-Si/*P*-3C-SiC heterojunction DDR structure not only excels its complementary counterpart but also that based on bulk Si.

2. LARGE-SIGNAL SIMULATION TECHNIQUE

One-dimensional model of reverse biased $n^+ - n - p - p^+$ structure of DDR IMPATT device, shown in Figure 1(a) is used for large-signal simulation since the physical phenomena take place in the semiconductor bulk along the symmetry axis of the mesa structure of the device. The fundamental time and space dependent device equations, i.e., Poisson's equation, continuity equations, current density equations involving mobile space charge are simultaneously solved subject to appropriate boundary conditions [14] under large-signal condition to obtain the snap-shots of electric field ($\xi(x, t)$) and normalized current density ($P(x, t) = (J_p(x, t) - J_n(x, t))/J_t(t)$; where, $J_t(t) = J_p(x, t) + J_n(x, t)$) for different bias current densities (J_0) at several instants of time of one complete cycle of steady-state oscillation. The large-signal simulation is carried out by considering 500 space steps and 150 time steps. In the present simulation method, the IMPATT device is considered as a non-sinusoidal voltage driven source as shown in Figure 1(b). The input AC voltage is taken as

$$V_{RF}(t) = V_B \sum_{p=1}^n m_x^p \sin(p\omega t). \quad (1)$$

The bias voltage is applied across the device through a coupling capacitor (C) to study the performance of the device at a given fundamental frequency ($f = \omega/2\pi$) with its n harmonics.

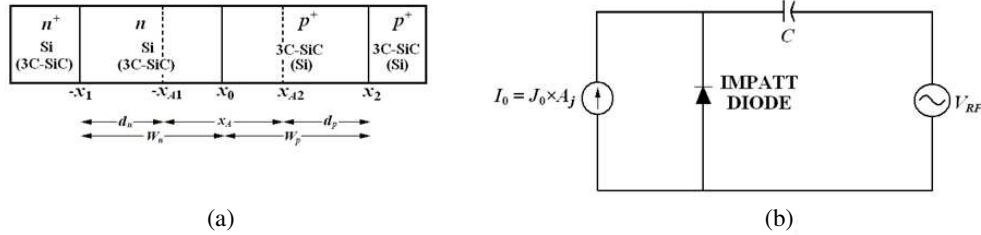


Figure 1: (a) One-dimensional model of DDR IMPATT device, (b) voltage driven IMPATT oscillator and associated circuit.

The large-signal program is run till the limit of one complete cycle (i.e., $0 \leq \omega t \leq 2\pi$) is reached. The bias current density, RF voltage amplitude and frequency are J_0 , V_{RF} and f respectively. The terminal current ($J_t(t)$) and voltage ($V_{RF}(t)$) waveforms for a complete cycle of oscillation are Fourier analyzed to study the RF performance of the device at different phase angles of one complete cycle of oscillation, i.e., $\omega t = 0, \pi/2, \pi, 3\pi/2, 2\pi$.

3. DESIGN OF STRUCTURAL DOPING AND OTHER PARAMETERS

The active layer widths (W_n, W_p) and background doping concentrations (N_D, N_A) of DDR IMPATTs are initially chosen by using the transit time formula of Sze and Ryder [15]. The structural and doping parameters of homojunction DDR Si IMPATTs (Si HMDDR) and heterojunction DDR N -3C-SiC/ p -Si and n -Si/ P -3C-SiC IMPATTs (N -3C-SiC/ p -Si HTDDR and n -Si/ P -3C-SiC HTDDR) are first designed for optimum performance at 0.3 and 0.5 THz design frequencies (f_d) by using the method reported in an earlier paper [14]. In the present simulation, the doping concentrations of n^+ - and p^+ -layers (N_{n+} and N_{p+}) are taken to be $\sim 10^{25} \text{ m}^{-3}$. These parameters are given

Table 1: Structural and doping parameters.

DEVICE	f_d (THz)	W_n (μm)	W_p (μm)	N_D ($\times 10^{23} \text{ m}^{-3}$)	N_A ($\times 10^{23} \text{ m}^{-3}$)	N_{n+} ($\times 10^{25} \text{ m}^{-3}$)	N_{p+} ($\times 10^{25} \text{ m}^{-3}$)	D_j (μm)
Si HMDDR	0.300	0.118	0.110	6.500	7.300	5.000	2.700	13.5
	0.500	0.068	0.066	15.500	16.700	5.000	2.700	10.0
N -3C-SiC/ p -Si HTDDR	0.300	0.220	0.210	3.100	4.600	5.000	2.700	13.5
	0.500	0.140	0.130	5.200	7.400	5.000	2.700	10.0
n -Si/ P -3C-SiC HTDDR	0.300	0.114	0.125	5.000	4.920	5.000	2.700	13.5
	0.500	0.063	0.065	5.400	5.300	5.000	2.700	10.0

in Table 1. A rigorous thermal analysis of the device is carried out in continuous-wave (CW) mode by considering proper heat sink arrangement to avoid thermal runaway phenomenon and burn out problem of the device [16]. The effective junction areas (A_j) and corresponding junction diameters (D_j) of the devices are obtained from the said thermal analysis (Table 1). Material parameters such as field dependence of ionization rates ($\alpha_{n,p}$ vs. ξ) and drift velocities ($v_{n,p}$ vs. ξ) of charge carriers, bandgap (E_g), intrinsic carrier concentration (n_i), effective density of states of conduction and valance bands (N_c, N_v), effective masses (m_n^*, m_p^*), diffusion coefficients (D_n, D_p), mobilities (μ_n, μ_p) and diffusion lengths (L_n, L_p) of charge carriers in Si and 3C-SiC are taken from published experimental reports [17–21].

4. RESULTS AND DISCUSSION

The large-signal simulation program [14] has been used to study the high frequency properties of n -Si/ p -Si (homojunction), N -3C-SiC/ p -Si and n -Si/ P -3C-SiC (heterojunction) DDR IMPATTs designed to operate at 0.3 and 0.5 THz. The static or DC analysis of the devices are then carried out as a limiting case of large-signal analysis where the voltage modulation factor, $m_x = 0$. The various important DC parameters such as peak electric field (ξ_p), breakdown voltage (V_B), avalanche voltage (V_A), ratio of drift layer voltage to breakdown voltage (V_D/V_B ; where $V_D = V_B - V_A$), avalanche layer width (x_A ; where $x_A = |x_{A1}| + x_{A2}$), ratio of avalanche layer width to total drift layer width (x_A/W ; where $W = W_n + W_p$) of the devices under consideration are obtained by taking the time averages of respective time varying parameters. The DC parameters of the devices under consideration are given in Table 2. It is observed from Table 2 that the peak electric field (ξ_p) and breakdown voltage (V_B) of heterojunction devices are higher as compared to those of Si homojunction device at both 0.3 and 0.5 THz. The ratio of V_D/V_B is found to be largest (63.51 and 58.97% at 0.3 and 0.5 THz respectively) in n -Si/ P -3C-SiC heterojunction IMPATTs. Also the avalanche region is narrowest in n -Si/ P -3C-SiC heterojunction IMPATTs at both 0.3 and 0.5 THz frequencies. Largest V_D/V_B (%) and smallest x_A/W (%) in n -Si/ P -3C-SiC heterojunction IMPATT indicates that the DC to RF conversion efficiency of the above mentioned heterojunction device is higher than its complementary heterojunction counterpart as well as its homojunction counterpart at both the design frequencies.

Table 2: Static parameters.

DEVICE	f_d (THz)	J_0 ($\times 10^8 \text{ Am}^{-2}$)	ξ_p ($\times 10^7 \text{ Vm}^{-1}$)	V_B (V)	V_A (V)	V_D/V_B (%)	x_A (μm)	x_A/W (%)
Si HMDDR	0.300	26.500	9.3341	11.52	9.29	19.34	0.138	60.52
	0.500	59.000	11.2603	8.78	7.75	11.68	0.090	66.18
N -3C-SiC/ p -Si HTDDR	0.300	25.590	9.7311	20.99	7.78	62.92	0.116	26.98
	0.500	50.180	12.1445	16.76	6.94	58.61	0.092	34.07
n -Si/ P -3C-SiC HTDDR	0.300	35.300	9.8871	15.84	5.78	63.51	0.061	25.31
	0.500	64.500	12.7040	11.31	4.64	58.97	0.039	30.08

The large-signal admittance characteristics of DDR Si homojunction, N -3C-SiC/ p -Si and n -Si/ P -3C-SiC heterojunction IMPATTs are shown in Figures 2 and 3 for operation at 0.3 and 0.5 THz. The voltage modulation is taken to be 60% in all cases under consideration. This is due to the fact that the optimum performance is obtained at the said voltage modulation for all the devices as shown in Figures 4 and 5. It is observed from Figures 2 and 3 that the magnitude of peak negative conductance ($|G_p|$) of DDR n -Si/ P -3C-SiC heterojunction IMPATTs is much higher than that of Si homojunction DDR IMPATTs as well as that of N -3C-SiC/ p -Si heterojunction IMPATTs operating at 0.3 and 0.5 THz. The largest magnitude of peak negative conductance of DDR n -Si/ P -3C-SiC heterojunction IMPATTs suggests that the device is capable of delivering highest RF power output ($P_{\text{RF}} = (1/2) \times V_{\text{RF}}^2 \times |G_p| \times A_j$) with highest DC to RF conversion efficiency ($\eta_L = P_{\text{RF}}/P_{\text{DC}}$; where $P_{\text{DC}} = V_B \times J_0 \times A_j$) at both the design frequencies.

Figures 4 and 5 show respectively the variations of RF power output (P_{RF}) and large-signal DC to RF conversion efficiency (η_L) with RF voltage for DDR Si homojunction, N -3C-SiC/ p -Si and n -Si/ P -3C-SiC heterojunction IMPATTs at both 0.3 and 0.5 THz. It is observed that the RF power output increases initially with the increase of voltage modulation (m_x), attains a peak value at 60% voltage modulation and then decreases. This holds good for the all homojunction

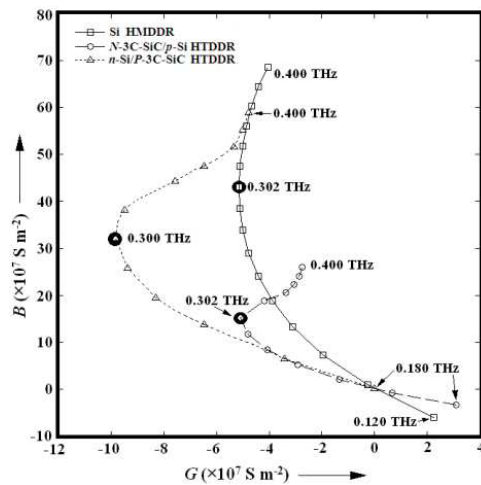


Figure 2: Admittance characteristics of 0.3 THz DDR Si homojunction, N -3C-SiC/ p -Si and n -Si/ P -3C-SiC heterojunction IMPATTs.

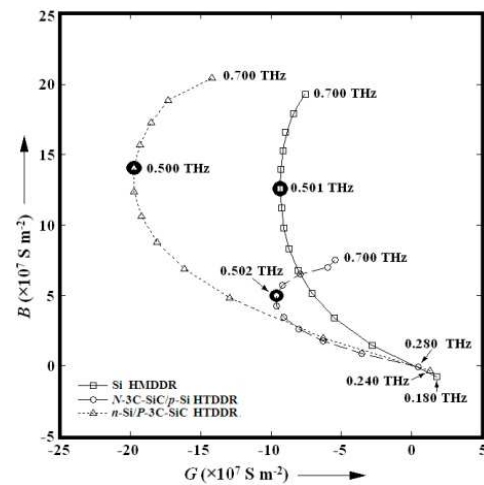


Figure 3: Admittance characteristics of 0.5 THz DDR Si homojunction, N -3C-SiC/ p -Si and n -Si/ P -3C-SiC heterojunction IMPATTs.

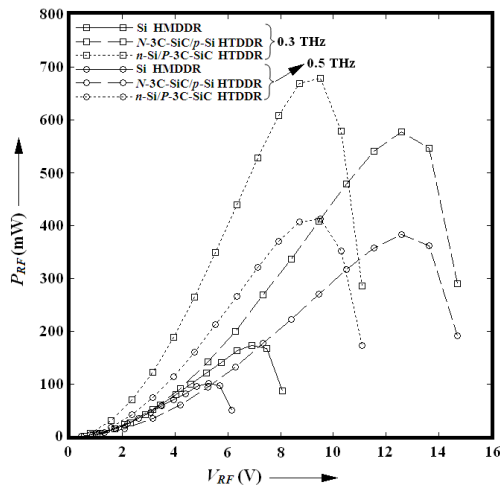


Figure 4: Variations of large-signal RF power output of 0.3, 0.5 THz DDR Si homojunction, N -3C-SiC/ p -Si and n -Si/ P -3C-SiC heterojunction IMPATTs with RF voltage.

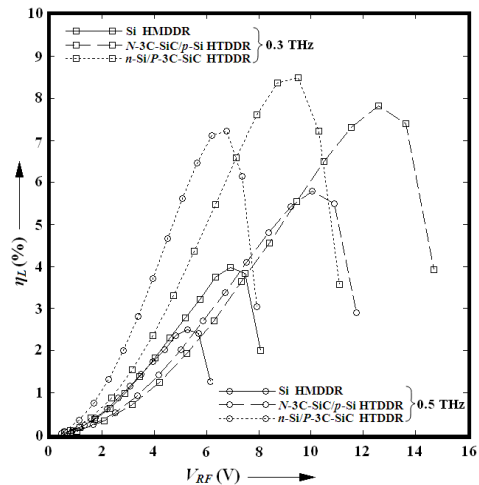


Figure 5: Variations of DC to RF conversion efficiency of 0.3, 0.5 THz DDR Si homojunction, N -3C-SiC/ p -Si and n -Si/ P -3C-SiC heterojunction IMPATTs with RF voltage.

and heterojunction devices under consideration at both 0.3 and 0.5 THz design frequencies. The nature of variation of large-signal efficiency (η_L) with RF voltage is similar to that of P_{RF} for all devices. The results clearly indicate that the optimum performance from the device is obtained when the voltage modulation is around 60%. It may therefore be concluded that the large-signal performance of DDR n -Si/ P -3C-SiC heterojunction IMPATT is the best among all other devices operating at 0.3 and 0.5 THz frequencies.

The important large-signal parameters such as optimum frequency (f_p), avalanche resonance frequency (f_a), peak negative conductance (G_p), corresponding susceptance (B_p), quality factor (Q_p), negative resistance (Z_R), RF power output (P_{RF}) and conversion efficiency (η_L) of DDR n -Si/ p -Si (homojunction), N -3C-SiC/ p -Si and n -Si/ P -3C-SiC (heterojunction) IMPATTs operating at 0.3 and 0.5 THz frequencies are obtained from the simulation for a fixed voltage modulation of 60%. The results are given in Table 3. Avalanche resonance frequency (f_a) of the IMPATT device is the frequency at which the conductance of the device changes its sign from positive to negative. It is observed from Table 3 that f_a of heterojunction DDR IMPATTs (N -3C-SiC/ p -Si and n -Si/ P -3C-SiC HTDDR) is higher than that of their Si homojunction counterpart. Thus the heterojunction devices start to oscillate at higher frequencies as compared to their homojunction counterpart. The Q -factor ($Q_p = -B_p/G_p$) of the device is a measure of the growth rate and

Table 3: Large-signal parameters ($m_x = 60\%$).

DEVICE	f_d (THz)	f_a (THz)	f_p (THz)	G_p ($\times 10^7 \text{ Sm}^{-2}$)	B_p ($\times 10^7 \text{ Sm}^{-2}$)	$Q_p =$ $-(B_p/G_p)$	Z_R/A_j ($\times 10^{-9} \Omega \text{ m}^{-2}$)	P_{RF} (mW)	η_L (%)
Si HMDDR	0.300	0.137	0.302	-5.0867	43.5257	8.56	-0.2649	173.9	3.97
	0.500	0.193	0.501	-9.3138	126.3743	13.57	-0.0580	101.7	2.49
<i>N</i> -3C-SiC/ <i>p</i> -Si	0.300	0.206	0.302	-5.0963	15.4966	3.04	-1.9151	578.7	7.83
	HTDDR	0.500	0.284	0.502	-9.6537	46.5228	4.82	-0.4276	383.4
<i>n</i> -Si/ <i>P</i> -3C-SiC	0.300	0.179	0.300	-10.5181	32.9046	3.13	-0.8814	680.0	8.49
	HTDDR	0.500	0.246	0.500	-22.8757	125.2786	5.48	-0.1413	413.7

stability of IMPATT oscillation. Lower the Q -factor closer to one (i.e., $Q_p \approx 1$) higher is the growth rate of oscillation and better is its oscillation stability. It is observed from Table 3 that the Q -factors of heterojunction DDR IMPATTs are smaller (and closer to 1) than those of homojunction DDR *n*-Si/*p*-Si IMPATTs operating at the same frequencies. This fact suggests that the high frequency performance of heterojunction DDR IMPATTs is better than their homojunction counterparts. Table 3 shows that the magnitudes of negative resistance (Z_R) of both types of heterojunction DDR IMPATTs are significantly higher than that of homojunction DDR Si IMPATTs at both 0.3 and 0.5 THz. All the above results clearly indicate that the DDR *n*-Si/*P*-3C-SiC IMPATT device excels all other DDR structures under consideration as regards power output and efficiency for both 0.3 and 0.5 THz frequencies.

5. CONCLUSION

Large-signal properties of anisotype heterojunction DDR 3C-SiC/Si IMPATT devices designed to operate at 0.3 and 0.5 THz are obtained by using a full scale large-signal simulation software based on non-sinusoidal voltage excitation model. Two different heterojunction structures such as *N*-3C-SiC/*p*-Si and *n*-Si/*P*-3C-SiC are considered for comparison with Si homojunction DDR IMPATTs. The results show that *n*-Si/*P*-3C-SiC heterojunction DDR structure of IMPATTs excels all other structures with respect to DC to RF conversion efficiency and RF power output at both 0.3 and 0.5 THz frequencies. The design parameters and corresponding large-signal properties presented in this paper will be useful for experimental realization of high power and high efficiency from DDR *n*-Si/*P*-3C-SiC heterojunction IMPATTs at both 0.3 and 0.5 THz design frequencies.

REFERENCES

1. Chan, W. L., J. Deibel, and D. M. Mittleman, "Imaging with terahertz radiation," *Rep. Prog. Phys.*, Vol. 70, 1325–1379, 2007.
2. Grischkowsky, D., S. Keiding, M. Exter, and C. Fattinger, "Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors," *J. Opt. Soc. Am. B*, Vol. 7, 2006–2015, 1990.
3. Debus, C. and P. H. Bolivar, "Frequency selective surfaces for high sensitivity terahertz sensing," *Appl. Phys. Lett.*, Vol. 91, 184102-3, 2007.
4. Yasui, T., T. Yasuda, K. Sawanaka, and T. Araki, "Terahertz paintmeter for noncontact monitoring of thickness and drying progress in paint film," *Appl. Opt.*, Vol. 44, 6849–6856, 2005.
5. Stoik, C. D., M. J. Bohn, and J. L. Blackshire, "Nondestructive evaluation of aircraft composites using transmissive terahertz time domain spectroscopy," *Opt. Express*, Vol. 16, 17039–17051, 2008.
6. Jördens, C. and M. Koch, "Detection of foreign bodies in chocolate with pulsed terahertz spectroscopy," *Opt. Eng.*, Vol. 47, 037003, 2008.
7. Fitzgerald, A. J., B. E. Cole, and P. F. Taday, "Nondestructive analysis of tablet coating thicknesses using terahertz pulsed imaging," *J. Pharm. Sci.*, Vol. 94, 177–183, 2005.
8. Siegel, P. H., "Terahertz technology in biology and medicine," *IEEE Trans. Microwave Theory & Tech.*, Vol. 52, 2438–2447, 2004.
9. Siegel, P. H., "THz instruments for space," *IEEE Trans. Antenn. Propag.*, Vol. 55, 2957–2965, 2007.

10. Yuan, L., A. James, J. A. Cooper, M. R. Melloch, and K. J. Webb, "Experimental demonstration of a silicon carbide IMPATT oscillator," *IEEE Electron Device Letters*, Vol. 22, 266–268, 2001.
11. Vassilevski, K. V., A. V. Zorenko, K. Zekentes, K. Tsagaraki, E. Bano, C. Banc, and A. Lebedev, "4H-SiC IMPATT diode fabrication and testing," *Technical Digest of International Conference on SiC and Related Materials*, 713–714, Tsukuba, Japan, 2001.
12. Acharyya, A. and J. P. Banerjee, "Prospects of IMPATT devices based on wide bandgap semiconductors as potential terahertz sources," *Applied Nanoscience*, 1–12, 2012, DOI 10.1007/s13204-012-0172-y.
13. Banerjee, S., A. Acharyya, and J. P. Banerjee, "Noise performance of heterojunction DDR MITATT devices based on $\text{Si}_{1-x}\text{Ge}_x$ at W-band," *Active and Passive Electronic Components*, Vol. 2013, 1–7, 2013.
14. Acharyya, A., S. Banerjee, and J. P. Banerjee, "Large-signal simulation of 94 GHz pulsed DDR silicon IMPATTs including the temperature transient effect," *Radioengineering*, Vol. 21, No. 4, 1218–1225, 2012.
15. Sze, S. M. and R. M. Ryder, "Microwave avalanche diodes," *Proc. of IEEE Special Issue on Microwave Semiconductor Devices*, Vol. 59, 1140–1154, 1971.
16. Acharyya, A., J. Mukherjee, M. Mukherjee, and J. P. Banerjee, "Heat sink design for IMPATT diode sources with different base materials operating at 94 GHz," *Archives of Physics Research*, Vol. 2, 107–126, 2011.
17. Grant, W. N., "Electron and hole ionization rates in epitaxial silicon," *Solid State Electron*, Vol. 16, No. 10, 1189–1203, 1973.
18. Bellotti, E., H. E. Nilsson, K. F. Brennan, and P. P. Ruden, "Ensemble Monte Carlo calculation of hole transport in bulk 3C-SiC," *J. Appl. Phys.*, Vol. 85, No. 6, 3211–3217, 1999.
19. Canali, C., G. Ottaviani, and A. A. Quaranta, "Drift velocity of electrons and holes and associated anisotropic effects in silicon," *J. Phys. Chem. Solids*, Vol. 32, No. 8, 1707, 1971.
20. Mickevicius, R. and J. H. Zhao, "Monte Carlo study of electron transport in SiC," *J. Appl. Phys.*, Vol. 83, No. 6, 3161–3167, 1998.
21. "Electronic archive: New semiconductor materials, characteristics and properties," 2013, <http://www.ioffe.ru/SVA/NSM/Semicond>.

A High Efficiency Broadband Supply Modulator Based on Switch-mode Operations for Polar Transmitters

Fei You, Bohai Zhang, Zhebin Hu, and Songbai He

School of Electronic Engineering, University of Electronic Science and Technology
Chengdu 611731, China

Abstract— In this paper, a high efficiency broadband supply modulator is proposed. This modulator is built based on class-E amplifiers and class-E inverters. The envelope reconstruction is realized by $\Delta\Sigma$ modulations. The feasibility of high efficiency operations is verified by numerical simulations and experiments. The simulation of theoretical WCDMA signals' envelope amplification shows over 70% average efficiency. The simulations of the modulator built by TSMC 130 nm CMOS devices shows 43.3% average efficiency for 40 MHz baseband signals' envelopes. The experimental board for 3.84 MHz WCDMA signals' envelope modulator built by discrete transistors shows 45.6% average efficiency.

1. INTRODUCTION

Because polar transmitters can be reconfigured for different modulation schemes by only alternating the baseband settings [1]. They are very suitable for integrated multi-mode transmitters. However, the design of high efficiency and high linear supply modulators are critical and demanding to the overall performance of polar transmitters.

Pulse modulation based supply modulators, such as $\Delta\Sigma$ [15], are generally utilized in the envelope path of polar transmitters to utilize high efficiency switch-mode modulators. However, high efficiency and linearity of switch-mode supply modulators cannot be constantly kept for wideband signal amplifications caused by pulsewidth and frequency variations [3]. Degraded efficiency of supply modulators becomes the most important factor which should be valued in the design of overall system efficiency of polar transmitters.

Although general switch-mode power supply modulators (SMP) can offer high efficiency performance at a fixed frequency, the average efficiency degraded sharply as the input pulsewidth varies with random baseband modulations. For the sake of wideband high efficiency supply modulators, a variety of hybrid configurations combining linear amplifier and SMP was proposed and studied [4–6]. In the hybrid supply modulator, the SMP provides low frequency heavy current and the linear amplifier complete the rest high frequency current to the load [7, 8]. Through measurements, this configuration was verified to be of efficiency higher than 40% when amplifying envelopes wider than 20 MHz bandwidth [4].

In this paper, a switch-mode only supply modulator based on a conventional class-E amplifier and a zero-voltages-itching (ZVS) rectifier is introduced to expand the operating bandwidth of high efficiency. By adopting thinned out methods [9–11] and a diode, output envelope can be restored with moderated efficiency degradation for the resonate class-E inverter operation at a fixed frequency.

This paper is organized as follows, the configurations and theoretical performance of the proposed supply modulator is introduced in Section 2. In Section 3, its wideband performance is verified by simulations in TSMC 130 nm CMOS process. In Section 4, an experiment is given for discrete transistors realized in printed circuit boards. Section 5 draws the conclusions.

2. BASIC CONFIGURATIONS AND THEORETICAL PERFORMANCE

The basic configuration of the proposed switch-mode supply modulator is shown in Figure 1. In this figure, the shaded block tagged 'class-E amplifier' is consisted of a general class-E amplifier whose component values can be determined according to [12]. The shaded block tagged 'ZVS Rectifier' is consisted of a class-E inverter which can be referenced to [13]. The whole supply modulator can be directly driven by 1-bit digital signals for the switch-mode only configurations. This modulator is suitable for integrated circuits.

The class-E amplifier is operating at a fixed frequency, and the output envelope modulation is realized by the on-off states of transistor T2 controlled by the $\Delta\Sigma$ modulator's outputs from the baseband digital circuits. When T2 is turned off, the circuit operates as a general high efficiency class-E inverter. When T2 is turned on, the diode D_a can prevent the class-E amplifier from sharp

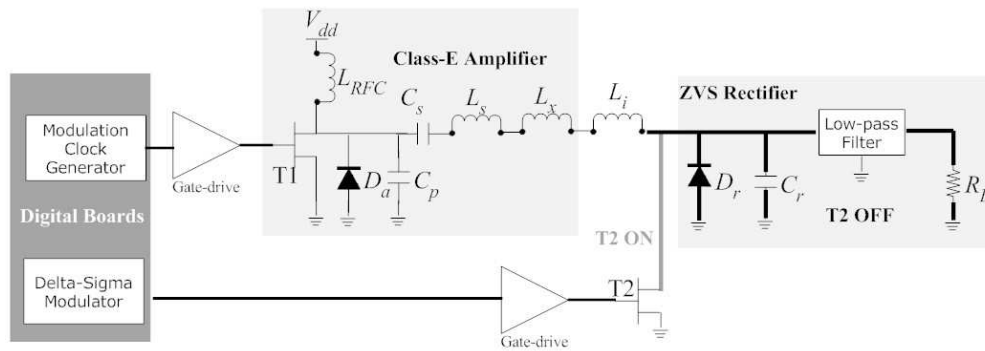


Figure 1: The basic configurations of the proposed supply modulator.

efficiency degradation as analyzed in [14]. So the proposed supply modulator can maintain low power loss for T2's on and off settings in steady states. However, the high efficiency performance of this modulator can be affected by the transient responses during T2's state transitions. This analysis will be included in a further extended publication, which will not be discussed here for simplicity.

The modulated envelope is realized by $\Delta\Sigma$ modulations [15]. This modulations from baseband circuit boards are restored in T2's drain node as shown in Figure 2. If the sampled pulses are filtered with a properly designed low-pass filter, the envelopes can be restored including negligible frequency aliasing due to infinite envelope bandwidth, which can be verified by simulations. However, the amplitude of practical sampled pulses will also be distorted by transient responses during T2's state transitions. This effect must be taken into considerations in linearized amplification. In this paper, only the efficiency performance is discussed considering the distortions can be linearized by digital predistortion methods.

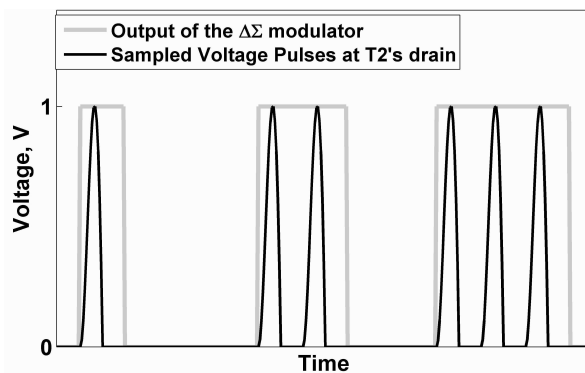


Figure 2: The envelope modulation schemes of the proposed supply modulator.

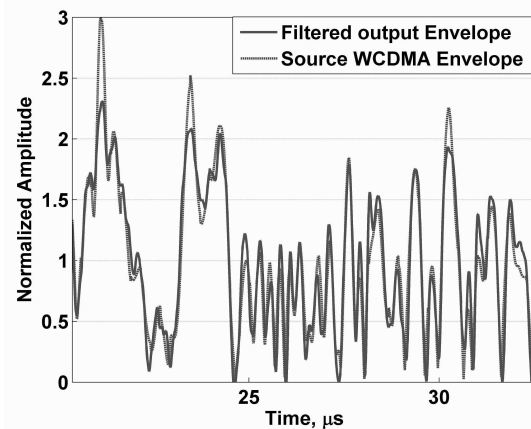


Figure 3: The simulated restored output 3.84 Mbps WCDMA envelopes of the modulator operating at 122.88 MHz.

A numerical simulation example is given as follows. The class-E amplifier of the supply modulator is designed operating at 122.88 MHz. The envelopes of 3.84 Mbps WCDMA baseband signals are used to generate the $\Delta\Sigma$ modulator's outputs. The supply modulator is designed according to the equations of the general class-E amplifiers and rectifiers. The low-pass filter's cut-off frequency is set to be 3.84 MHz which can allow 7.68 MHz envelope signals to pass. The simulated restored output WCDMA envelopes of the modulator is shown in Figure 3. The simulation shows that the proposed supply modulator can recover the input envelopes through digital $\Delta\Sigma$ modulations, except compressions for higher amplitudes. The compressions are caused by the distortions during state transitions. The average efficiency of this modulator is 76% for $L_{dc} = 0.5 \mu\text{H}$. The normalized mean square error (NMSE) between the source signals and the recovered envelopes is about -15.8 dB. Two carrier WCDMA signals' envelopes can also be amplified by the same modulator,

except that the cutoff frequency of the low-pass filter is changed to 6 MHz. The average efficiency is about 74.5%. The NMSE in this case is about -12 dB.

3. SIMULATIONS FOR 40 MHz BASEBAND BANDWIDTH SUPPLY MODULATOR IN TSMC 130 NM CMOS PROCESS

As mentioned above, the proposed modulator is configured for switch-mode operations only. It is suitable for integrated circuits. In order to verify the feasibility of the proposed modulator in integrated circuits, a simulation of the practical circuits built based on TSMC, *Taiwan Semiconductor Manufacturing Co.*, 130 nm CMOS device models are given in this section.

The modulator is designed at 1.28 GHz, and the baseband signal's bandwidth is 40 MHz. The bandwidth of the envelopes is higher than 80 MHz due to bandwidth expansion. The dc supply voltage is 3.3 V. The maximal output power is 0.5 W, and the load resistance is 21.78Ω . The schematic of this modulator is shown in Figure 4. The transistors are selected to be of 3.3 V NMOS with DNW, $0.35 \mu\text{m}$ channel length and $10 \mu\text{m}$ channel width. The parallel transistors are adopted to increase the drain current. The diodes are of 3.3 V native N+/PW. The cascaded diodes are utilized to increase the breakdown voltage. The cutoff frequency of the low-pass filter is 40 MHz. The inductor and the low-pass filter is realized by discrete lumped elements.

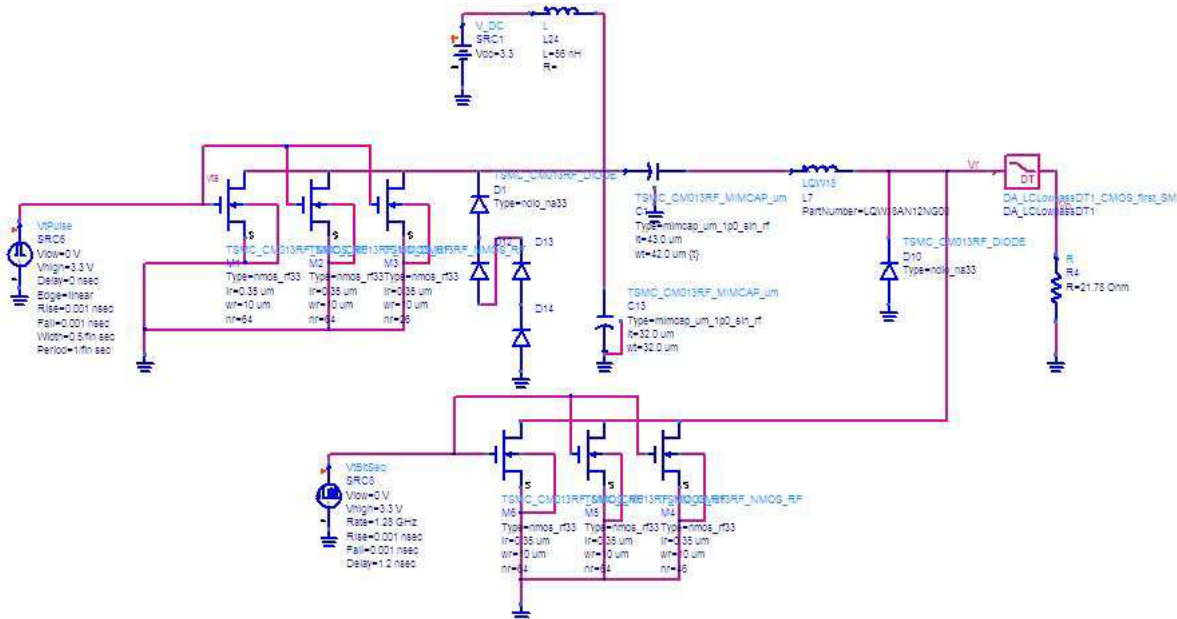


Figure 4: The schematic of the 40 MHz baseband bandwidth signals' envelope supply modulator operating at 1.28 GHz for TSMC 130 nm CMOS devices.

The highest efficiency of this modulator, when T2 is constantly off, is 56%. The output envelopes of this modulator in comparison with the source envelopes, which are normalized to their mean values, are shown in Figure 5. The AM-AM relation between sources and outputs are shown in Figure 6. The 4th degree fitted curve is the fitted AM-AM relation of the expansion one. Because of lower inductance used in this CMOS modulator compared to the one operating at 122.88 MHz, less distortions are observed due to lower transients responses during state transitions. The results show that this proposed configuration is more suitable for high rate envelopes amplification. The average efficiency is measured to be 43.3% for amplifications of 40 MHz baseband signals' envelopes.

4. EXPERIMENTS FOR 3.84 MHz BASEBAND BANDWIDTH SUPPLY MODULATOR IN PRINTED CIRCUIT BOARDS

An experiment for envelope amplifications of WCDMA signal based on $\Delta\Sigma$ modulation is given in this section. The digital board is built based on Cyclone III FPGA device EP3C16Q240C8 obtained from Altera co.. To improve the drive capability for lower speed operations, a gate drive LM5112, obtained from National semiconductor, is inserted between the digital I/O port and the gate of T2.

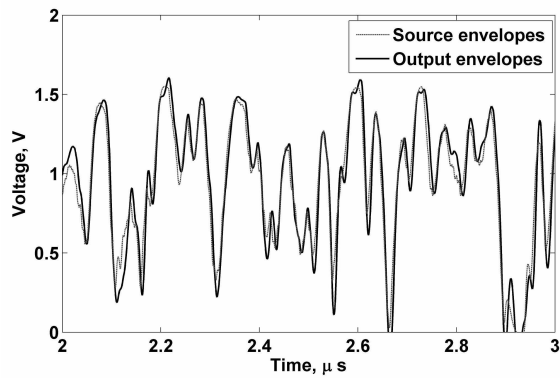


Figure 5: The output envelopes of this modulator in comparison with the source envelopes.

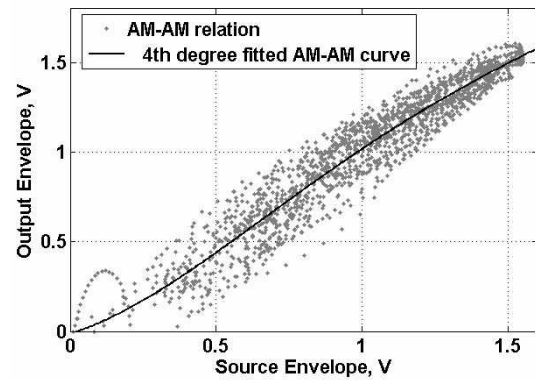


Figure 6: The AM-AM relation between source envelopes and output envelopes.

The schematic of the supply modulator in this experiment is shown in Figure 1. The supply modulator is designed for $V_{dd} = 10$ V and $P_o = 2$ W at 122.88 MHz. LDMOS transistors MRF21010, from Freescale Co., are used to build the supply modulator. Diodes BAT54, obtained from STMicroelectronics Co., are adopted for D_a and D_r in Figure 1. The low-pass filter is tuned for about 3.84 MHz 3 dB cut-off frequency by measurements using practical components. The photo of the FPGA board and the supply modulator is shown in Figure 7.

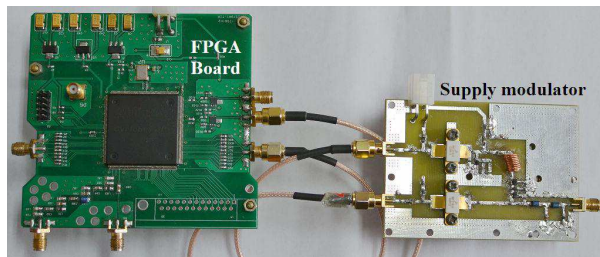


Figure 7: The picture of the experimental supply modulator operating at 122.88 MHz.

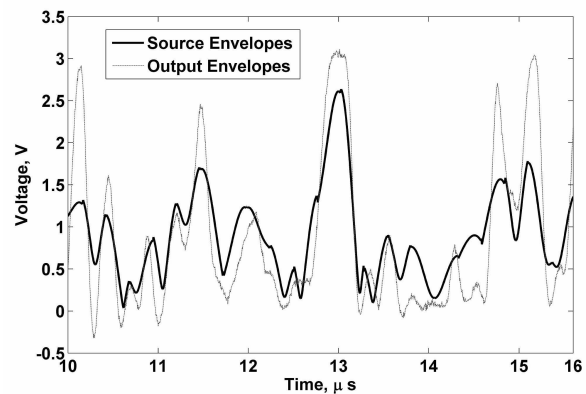


Figure 8: The measured output envelopes in comparison with source envelopes.

The measured efficiency of the supply modulator, when connected transistor T2 is turned off, is 57.8%. The average efficiency is measured to be 45.6% for WCDMA signal's envelope of 5.8 dB PAPR. The NMSE is about -9.5 dB. The measured output envelopes in comparison with source envelopes are shown in Figure 8. Efficiency degradation is mainly caused by power loss in diodes and inductors. Distortions caused by transients are serious due to high dc feed inductance.

5. CONCLUSIONS

The proposed high efficiency broadband supply modulator is verified by simulations and experiments in this paper. The simulations of theoretical supply modulator shows over 70% average efficiency for WCDMA signals' envelopes. The simulation of the modulator built by TSMC 130 nm CMOS devices shows 43.3% average efficiency for 40 MHz baseband signals' envelopes. The experiment for 3.84 MHz WCDMA signals' envelope modulator shows 45.6% average efficiency. Although distortions can be found especially for low frequency operations, this modulator can be linearized by predistortion methods like general transmitters.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under Grant 61001032/F010501.

REFERENCES

1. Yang, H.-S., J.-H. Chen, and Y.-J. E. Chen, “A polar transmitter using interleaving pulse modulation for multimode handsets,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 59, No. 8, 2083–2090, 2011.
2. Lee, S., J. Lee, H. Park, K.-Y. Lee, and S. Nam, “Self-calibrated two-point $\Delta\Sigma$ modulation technique for RF transmitters,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 58, No. 7, 1748–1757, 2010.
3. Kim, D., et al., “Optimization for envelope shaped operation of envelope tracking power amplifier,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 59, No. 7, 1787–1795, 2011.
4. Shrestha, R., R. van der Zee, A. de Graauw, and B. Nauta, “A wideband supply modulator for 20 MHz RF bandwidth polar PAs in 65 nm CMOS,” *IEEE Journal of Solid-State Circuits*, Vol. 44, No. 4, 1272–1280, 2009.
5. Lopez, J., et al., “Design of highly efficient wideband RF polar transmitters using the envelope-tracking technique,” *IEEE Journal of Solid-State Circuits*, Vol. 44, No. 9, 2276–2294, 2009.
6. Kwak, T.-W., M.-C. Lee, and G.-H. Cho, “A 2 W CMOS hybrid switching amplitude modulator for EDGE polar transmitters,” *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 12, 2666–2676, 2007.
7. Minnis, B. J., P. A. Moore, P. N. Whatmough, P. G. Blanken, and M. P. van der Heijden, “System-efficiency analysis of power amplifier supplytracking regimes in mobile transmitters,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 56, No. 1, 268–279, 2009.
8. Yan, L., et al., “Circuits and system design of RF polar transmitters using envelope-tracking and SiGe power amplifiers for mobile WiMAX,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, Vol. 58, No. 5, 893–901, 2011.
9. Fujii, M., T. Suetsugu, H. Koizumi, K. Shinoda, and S. Mori, “Resonant DC/DC converter with class E inverter and class E rectifier using thinnedout method,” *Tenth Annual Applied Power Electronics Conference and Exposition*, 510–515, Dallas, TX, 1995.
10. Fujii, M., T. Suetsugu, K. Shinoda, and S. Mori, “Class-E rectifier using thinned-out method,” *IEEE Transactions on Power Electronics*, Vol. 12, No. 5, 832–836, 1997.
11. Koizumi, H., H. Sekiya, M. Matsuo, S. Mori, and I. Sasae, “Resonant DC/DC converter with class DE inverter and class E rectifier using thinned-out method (deleting some of the pulses to the rectifier),” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 48, No. 1, 123–126, 2001.
12. Albulet, M., *RF Power Amplifiers*, Noble, Atlanta, GA, 2001.
13. Kazimierczuk, M. K., “Analysis of class E zero-voltage-switching rectifier,” *IEEE Transactions on Circuits and Systems*, Vol. 37, No. 6, 747–755, 1990.
14. Kazimierczuk, M. K. and X. T. Bui, “Class E DC/DC converters with an inductive impedance inverter,” *IEEE Transactions on Power Electronics*, Vol. 4, No. 1, 124–135, 1989.
15. Lee, S., J. Lee, H. Park, K.-Y. Lee, and S. Nam, “Self-calibrated two-point $\Delta\Sigma$ modulation technique for RF transmitters,” *IEEE Transactions on Microwave Theory and Techniques*, Vol. 58, No. 7, 1748–1757, 2010.

Separation of Overlapped Compressed Cyclostationary Signals Using Adaptive FRESH Filtering

Said E. El Khamy¹, Amr El-Helw², and Azza Mahdy²

¹Department of Electrical Engineering, Faculty of Engineering
Alexandria University, Alexandria, Egypt

²Department of Electronics and Communication Engineering
Arab Academy for Science & Technology & Maritime Transport, Cairo, Egypt

Abstract— FRESH filters are adaptive filters and their operation depends on the cyclostationarity of digital signals. By knowing the required set of cycle frequencies of the signal of interest, these filters can remove an interfering signal. In this paper, the optimum adaptive FRESH filter is implemented in the wavelet domain to compress the filter's input signal. The purpose of reducing the sequence length is to reduce the processing time taken to separate spectrally overlapped signals. The idea is illustrated by applying multilevel wavelet decomposition to the input signal. These results were compared with the case of implementing the FRESH filter without compressing the input sequence. The results showed an improvement on the processing time but on the expense of increased average error and degradation in the error convergence rate.

1. INTRODUCTION

The co-channel interference problem is a commonly encountered problem in multiuser communication systems. This problem is caused by spectrally overlapped signals one of which is desired signal and the other one is the interference, Conventional filters cannot solve this problem [1], however, by using the cyclostationary properties of modulated signals, the adaptive Frequency Shift filtering techniques (FRESH) filters, proposed by W. A. Gardner, can solve this problem. It only requires cyclostationary information of signals, the cycle frequencies, to extract the desired signal [1,2]. This requires statistical knowledge of SOI training signal to design the Cyclic Wiener Filter. However, these requirements are difficult to be available. Alternatively, Blind Adaptive FRESH filtering is another technique that does not need such training signal. Some researcher had investigated FRESH filters in many applications [5–7,9]. This paper is organized as follows: Section 2 introduces the concept of adaptive FRESH filter and illustrates the main equations that govern the technique as well as Blind Adaptive FRESH filtering (BA-FRESH). In Section 3, the proposed technique is introduced to study the effect of applying FRESH filtering to wavelet transformed signal. Simulation results are illustrated in Section 4. Finally, conclusions are drawn in Section 5.

2. THEORITICAL BACKGROUND

2.1. Optimum Adaptive FRESH Filtering

The Adaptive FRESH filtering is a technique that takes advantage of the spectral correlation of cyclostationary signals for interference rejection as well as reducing the effect of noise. The basic idea of FRESH filters depends on some properties of cyclostationary signals such that for a cyclic signal, frequency shifted versions of it are correlated at various frequencies. When two spectral components are correlated, we say that there exists spectrum redundancy between the two components. If these components are linearly dependent, then we can use one of them to cancel or recover the other. FRESH filters can exploit the cycle frequencies of the interferer signal, the desired or signal of interest (SOI) or both [3]. Good performance is possible if both signals have different baud rates, different carrier frequencies and different modulation schemes.

Assume that the input $x(t)$ of the FRESH filter consists of SOI $s(t)$, the interfering signal $i(t)$, and white noise $v(t)$:

$$x(t) = s(t) + i(t) + v(t) \quad (1)$$

The output of a FRESH filter is given by:

$$y(t) = \int_{-\infty}^{\infty} h(t, \tau) x(\tau) d\tau \quad (2)$$

where $h(t, \tau)$ is the impulse response of the time variant filter which is also called Linear Periodic Time Variant filters (LPTV). These filters are the optimum filters for cyclostationary signals whereas Linear-Time-Invariant filters are the optimum filters for stationary signals. Thus, $h(t, \tau)$ is assumed to be a periodic function. By taking the Fourier Transform (FT) of Equation (2) we get:

$$Y(f) = \sum_{m=1}^M H_m(f) X(f - \alpha_m) \quad (3)$$

where $H_m(f)$ is the FT of the impulse response of the m th FIR filter of the FRESH filter. Equation (3) the input is first frequency shifted then followed by (convolution) that is why it is called frequency-shift filter (FRESH). The design equation for the optimum FRESH filters is given by:

$$S_{xx}^{\alpha_k}(f) \cdot H(f) = S_{sx}^{\alpha_k}(f) \quad (4)$$

$S_{xx}^{\alpha_k}(f)$ is the auto-spectral density function of $x(t)$ and $S_{sx}^{\alpha_k}$ is the cross-correlation spectral density of $s(t)$ and $x(t)$. LMS and RLS algorithms can be used to design the optimum adaptive FRESH filter [8].

2.2. Blind Adaptive FRESH Filtering

In order to blindly adapt the FRESH filter [4], we need a prior knowledge of only its modulation type, carrier frequency, and its baud rate. We will work with input sampled signal such that:

$$x(n) = s(n) + i(n) + v(n) \quad (5)$$

A suitable reference is created by shifting the input by a certain cycle frequency α' which is chosen carefully such that $\alpha' \neq \alpha_m$. This forms a second branch to the optimum FRESH filter. The output of the first branch is fed into a correlator with the output of the second branch. The correlator output gives an indication to the percentage of suppression of the interference or, in other words, how close the output of the filter to the desired signal is. We need to maximize the following cost function:

$$\max J = \max \frac{IR_{yr}I^2}{IR_{yr}IR_{rr}I} \quad (6)$$

where R_{yr} is the cross-correlation between the output of the optimum adaptive FRESH filter and the output of the second branch which is given by:

$$r(n) = x(n) e^{j2\pi\alpha'n} \quad (7)$$

3. COMPRESSED-FEATURES FRESH FILTER

The optimum Adaptive FRESH filter shown in Figure 1 is implemented. The SOI is chosen to be a passband raised-cosine BPSK signal, whereas the interference signal is a passband square-root raised-cosine BPSK signal. The carrier frequency of SOI is $f_c = 50$ Hz, and the interferer carrier frequency $f_i = f_c + \Delta f$ where Δf is the frequency offset and it is set to 16 Hz. The signal-to-noise ratio is set to 3 dB. These two signals are spectrally overlapped and it is required to separate the SOI from the interferer using adaptive FRESH filtering as well as reducing the effect of noise on the extracted desired signal. The job of the adaptive algorithm is to change the filter weights in order to enhance the SOI component and weaken the noise and the interferer components. The step size is set to 0.0001. This is accomplished by using the Least Mean Square algorithm (LMS) algorithm. The error triggers the adaptive algorithm to continue its adaptation process. Once the error reaches a minimum value (in this case almost zero), the adaptation process stops. The output is the summation of the individual outputs from each FIR filter. The output must be a close approximation of the SOI to measure the effectiveness of the filter in the interference removal.

Figure 2 shows the power spectral density (PSD) of SOI as well as that of the interferer signal spectrally overlapped.

In the proposed technique, we suggest compressing the input signal using multilevel wavelet transform. This is shown in Figure 3 where α_{wm} is the cycle frequency parameter of the wavelet level w of the filter branch m . It is required to find relation between the carrier frequency and the periodicities appear in the cyclic spectrum of the wavelet transformed signal. The cyclic spectrum of the wavelet transformed signal is indeed related to that of the time domain signal. However, the

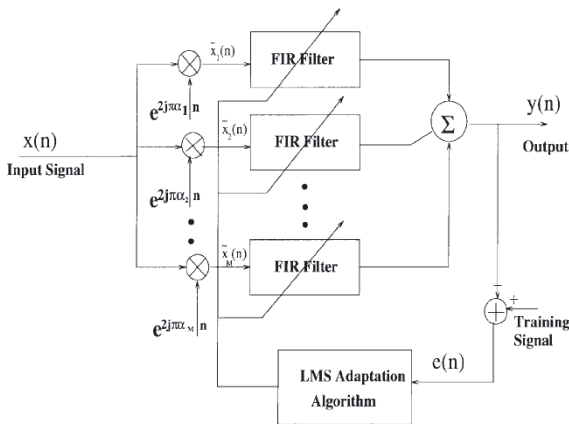


Figure 1: Structure of the optimum adaptive FRESH filter [3].

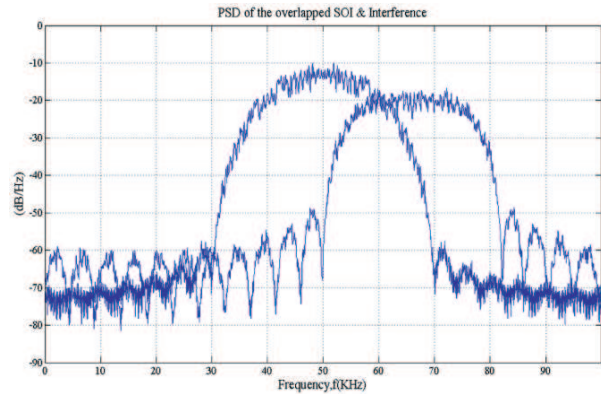


Figure 2: The power spectral density of the spectrally overlapped signals.

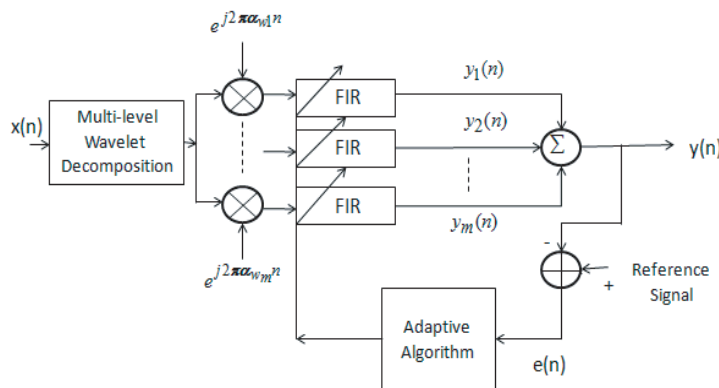


Figure 3: The structure of the Compressed-Features FRESH filter.

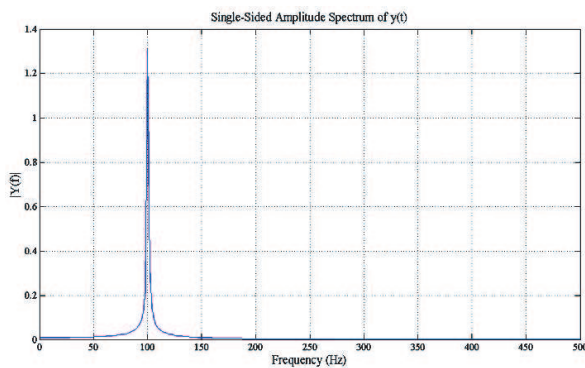


Figure 4: The spectrum of first-level wavelet-transformed signal.

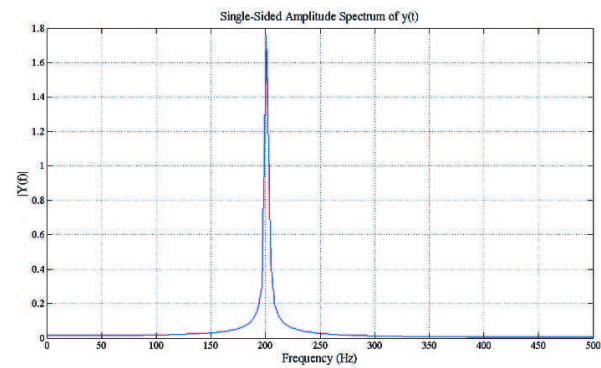


Figure 5: The spectrum of second-level wavelet-transformed signal.

wavelet transform is a multirate transform, so that the signal periodicities change at each level of the wavelet decomposition. Because the signal is decimated by two at each level, its periodicities are multiplied by two. Thus, the periodicities of the coefficients at level 1 are those of the signal but multiplied by two; the periodicities at level two are the signal ones multiplied by four; and so on. Figure 4 and Figure 5 illustrate this idea. In these figures a sinusoidal signal is wavelet-transformed (first level decomposition) and the approximation coefficients are used only. Its spectrum, only one-sided spectrum is shown, is determined. The signal frequency is 50 Hz.

This scaling affects, firstly, the spectrum of the wavelet transformed signal, and therefore the scaling of the cyclic spectrum is a necessary consequence. As a result, a scaled frequency shifting is required at each wavelet level. Figure 6 shows the cyclic spectrum of BPSK signal with a carrier frequency of 50 Hz. The cyclic features occur at double the carrier frequency.

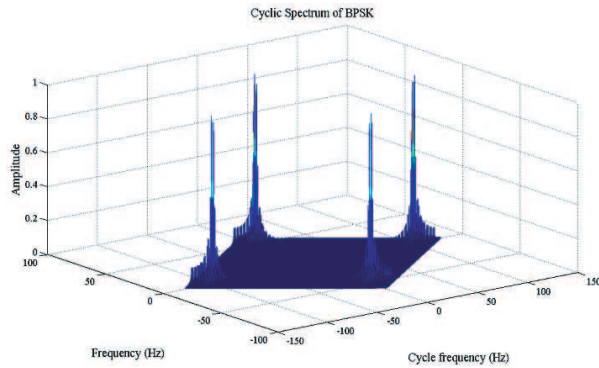


Figure 6: The cyclic spectrum of BPSK signal.

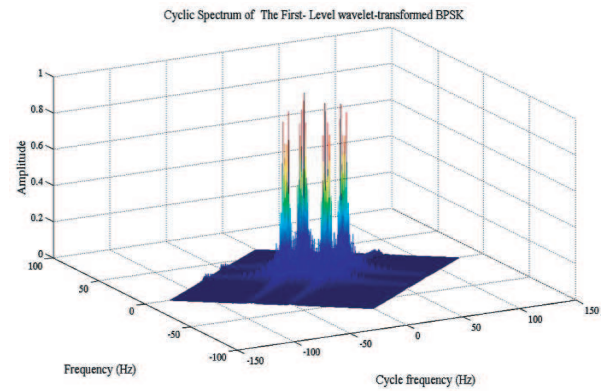


Figure 7: The cyclic spectrum of the first-level wavelet-transformed BPSK.

Figure 7 shows the cyclic spectrum of the approximation coefficients of the first-level wavelet-transformed BPSK signal. The cyclic features occur at half the carrier frequency. This result can be used in the FRESH filter structure by using the frequency shifting parameter $\alpha_w = f_c/2$.

4. SIMULATION RESULTS

The sequence length is chosen to be 50025 samples. The signal duration is 10 sec. The sampling frequency is 1 kHz. The carrier frequency of SOI is $f_c = 50$ Hz, and the interferer carrier frequency $f_i = f_c + \Delta f$ where Δf is the frequency offset and it is set to 16 Hz. The signal-to-noise ratio is set to 3 dB. By applying the first-level wavelet decomposition, the simulation results showed that the two overlapped signals can be successfully separated. However, the more the input signal is wavelet-decomposed, the worse the estimated signal-of-interest (SOI). This is because we chose to consider only the approximation coefficients, some information, despite of its insignificance, is discarded. This is because only the approximation coefficients contain the significant information of the input signal. Figure 8 shows the error convergence without applying wavelet transform to the input signal. The average value of error in this case is 0.0968 and the processing time is 10.9474 sec. Figure 9 illustrates the spectrum of the approximation coefficients of first-level wavelet-transformed SOI. As well as the spectrum estimated approximation coefficients of first-level wavelet-transformed SOI. Figure 10 shows the error convergence rate in this case. We notice that the error convergence rate becomes worse or, in other words, the average value of error becomes higher. However, the processing time taken to remove the interfering signal equals in case of applying first-level wavelet decomposition equals 4.17 sec. This is a reduction percentage of the processing time by 62%. Table 1 illustrates the relation between the processing time and the sequence length as well as the average error value at each wavelet-decomposed level. It is obvious that as the sequence length decreases, the processing time decreases and the average value of the adaptation error increases.

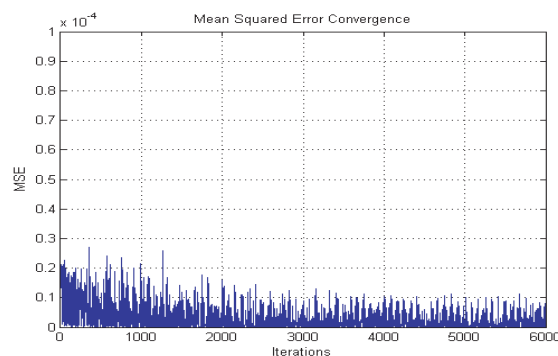


Figure 8: Error convergence rate without applying wavelet transform.

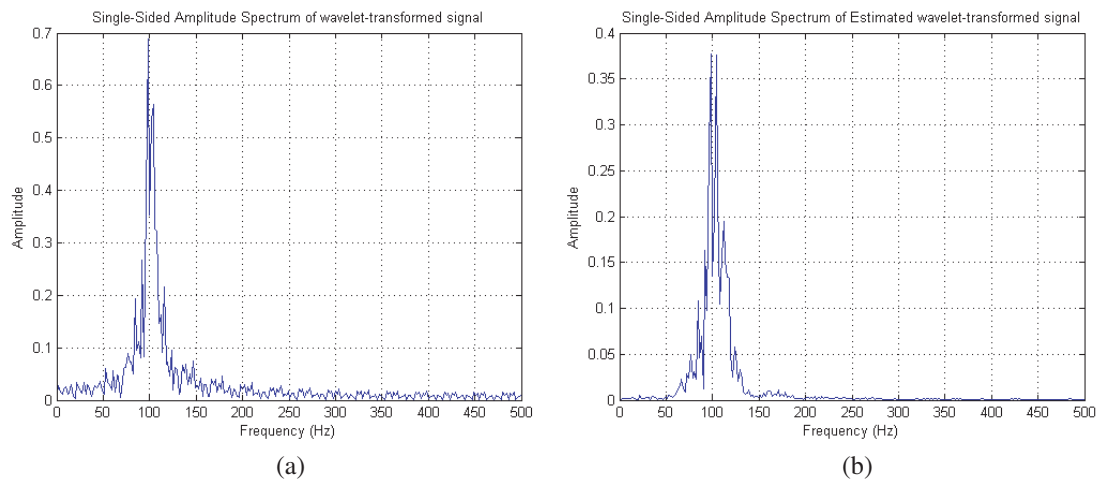


Figure 9: The spectrum of the approximation coefficients of the first-level wavelet transformed signal: (a) Desired, (b) estimated.

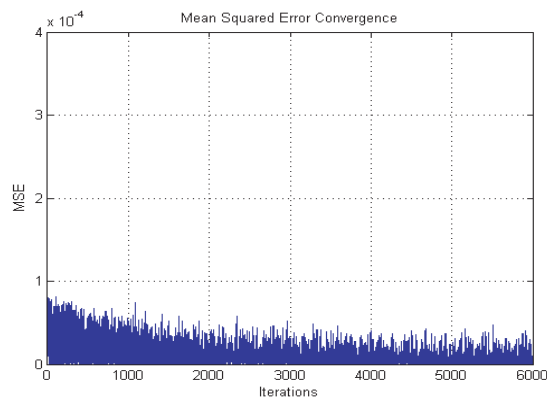


Figure 10: Error convergence rate when applying first-level wavelet transform.

Table 1: Relation between sequence length, processing time, and average value of error.

Wavelet-Decomposition level	Sequence Length	Processing Time (Sec)	Average Error
First level	25013	4.170	0.1886
Second level	12507	2.003	0.3226
Third level	6254	1.983	0.4702

5. CONCLUSION

In this paper, the optimum adaptive FRESH filter is implemented after compressing the input signal to the filter for the SOI being a raised-cosine bandpass BPSK signal and the interference signal being a square-root raised-cosine bandpass BPSK signal. Their carrier frequencies are defined to be overlapped. It is shown that this filter is able to separate spectrally overlapped signals after first level wavelet decomposition of the input signal. Further, the processing time taken to separate the signals after first-level wavelet decomposition equals 4.17sec which is shorter than that in case of implementing the FRESH filter without compressing the input signal which equals 10.9474sec. A reduction percentage of the processing time of 62% after applying first-level wavelet decomposition is achieved. Accordingly, the average value of error increases from 9.8% to 18.8%. Also the more we compress the input signal, the worse the average error.

REFERENCES

1. Gardner, W. A., "Cyclic wiener filtering: Theory and method," *IEEE Transaction on Communications*, Vol. 41, No. 1, 151–163, Jan. 1993.

2. Gardner, W. A., *Introduction to Random Processes with Applications to Signals and Systems*, 2nd Edition, McGraw-Hill, New York, 1989.
3. Zhang, J., K. M. Wong, Q. Jin, and Q. Wu, “A new kind of adaptive frequency shift filter,” *International Conference on Acoustics, Speech, and Signal Processing*, Detroit, USA, May 1995.
4. Zhang, J., K. M. Wong, Z. Q. Luo, and P. C. Ching, “Blind adaptive FRESH filtering for signal extraction,” *IEEE Transactions on Signal Processing*, Vol. 47, No. 5, 1397–1402, May 1999.
5. Guo, J., Y. Liu, Z. H. Song, T. C. Ye, and L. F. Shen, “Separating cyclostationary signals from spectrally overlapping interference,” *Frontiers of Electrical and Electronic Engineering*, Vol. 1, No. 3, 307–312, 2006.
6. Yeste, O. A. and J. Grajal, “Adaptive-FRESH filters for compensation of cycle-frequency errors,” *IEEE Transaction on Signal Processing*, Vol. 58, No. 1, 1–5, Jan. 2010.
7. Whitehead, J. B. and F. Takawira, “Low complexity constant modulus based cyclic blind adaptive multiuser detection,” *7th AFRICON Conference*, Bostwana, Sept. 2004.
8. Haykin, S., *Adaptive Filter Theory*, Prentice Hall, 2001.
9. Yeste, O. A. and J. Grajal, “Cyclostationarity-based signal separation in interceptors based on a single sensor,” *8th IEEE Radar Coneference*, Rome, May 2008.

Depth-image Coding Using Neural Networks for 3D Video Transmission

Yih-Chuan Lin and Pu-Jian Hsu

Department of Computer Science and Information Engineering
National Formosa University, Yunlin 63201, Taiwan

Abstract— In this paper, neural networks are used to classify the content of depth images of 3D videos into three kinds of content classes with distinct characteristics in order to reduce coded data-rate. One type of neural network is first employed to identify the boundary information across the foreground and background objects based on the information from the video texture frame and the associated depth image. Another type of hybrid neural network is trained to fit the surface of depth values over the foreground objects. The similarity between successive depth-image frames is captured by using one other type of neural network to reduce the amount of data required to deliver them. With experimental results, the transmission bit-rate and frame reconstruction quality are evaluated by comparisons to that of using H.264/AVC codec on the same depth image sequence.

1. INTRODUCTION

With the advances of multimedia technology, a variety of 3D video applications have been growingly penetrated in our daily live. From the human visual perception, 3D videos are represented in a way such that audiences are able to receive at least two pictures that are snapshot from different angles at the same time in order for generating the visual perception of depth dimension information from the scene [1]. Owing to its high transmission bit-rate, delivering the 3D video sequences across wired or wireless networks is a challenging task for 3D video communications. Depth image based rendering technique [2] was developed to construct the stereoscopic images at the receiver side based on one representative image (or called texture image) and the associated depth image which represents the depth or distance to the objects of scene from the camera device. While 2D-to-3D conversion techniques [3] were studied to automatically generate the depth image from the 2D frames in a video, the true depth information on the generated depth image is not guaranteed. Therefore, to reduce the data-rate of transmitting 3D videos, the texture and the associated depth information are encoded separately, in which H.264/AVC standard [4] is used for encoding the texture part of 3D video, and a special designed coder is applied to the encoding for depth-image.

Neural networks (NNs) are studied to take the advantage of higher correlation between the texture and depth information for increasing the compression ratio. Neural network is a kind of artificial intelligence which uses machines or software to learn the relationship between the input and output of a dynamic system. With their good ability in dealing with the incomplete data or non-linear cases, neural networks have been popularly used in image processing and compression tasks [5–9].

In this paper, we propose a compression scheme based on feedforward multi-layer neural network with backwardpropagation learning method on the depth-image sequence in a 3D video. In order to release the difficulties of modeling the whole depth image by a single neural network, we try to divide the depth-image content into different parts and encode them in a sequence of steps by distinct types of neural networks. Figure 1 shows the whole 3D video coding and transmission scheme. The incoming 3D video, which includes one video-texture sequence and the associated depth-image sequence, is separated into a series of group-of-picture (GOP), each consisting of one I-frame and several P-frames. Based on the structure, NNs are employed to identify the foreground objects in the depth image associated to the I-frame of each GOP, and to model the depth values over the foreground object, respectively. Beside the bitstream of H.264/AVC for the video texture information, all of the weight values of NNs for depth information of each GOP are transmitted to the receiver for rendering stereoscopic-views of 3D video.

The remainder of this paper is organized as follows. Description of how the neural networks are applied to the coding of depth-image is presented in Section 2. Section 3 demonstrates the results of using the proposed scheme on several test 3D videos. Finally, the paper concludes with some remarks and future work in Section 4.

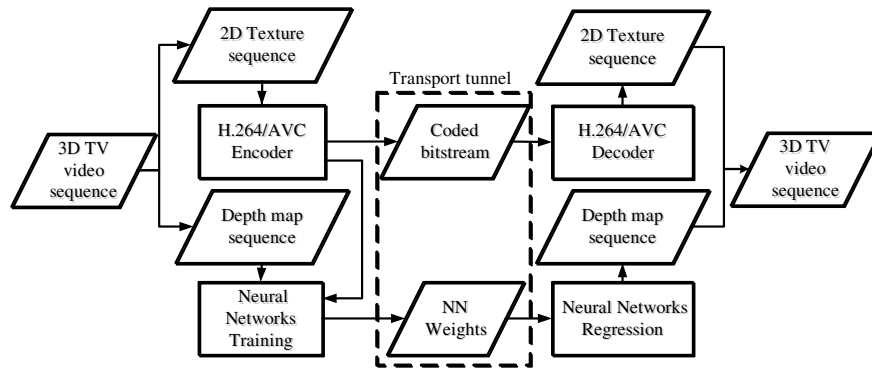


Figure 1: 3D video coding and transmission system.

2. NEURAL NETWORKS FOR DEPTH-IMAGE CODING

3D videos that are considered in the paper are viewed as a set of including one 2D color texture sequence and the associated depth-image sequence. In most cases, the data distribution of pixel values in each depth image is found being irregular. Therefore, to learn the nonlinear and somewhat complex relationship existed among the data in a 3D video, neural networks are employed here to deal with depth-image coding task. Figure 2 depicts an example of the feed-forward multi-layer neural network model with P -dimensional input vector \mathbf{X} and one output variable y . Assume the pixel value at position (m, n) in a depth image that is to be encoded be denoted by $d(m, n)$. With an appropriately chosen input vector \mathbf{X} for $d(m, n)$, the output y of the network is set to be the estimation of $d(m, n)$, which is calculated by

$$\bar{d}(m, n) = y = \sum_{j=1}^Q v_j \sum_{i=1}^P x_i w_{ji} \quad (1)$$

where $\bar{d}(m, n)$ represents the estimation of $d(m, n)$, and Q is the number of neurons in the hidden layer of the network. For ease of presentation, the nonlinear sigmoid function involved in calculating the output of the network is omitted in (1). The estimation error caused by (1) is evaluated as follows,

$$E = \sum_{m=1}^H \sum_{n=1}^W (d(m, n) - \bar{d}(m, n))^2 \quad (2)$$

where H and W are used to stand for the image height and width, respectively. To minimize the estimation error, the back-propagation training algorithm is employed to update the weight values of (1) with a set of training vectors \mathbf{X} and the associated depth value $d(m, n)$ as the target value for supervising the training process.

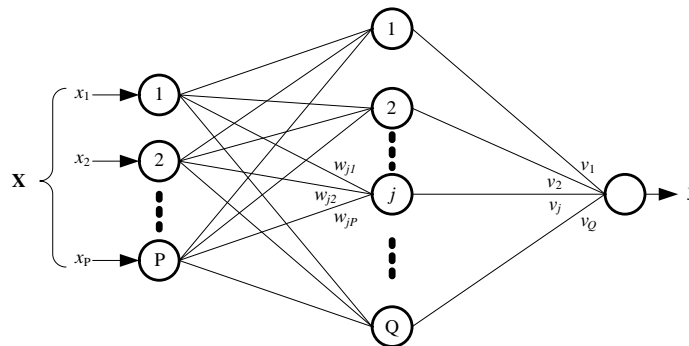


Figure 2: Three-layer feed-forward neural networks.

In order to simplify the sophistication of using a single network to accurately model the whole depth image, the encoding process is separated into four steps, as that shown in Figure 3, each of which is designed to encode the different part of the depth-image sequence.

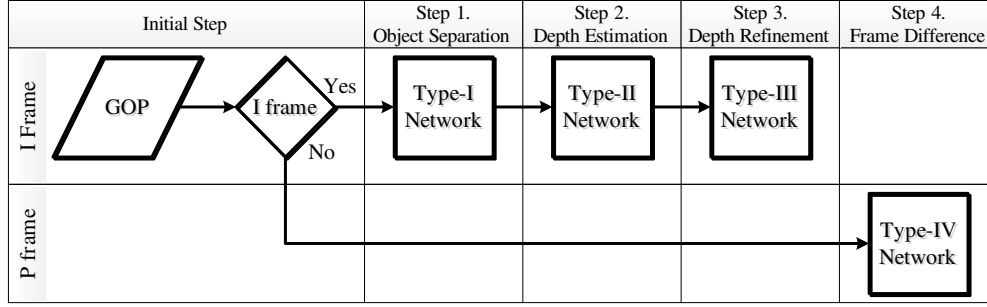


Figure 3: Four-step encoding process for depth-image sequence.

For each I-frame in the depth image sequence, three steps are performed to accomplish its encoding process. In Step 1, the I-frame first goes through a thresholding stage to distinguish the background (far depth) and foreground (near depth) parts. The processed result is denoted as

$$b(m, n) = \begin{cases} 1 & d(m, n) > t \\ 0 & \text{otherwise} \end{cases}, \quad t \geq 0. \quad (3)$$

As $b(m, n)$ is needed to be known by the receiver, a neural network (Type-I network) with a binary output is trained to represent $b(m, n)$ using a training set containing the pairs $(\mathbf{X}, b(m, n))$, and sends the weights of the trained network to the receiver. The elements in the input vector for each target output $b(m, n)$ is designated as follows,

$$\mathbf{X} = (m, n, g(m, n)), \text{ where } g(m, n) = \sqrt{(l(m, n) - l(m-1, n))^2 + (l(m, n) - l(m, n-1))^2} \quad (4)$$

In (4), $l(m, n)$ stands for the luminance component of the corresponding video texture part in the incoming 3D video. In Step 2, those depth values $d(m, n)$ that are classified as the foreground part ($b(m, n) = 1$) are used to train another neural network (Type-II network) and the converged network weights are sent to represent these foreground depth values. The training set is as follows,

$$\{(\mathbf{X}, d(m, n)) | b(m, n) == 1 \text{ and } \mathbf{X} = (m, n, I_R(m, n), I_G(m, n), I_B(m, n))\} \quad (5)$$

$I_R(m, n)$, $I_G(m, n)$, and $I_B(m, n)$ are the three primary color components of the texture pixels in the video. In Step 3, T successively cascaded neural networks (Type-III network) are trained to refine the accuracy of the output depth value from Step 2. The weights of these T networks are also needed by the receiver. The i -th network of these T neural networks is used to generate the i -th refinement $d^{(i)}(m, n)$ of code depth value by employing the approximate error from the preceding networks $e^{(i)}(m, n)$, which is computed by

$$e^{(i)} = \begin{cases} \bar{d}(m, n) & i = 1 \\ d(m, n) - d^{(i-1)}(m, n) & i > 1 \end{cases} \quad (6)$$

Thus, the input vector \mathbf{X} to the i -th Type-III network consists of the following elements,

$$\mathbf{X} = (m, n, e^{(i)}(m, n)) \quad (7)$$

And the target value associated to the input vector \mathbf{X} is the original depth value $d(m, n)$.

As for the coding of the k -th P-frame in a GOP of depth-image sequences, the frame difference $fd(m, n)$ between the current P-frame and the starting I-frame of the GOP is used as the target output value for training the Type-IV network with the input vectors \mathbf{X} , each of which consists of the following elements,

$$\mathbf{X} = (m, n, k, g(m, n), MV(m, n)) \quad (8)$$

where $MV(mn)$ is the motion vector of the texture macroblock that covers the position (m, n) of pixel at the corresponding depth image.

3. EXPERIMENT RESULTS

To verify the effectiveness of the proposed coding scheme, experiments have conducted on a set of 3D videos and comparisons with H.264/AVC codec are made in terms of rate-distortion performance. The 3D video sequences which are considered in the experiments include “Interview”, “Cg”, “Orbi”, “Puppy”, and “Soccer”, each of which contains 125 color texture frames in YUV 4 : 2 : 0 format and 125 frames of depth information. Frames size for the first three videos is of 720×576 pixels and 720×480 pixels for the rest.

As an example, Figures 4(a) and 4(b) show the fifteenth color texture and depth image of “interview” video, which are designated as the I-frames in the experiment, respectively Figure 4(c) demonstrates the reconstructed depth image of Figure 4(b) by the proposed scheme. When considering the loss of NN weights during transmission, the corresponding NN weights of previous frame are used to replace the lost weights, and Figure 4(d) demonstrates the reconstructed image with loss weight correction method.

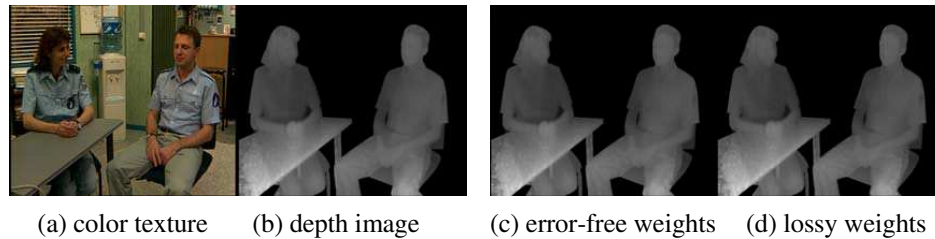


Figure 4: Results of “interview” 3D video sequence.

Figure 5 shows the comparisons of rate-distortion performance between our proposed method and H.264/AVC for each of the five test 3D videos. In the figure, those red lines indicate the results of our proposed method and the blue lines stand for that of H.264/AVC. As compared to the H.264/AVC, the proposed algorithm performs better compression ratios in terms of the amount of coded bits at the same PSNR (peak-signal-to-noise ratio) value.

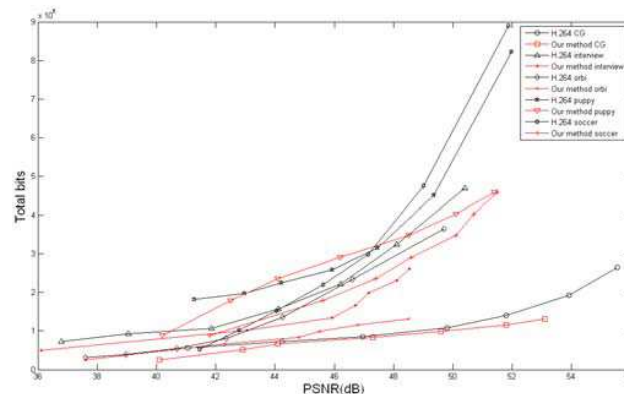


Figure 5: Rate-distortion performance comparison with H.264/AVC.

4. CONCLUSION

In this paper, we have proposed a novel compression method based on multi-layer feed-forward neural networks for the depth-image sequence in 3D videos. The proposed compression scheme is able to exploit the correlation between texture frame and the associated depth image to gain the compression ratio. In order to have accurate reconstructed quality, the foreground part of the depth image is segmented to be represented by a neural network and a cascaded network networks are utilized to refine approximation accuracy. In the experiment, the effect of weight loss when transmitting the coded bits was also simulated and can be concealed or corrected by a replication of weight values from the previous frame. In the future, other feasible neural networks, such as generalized regression neural network (GRNN) and support vector machine approach (SVM) are considered for gaining higher rate-distortion performance.

ACKNOWLEDGMENT

This research is supported in part by the National Science Council of Taiwan under the contract NSC 101-2221-E-150-082.

REFERENCES

1. Redert, A., M. O. de Beeck, C. Fehn, W. Ijsselsteijn, M. Pollefeys, L. Van Gool, E. Ofek, I. SEXTON, and P. Surman, "ATTEST — Advanced three-dimensional television system technologies," *IEEE First International Symposium on 3D Data Processing Visualization and Transmission*, 313–319, Jun. 2002.
2. Fehn, C., "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," *Proceedings of the SPIE of Stereoscopic Displays and Virtual Reality Systems XI*, Vol. 5291, 93–104, Jan. 2004.
3. Harman, P., J. Flack, S. Fox, and M. Dowley, "Rapid 2D to 3D conversion," *Proceedings of SPIE*, 78, 2002.
4. "Advanced video coding for generic audiovisual services, ITU-T Rec. H.264, or ISO/IEC 14496-10 information technology — Coding of audio-visual objects — Part 10: Advanced video coding (MPEG-4 AVC)," 5th Edition, May 2009.
5. Durai, S. A. and E. Anna Saro, "Image compression with back-propagation neural network using cumulative distribution function," *World Academy of Science, Engineering and Technology*, Vol. 17, 60–64, 2006.
6. Zhang, Y. and M. M. Zhang, "Application of artificial neural network in video compression coding," *2008 IEEE ICIII*, 207–210, 2008.
7. Ghorbel, S., M. B. Jemaa, and M. Chtourou, "Object-based Video compression using neural networks," *IJCSI*, Vol. 8, Issue 4, No. 1, 2011.
8. Grossberg, S., L. Kuhlmann, and E. Mingolla, "A neural model of 3D shape-from-texture: Multiple-scale filtering, boundary grouping, and surface filling-in," *Vision Research*, Vol. 47, No. 5, 634–672, 2007.
9. Fanany, M. I. and I. Kumazawa, "A neural network for recovering 3D shape from erroneous and few depth maps of shaded images," *Pattern Recognition Letters*, Vol. 25, No. 4, 377–389, 2004.

Characterization of Dissipation Due to Quasiparticles and Dielectric Loss in a Josephson Metamaterial

J. Hassel¹, P. Lähteenmäki², A. Timofeev¹, G. S. Paraoanu², and P. J. Hakonen²

¹VTT Technical Research Centre, Finland

²O. V. Lounasmaa Laboratory, Aalto University, Finland

Abstract— Josephson junctions can be arranged in a metamaterial configuration, forming an artificial one-dimensional medium for microwave propagation, which enables fast tuning of the speed of light by an external magnetic field. This metamaterial has been used for parametric amplification and as a demonstrator of the dynamical Casimir effect. An important property of metamaterial-based devices is damping, which is fundamentally limited by mechanisms related to the Josephson junctions themselves. In this paper we derive theoretical models for two such mechanisms, the quasiparticle tunneling and the dielectric loss within the material of the tunnel barrier. We discuss whether the junction-based effects are likely to be the limiting factors in the experimentally observed quality factors of metamaterial-based resonators.

1. INTRODUCTION

The speed of light in a Josephson metamaterial is tuned by the magnetic flux going through the loops of superconducting interference devices (SQUIDs). The SQUIDs are distributed circuit elements with dimensions much smaller than the wavelength. Most typically, metamaterial-based circuits are used as parametric amplifiers at microwave frequencies, with the ability to approach the standard quantum limit when operated as phase-preserving amplifiers or to squeeze below the vacuum level when operated as phase-sensitive amplifiers [1]. In the recent demonstration of the dynamical Casimir effect in a Josephson metamaterial, vacuum fluctuation became observable by rapidly modifying the index of refraction, or, equivalently, the speed of light [2]. An alternative realization of this effect uses a single SQUID to emulate a time-varying boundary condition [3].

In general, the potential sources of dissipation in Josephson devices include quasiparticle losses in the superconducting leads and dielectric losses in the capacitive circuit elements [4, 5]. The mechanisms related to the Josephson junctions themselves are the dielectric loss of the tunnel barrier material and the loss induced by the quasiparticle tunneling through the tunnel barrier [6]. In this paper we aim to quantify the contribution of these two dissipation mechanisms in metamaterial circuits.

2. THEORY

The dissipative current component in superconducting tunnel junctions, the quasiparticle current, is effectively in parallel with the Josephson current. It is expressed as a function of the voltage across the junction as

$$I = \frac{1}{eR_n} \int_{-\infty}^{\infty} n_s(E' - eV) n_s(E') [f(E' - eV) - f(E')] dE', \quad (1)$$

where R_n is the normal-state resistance of the junction, e is the electron charge, and $n_s(E)$ and $f(E)$ are the superconducting density of states (scaled to normal metal density of states), and the Fermi function, respectively. The Fermi function is expressed as $f(E) = [\exp(E/k_B T) + 1]^{-1}$ with E referred to the Fermi level. For ideal superconductors the density of states is expressed as

$$n_s(E) = \frac{E}{\sqrt{E^2 - \Delta^2}}, \quad (2)$$

for $|E| > \Delta$ and zero otherwise. The superconducting gap is Δ , and the density of states diverges at the edge of the superconducting gap when $|E| = \Delta$. However, in real superconductors the density of states can be considered finite within the gap, as captured by the Dynes model [7]

$$n_s(E) = \left| \operatorname{Re} \left(\frac{E + i\gamma\Delta}{\sqrt{(E + i\gamma\Delta)^2 - \Delta^2}} \right) \right|, \quad (3)$$

where γ is a phenomenological coefficient introduced to quantify the smearing. As the Josephson metamaterial is typically operated with zero DC voltage, the relevant quantity is the zero-bias resistance

$$R_{q0} = \left(\left. \frac{\partial I}{\partial V} \right|_{V=0} \right)^{-1}. \quad (4)$$

The other mechanism related to the tunnel junction element itself is the dielectric loss of the junction barrier. From the point of view of microwave damping, it plays the same role as the quasiparticle dissipation. Therefore it is natural to present it as equivalent resistance, comparable to R_{q0} parallel to the Josephson junction as

$$R_{d0} = \frac{1}{\omega C_J \tan \delta}, \quad (5)$$

where ω is the frequency, C_J is the Josephson junction capacitance and $\tan \delta$ is the loss tangent of the junction barrier material. The total loss can be represented by the parallel connection of R_{d0} and R_{q0} as

$$R_0 = \frac{R_{d0} R_{q0}}{R_{d0} + R_{q0}} \quad (6)$$

In Fig. 1(a) we present an equivalent circuit model of a unit cell of Josephson metamaterial consisting of a SQUID loop embedded in a microwave transmission line. Fig. 1(b) presents a microscope photograph of a few unit cells of the metamaterial realised in coplanar waveguide geometry. Assuming a small loop inductance $L_s \ll L_J$, the absence of DC current allows us to linearize the supercurrent effects as described by the Josephson inductance $L_J = \Phi_0 / (2\pi I_c)$ in parallel with the junction capacitance C_J and the dissipation given by R_0 . The Josephson inductance (or I_c) can be tuned by a flux induced by current indicated as I_ϕ in Fig. 1.

The relevant properties for our analysis are contained in the propagation constant of the transmission line $\eta = \alpha + j\beta$, and the transmission line impedance Z_t . In a regular microwave transmission line represented by a circuit equivalent of Fig. 1(c) the quantities are expressed as $\beta = \omega \sqrt{L_t C_t}$, $\alpha = (1/2)(R/Z_t + Z_t/G_t)$, and $Z_t = \sqrt{L_t/C_t}$, where L_t , C_t , R_t and G_t are transmission line inductance, capacitance, series resistance describing conductor loss, and parallel conductance describing dielectric loss per unit length. This is valid in the low-loss limit $\alpha \ll \beta$. Here we have discretized the system to the physical length of the unit cell ℓ , assumed much shorter than the wavelength $\lambda = 2\pi/\beta$. To obtain η and Z_t for the metamaterial we set the impedance of the center conductor of the coplanar line $R_c + jX_c$ (impedance between points A and B within the dotted rectangle in Figs. 1(a) and 1(b) equal to the corresponding impedance of the transmission line equivalent

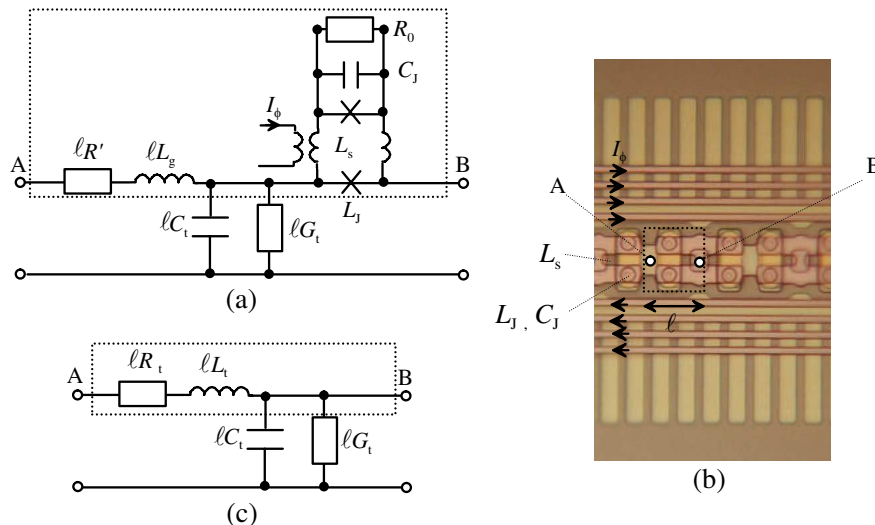


Figure 1: (a) The circuit schematic of a unit cell. (b) A microscopic photograph of a few unit cells of the metamaterial fabricated with Nb trilayer process. The symbols are defined in the main text. (c) The circuit equivalent of the reduced transmission line model.

$\ell R_t + j\omega\ell L_t$. The effective $L_t = X_c/\omega$ and $R_t = R_c$ are found through straightforward circuit analysis. Substituting them back to the definitions of η and Z_t yields

$$\beta = \omega \sqrt{\left(L_g + \frac{1}{\ell} \frac{\frac{1}{\omega^2 L_J} - C_J}{\frac{1}{R_0^2} + \left(\omega C_J - \frac{1}{\omega L_J} \right)^2} \right) C_t} \approx \omega \sqrt{\left(L_g + \frac{L_J}{\ell} \right) C_t}, \quad (7)$$

$$\alpha = \frac{1}{2\ell Z_t} \frac{R_0}{1 + R_0^2 \left(\omega C_J - \frac{1}{\omega L_J} \right)^2} \approx \frac{1}{2\ell Z_t} \frac{\omega^2 L_J^2}{R_0}, \quad (8)$$

and

$$Z_t = \sqrt{\frac{L_g + \frac{1}{\ell} \frac{\frac{1}{\omega^2 L_J} - C_J}{\frac{1}{R_0^2} + \left(\omega C_J - \frac{1}{\omega L_J} \right)^2}}{C_t}} \approx \sqrt{\frac{L_g + \frac{L_J}{\ell}}{C_t}}. \quad (9)$$

Here we have neglected the loss mechanisms not contained in R_0 (i.e., we assumed $R' \approx 0$, $G_t \approx 0$). It has been also assumed that the frequency of the propagating wave is much below the plasma frequency ω_p of the Josephson junction, i.e., $\omega \ll \omega_p = 1/\sqrt{L_J C_J}$, and the low-loss limit $R_0 \ll \omega L_J$.

The phase velocity of light in the metamaterial is expressed as $v_p = \omega/\beta$. If the metamaterial is to be used as a distributed resonator the intrinsic quality factor of the resonance is readily expressed through above transmission parameters. The quality factor of the non-terminated $\lambda/2$ or short-terminated $\lambda/4$ resonance is expressed as

$$Q_I = \frac{\beta}{2\alpha}. \quad (10)$$

3. NUMERICAL RESULTS AND DISCUSSION

Quasiparticle IV curves were computed from Eq. (1) with the aid of Eq. (3), see, e.g., [8] for a review. For Niobium a superconducting energy gap $\Delta = 1.4$ meV was assumed. One example is shown in Figs. 2(a) and 2(b). The former shows the IV curve and the latter its derivative. The feature of interest in our case is the zero-bias conductance dip with height $1/R_{q0}$. The zero-bias resistance R_{q0} is plotted at different temperatures as function of junction quality parameter in Fig. 2(c). It is evident that decreasing the temperature has a huge effect as the resistance increases by 4 to 5 orders of magnitude when the temperature decreases from 4.2 K to 1.0 K. The junction quality in the range of interest has an effect up to about one order of magnitude in the range of the simulation of $\gamma = 10^{-6} \dots 10^{-2}$. The zero-bias dip is characteristic to superconductor-insulator-superconductor junctions even in case of ideal or near-ideal junctions as experimentally demonstrated with our junction technology in [9]. This is in contrast to normal metal-insulator-superconductor junctions, where γ quantifies the zero-temperature limit of the leakage resistance in a straightforward fashion [10].

To relate the results to an experimental case we consider the parameters of the device described in [11], i.e., a metamaterial parametric reflection amplifier consisting of 583 junctions connected as open-ended $\lambda/2$ resonator. With the device parameters of $L_g \approx 0.3$ $\mu\text{H}/\text{m}$, $C_t \approx 0.2$ nF/m, $L_J = 16$ pH, and $\ell \approx 16$ μm we obtain $\beta \approx 300$ 1/m at the operating frequency of $\omega/2\pi \approx 3.1$ GHz, and the transmission line impedance of $Z_t \approx 81$ Ω from Eqs. (7) and (9), respectively. The constant of attenuation from Eq. (8) yields $\alpha R_0 \approx 38$ Ω/m . At the base temperature of 100 mK the maximum experimental intrinsic Q-value was 6000. From Eq. (10) it then follows $\alpha \approx 0.025$ 1/m. This maps to $R_0 \approx 1500$ k Ω assuming that the attenuation is limited by the tunnel junction specific mechanisms. At the temperature of 4.2 K the experimental Q-value was about 50. A similar analysis in this case leads to $R_0 \approx 13$ Ω .

The computational normal state resistance R_{q0} assumes values of $10^4 R_n \dots 10^5 R_n$ at $T = 1.0$ K. The normal state resistance of the SQUID junctions in parallel is about 40 Ω leading to the theoretical value of $R_{q0} \gtrsim 4 \times 10^5$ Ω with the data of Fig. 2(c). As at $T = 100$ mK one expects even higher R_{q0} it appears highly unlikely that quasiparticle loss limits the dissipation at sub-K temperatures since R_{q0} significantly exceeds R_0 estimated from the experiment. At $T = 4.2$ K it follows from the data of Fig. 2 that $R_{q0} \gtrsim 40$ Ω . This is still a factor of about 3 higher than the experimental value of R_0 .

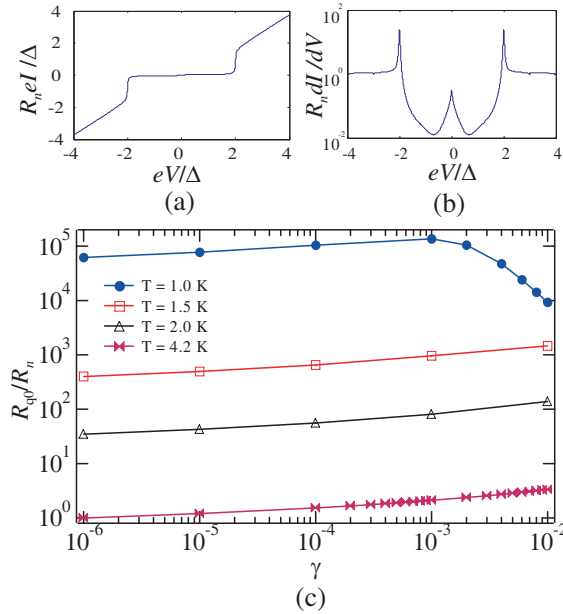


Figure 2: Numerical results to obtain quasiparticle dissipation of the superconducting tunnel junctions. (a) and (b): example current-voltage curve and its derivative, respectively, with example parameters ($T = 4.2$ K, $\gamma = 0.01$). (c): Zero-bias resistance R_{q0} obtained from the zero-voltage resistance dip as function of junction quality parameter γ at different temperatures.

If we instead consider R_{d0} through Eq. (5) we get the loss tangent of $\tan \delta = 2.7 \times 10^{-2}$ assuming $R_0 \approx R_{d0} \approx 1500 \Omega$ as it was estimated for $T = 0.1$ K, and using nominal junction capacitance of $C_J \approx 1.3$ pF. Similarly, at $T = 4.2$ K similar estimation yields $\tan \delta$ in the order of unity.

4. SUMMARY AND CONCLUSION

In summary we studied the dissipation in niobium based Josephson metamaterial devices. We excluded the possibility of quasiparticle tunneling to act as the limiting mechanism in the experiments involving niobium-based Josephson metamaterial at sub-K temperatures. The damping induced by the dielectric dissipation of the junction barrier material appears to be a plausible explanation at sub-K temperatures. However, at the liquid-He temperature the quasiparticle dissipation becomes comparable within a factor of about 3 to the dielectric dissipation within our model.

ACKNOWLEDGMENT

The work was supported by the European Commission Grant EU-FP7-NMP-246026 and by the Academy of Finland through projects 135135, 141559, 263457, and the Center of Excellence “Low Temperature Quantum Phenomena and Devices” project 250280.

REFERENCES

1. Castellanos-Beltran, M. A., K. D. Irwin, G. C. Hilton, L. R. Vale, and K. W. Lehnert, “Amplification and squeezing of quantum noise with a tunable Josephson metamaterial,” *Nat. Phys.*, Vol. 34, No. 10, 929–931, 2008.
2. Lähteenmäki, P., G. S. Paraoanu, J. Hassel, and P. Hakonen, “Dynamical Casimir effect in a Josephson metamaterial,” *Proc. Natl. Acad. Sci.*, Vol. 110, No. 11, 4234–4238, 2013.
3. Wilson, C. M., Johansson, G., Pourkabirian, A. M. Simoen, J. R. Johansson, T. Duty, F. Nori, and P. Delsing, “Observation of the dynamical Casimir effect in a superconducting circuit,” *Nature*, Vol. 479, No. 11, 376–379, 2011.
4. Martinis, J. M., K. B. Cooper, R. McDermott, M. Steffen, M. Ansmann, K. D. Osborn, K. Cicak, S. Oh, D. P. Pappas, R. W. Simmonds, and C. C. Yu, “Decoherence in Josephson Qubits from dielectric loss,” *Phys. Rev. Lett.*, Vol. 95, No. 21, 210503, 2005.
5. Gunnarsson, D., Pirkkalainen, J.-M., J. Li, G. S. Paraoanu, P. Hakonen, M. Sillanpää, and M. Prunnila, “Dielectric losses in multi-layer Josephson junction qubits,” *Supercond. Sci. Tech.*, in press, arXiv:1206.1138, 2012.

6. Catelani, G., S. E. Nigg, S. M. Girving, R. J. Schoelkopf, and L. I. Glazman, “Decoherence of superconducting qubits caused by quasiparticle tunneling,” *Phys. Rev. B*, Vol. 86, No. 18, 1841514, 2012.
7. Dynes, R. C., J. P. Garno, G. B. Hertel, and T. P. Orlando, “Tunneling study of superconductivity near the metal insulator transition,” *Phys. Rev. Lett.*, Vol. 53, 2437, 1984.
8. Kühn, T. and G. S. Paraoanu, “Electronic and thermal sequential transport in metallic and superconducting two-junction arrays,” *Trends in Nanophysics: Theory, Experiment, Technology*, edited by V. Barsan and A. Aldea, 99–131, Engineering Materials Series, Springer-Verlag, Berlin, 2010.
9. Castellano, M. G., L. Grönberg, P. Carelli, F. Chiarello, C. Cosmelli, R. Leoni, S. Poletto, G. Torrioli, J. Hassel, and P. Helistö, “Characterization of a fabrication process for the integration of superconducting qubits and rapid-single-flux-quantum circuits,” *Supercond. Sci. Tech.*, Vol. 19, 860–864, 2006.
10. Pekola, J. P., V. F. Maisi, S. Kafanov, N. Chekurov, A. Kemppinen, Y. A. Pashkin, O.-P. Saira, M. Möttönen, and J. S. Tsai, “Environment-assisted tunneling as an origin of the Dynes density of states,” *Phys. Rev. Lett.*, Vol. 105, 026803, 2010.
11. Lähteenmäki, P., “Near quantum limited microwave amplifiers based on superconductive tunnel junctions,” Masters Thesis, Aalto University, 2010.

RFID Tag on PCB in UHF Band

Won-Kyu Choi¹, Seung-Hwan Jeong¹, Chan-Won Park¹, and Hae-Won Son²

¹Electronics and Telecommunications Research Institute (ETRI), Republic of Korea

²Chonbuk National University, Republic of Korea

Abstract— This paper presents a tiny RFID tag which can be directly mounted on a PCB in electronic products as a solution to tracking PCBs within the electronics industry. The designed RFID tag consists of the tag chip, two capacitors and the 3-turn rectangular loop antenna. The RFID tag with the loop antenna utilizes a ground plane of a PCB as an ancillary radiator by an inductive coupling. Because the tag does not necessarily require direct connection to a PCB ground plane it can easily be mounted on a PCB using normal adhesive or sticky tape. The tag is $1.8 \times 10 \times 1$ mm in size and the fractional bandwidth for a return loss of more than 3 dB is about 4.2% at 920 MHz. The reading distance of the tag ranges from 0.3 to 0.7 m depending on a size of a PCB ground plane. Two capacitors of 3.6 and 3.9 pF are used in series for matching between the 3-turn loop antenna and tag chip.

1. INTRODUCTION

Radio frequency identification (RFID) is a rapidly developing technology that uses RF signals for the automatic identification of objects. This technology can be used not only for identification but also for tracking individual items in a supply chain. An antenna is one of the key factors in an RFID system [1]. The detection range and accuracy of an RFID system are directly dependent on the performance of a reader and tag antenna. Recently, various small RFID tag antenna have been widely studied for UHF RFID system in order to identify item-level small products for instance a medical vial [2] and metallic goods [3]. Also adoption of RFID technology has been profoundly considered for tracking printed circuit boards (PCBs) populated with electric components such as resistors, capacitors or active devices. A tiny RFID tag to be mounted on PCBs can be used to manage the tracking information and perceive bottlenecks within the manufacturing process. Some companies have lately developed the RFID tag module to track PCBs in electronic products around the manufacturing process [4, 5]. Important challenges associated with design of the RFID tag in electronics are to reduce the size of the tag, to realize the matching circuits, and to minimize the adverse effect on other electronic components in the PCB. In this paper, a very simple and tiny UHF RFID tag is proposed for applications to track PCBs in electronics. Since the designed tag recognizes a ground plane of a PCB as an additional radiator, the reading distance of the tag is dependent on the size of a PCB. The antenna of the tiny tag is comprised of the 3-turn rectangular loop in the dielectric body, FR-4.

2. DESIGN OF THE RFID TAG

The designed tiny RFID tag is made up of two-layered dielectric substrate, FR-4. The lower dielectric substrate is 0.2 mm in thickness and the upper one 0.8 mm. On the top side of the dielectric body for the RFID tag module is a part of the 3-turn rectangular loop patterned and the ground plane is formed on the bottom side of. Figure 1(a) shows the structure of the designed tiny RFID tag. The RFID tag chip and the capacitors for matching are assembled on the metallic antenna pattern which is etched on the top side of the dielectric body. The structure of the 3-turn loop antenna is completely realized by the metallic strips etched on the top side of the dielectric body, the strips inside the body, the ground plane formed on the bottom side of the body and the six via-holes. As shown in Figure 1(a), the dielectric body for the RFID tag consists of two dielectric layers. The ground plane is formed on the dielectric of layer 1 and the metallic strips for the rectangular loop antenna are etched on the dielectric of layer 2. Also four of the six via-holes just go through layer 2 and the others concurrently through layer 1 and 2. The designed RFID tag is $1.8(W) \times 10(L) \times 1(H)$ mm in size.

The tag chip employed in this paper is the commercial product, Higgs 3, of Alien Technologies and the chip data sheet represents that the chip is made up of the resistor of resistance $1.5 \text{ k}\Omega$ in parallel with the capacitor of capacitance 0.85 pF in the equivalent circuit diagram [6]. However we designed the tag antenna regarding the tag chip which equivalently consists of the resistor of $1.5 \text{ k}\Omega$ and the capacitor of 1.2 pF in parallel considering a parasitic capacitance caused by bonding the tag

chip directly to the antenna. So the measured input impedance of the tag chip is about $(14 - j143) \Omega$ at 920 MHz. The impedance of the tag antenna is designed to be conjugately matched with that of the tag chip to supply the maximum power to the tag chip. The required input impedance of the antenna for the tag chip can be easily achieved by the metallic 3-turn loop and two capacitors. In order to match between the 3-turn loop antenna and tag chip two capacitors of 3.6 and 3.9 pF are serially assembled on the loop antenna and the designed tag is mounted near the edge of a PCB in an electronic product as shown in Figure 1. The designed tag is inductively coupled with a PCB ground plane through magnetic fields generated by the 3-turn loop so a PCB ground plays an important role as an ancillary radiator of the tag. Figure 2 shows the variations of the return loss of the tag antenna depending on the PCB ground size and the magnitude distribution of the surface current induced on the ground plane at 920 MHz when the size of the PCB ground plane is 40 mm in width (W) and 80 mm in length (L). The fractional bandwidth for a return loss of more than 3 dB is approximately 4.2% at 920 MHz, ranging from 900 to 938.5 MHz. In Figures 2(a) and (b), we can examine that the variations of the return loss is just a little as a function of the PCB ground size and the surface current induced by the tag flows on the ground plane. This induced current causes the PCB ground to work as an ancillary radiator. Therefore the size of the PCB ground strongly influences the read range of the tag.

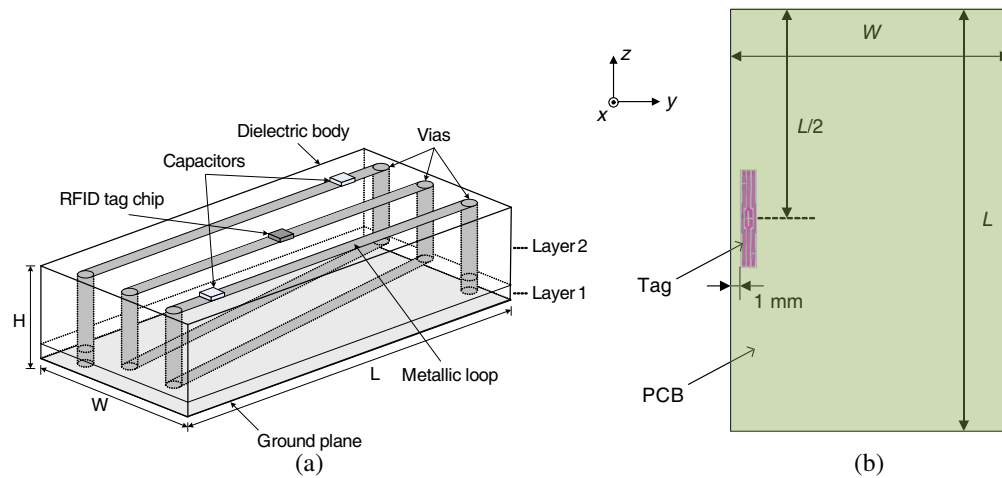


Figure 1: The designed tiny RFID tag. (a) Structure of the tag and (b) configuration for the tag to be assembled on a PCB.

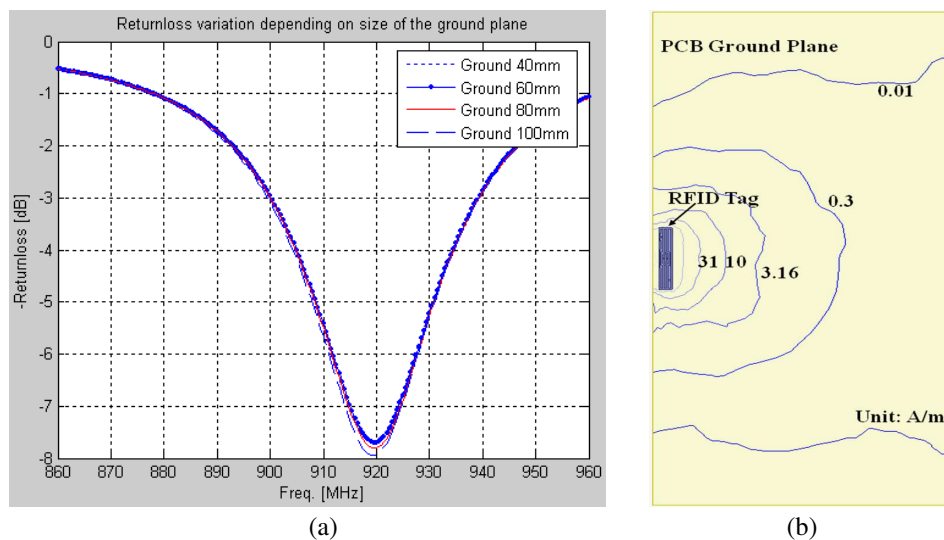


Figure 2: Return loss and surface current distribution. (a) Return loss depending on PCB ground plane. (b) Magnitude distribution of surface current.

Figure 3 shows the normalized radiation pattern of the electric field radiated by the tag in xy - and xz -plane, respectively. The radiation pattern of the tag inductively coupled with the PCB

ground plane is similar to that of dipole-type antenna which is parallel to the z -axis. The maximum radiation occurs at $\theta = 90^\circ$ and $\varphi = -60^\circ$ and then the gain is about -16.8 dBi. We used the commercial simulator HFSS, finite-element method (FEM), to compute the return loss, surface current distribution, and radiation pattern of the tag.

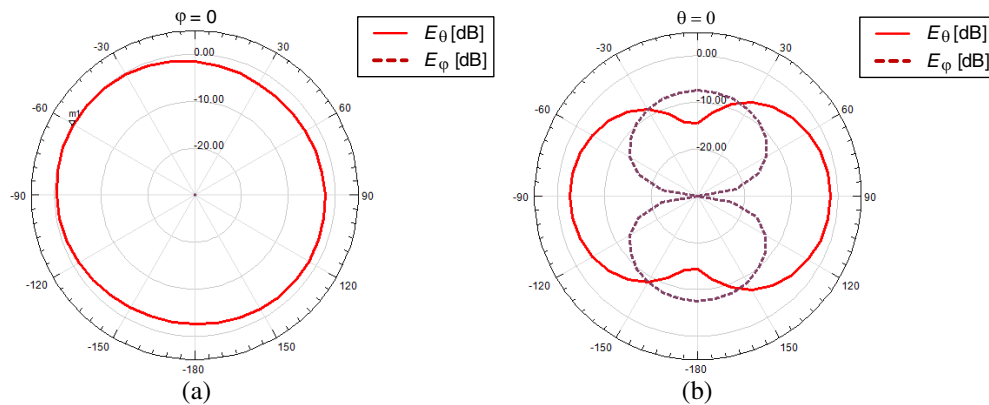


Figure 3: Radiation pattern (E_θ and E_φ). (a) xy -plane, (b) xz -plane.

3. EXPERIMENTAL RESULTS

The reading distance of the tag is measured in a long corridor. We realized the measurement environment which can minimize the effect of lots of reflection waves on the reading distance in a corridor. Figure 4 shows the measurement setup for reading distance and the fabricated tag which is assembled on the PCB of an electronic product. The PCB is 50×50 mm in size. In order to test the reading performance of the tag we used the commercial RFID reader, XCODE-IU9003, delivered by LSIS CO. in Korea [7]. The reading distance of the tag is roughly 0.45 m when the output power of the reader is 30 dBm and the reader antenna has 6 dBic of gain and linear polarization.

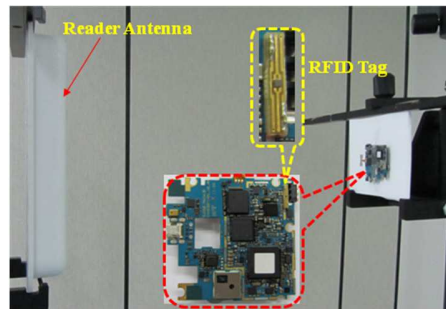


Figure 4: Measurement setup for reading distance.

4. CONCLUSIONS

The tiny RFID tag is proposed as a solution to identify PCBs in electronic products. The designed RFID tag consists of the tag chip, two capacitors and the 3-turn rectangular loop antenna. The RFID tag with the loop antenna utilizes a ground plane of a PCB as an ancillary radiator by an inductive coupling. So the reading distance is dependent on the size of the PCB ground plane. The reading distance of the tag ranges from 0.3 to 0.7 m depending on a size of a PCB ground plane. The fractional bandwidth for a return loss of more than 3 dB is approximately 4.2% at 920 MHz, ranging from 900 to 938.5 MHz. The designed tag can be directly assembled on the PCB in electronic products early in the manufacturing process. The tag can be used to not only prevent a replica of a product from being distributed but also trace a product in real time.

ACKNOWLEDGMENT

This work was supported by ETRI R&D Program [13ZC1100, The Development of an Intelligent Situation Cognition and an IoT Basic Technology] funded by the Government of Korea.

REFERENCES

1. Rao, K. V. S., P. V. Nikitin, and S. F. Lam, “Antenna design for UHF RFID tag: A review and a practical application,” *IEEE Trans. Antennas Propag.*, Vol. 53, No. 12, 3870–3876, 2005.
2. Lee, J. N., M. Y. Hwang, S. I. Lee, K. C. Lee, and J. K. Park, “Design of an ultra-compact UHF passive RFID tag antenna for a medical sample tube,” *ETRI J.*, Vol. 34, No. 6, 974–977, 2012.
3. Björninen, T. and L. Ukkonen, “Small slot antenna for metal mountable UHF RFID tags,” *Proceedings of IEEE International Symposium on Antennas and Propagation*, 1–2, 2012.
4. Murata Manufacturing Co., Ltd., Available at: <http://www.murata.com/products/rfid/index.html>.
5. NXP Semiconductors, Available at: http://www.nxp.com/documents/application_note/AN10-906.pdf.
6. Alien Technology, Available at: <http://www.alientechnology.com/wp-content/uploads/Alien-Technology-Higgs-3-ALC-360.pdf>.
7. LSIS Corporation, Available at: <http://www.lsis.biz/product/>.

Deep Defect Detection Using Eddy Current Testing with AMR Sensor

D. F. He and M. Shiwa

National Institute for Material Science, Sengen 1-2-1, Tsukuba 305-0047, Japan

Abstract— We developed high sensitive anisotropic magneto resistive sensor and used it for eddy current testing (ECT). Due to the high magnetic field sensitivity of AMR sensor at low frequency, the ECT system can be used to detect deep defect. The excitation coil was 50 turns with the diameter of about 30 mm. The amplitude of the excitation current was about 20 mA. The AMR sensor was put at the center of the excitation coil, and the sensing direction of the AMR sensor was parallel with the plane of the excitation coil. By this arrangement, the background field produced the excitation coil can be compensated. Using this ECT system, we successfully detect the deep crack defect in an aluminum plate with the depth of 15 mm, and the signal-to-noise ratio was best for the excitation frequency of 32 Hz.

1. INTRODUCTION

Eddy current testing (ECT) is an efficient surface and near surface inspection method. Inductive coil, hall sensor, flux gate, giant magneto resistive (GMR) sensor, anisotropic magneto resistive (AMR) sensor, and Superconducting quantum interference device (SQUID) have been used to construct ECT systems [1–6]. To detect deep defect in conductive materials, like aluminum, copper and steel, high sensitive magnetic field sensors at low frequency are needed. SQUID had the best magnetic field sensitivity at low frequency and was used to detect deep defect in conductive material [7], but the need of low temperature cooling made the SQUID-based ECT system complex and limited its practical applications for ECT. We developed high sensitive AMR sensor and used it for ECT [8–10]. Compared with other sensors, AMR sensor had the advantages of good sensitivity at low frequency, easy operation, big dynamic range, and flat frequency response in a big bandwidth. In this paper, we will describe the experiments of deep defect detection using ECT with AMR sensor.

2. AMR-BASED ECT SYSTEM FOR DEEP DEFECT DETECTION

Figure 1 shows the schematic block diagram of ECT system with AMR sensor. AC magnetic field was produced by the excitation coil as AC current flew in it; and eddy current was induced in the specimen. The AMR sensor was used to measure the magnetic field produced by the eddy current. AMR sensor of HMC1001 was used. The lock-in amplifier was used to measure the amplitude and phase signal of the signal. In our experiment, only amplitude signal was used.

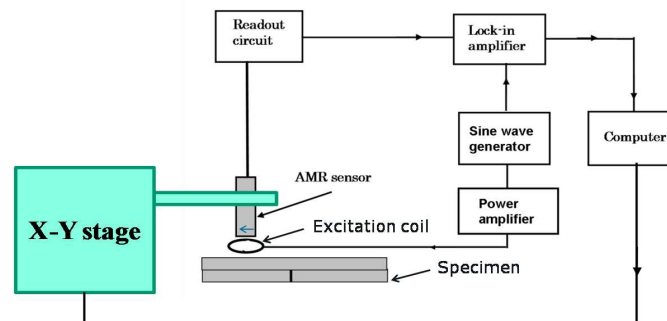


Figure 1: Schematic block diagram of ECT system with AMR sensor.

The excitation coil was a 50-turn circular coil with the diameter of about 30 mm. The AMR sensor was put in the center of the excitation coil, and the sensing direction was not perpendicular to, but parallel with the plane of the excitation. This configuration could compensate the background field produced by the excitation coil.

Figure 2 shows the magnetic field noise spectrum measured in our laboratory using the AMR sensor. The magnetic field resolution was about $40 \text{ pT}/\sqrt{\text{Hz}}$ at 32 Hz. The maximum magnetic

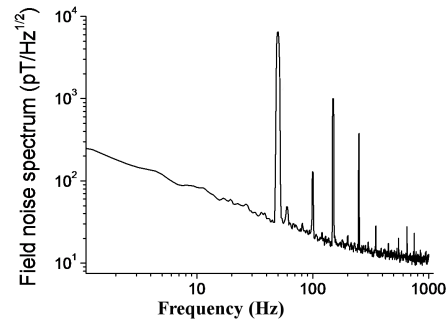


Figure 2: Magnetic field noise spectrum measured in our laboratory using AMR sensor.

field detected by AMR was about 2 Gauss, so the dynamic range of the AMR sensor was over 120 dB at 32 Hz.

3. ECT EXPERIMENTS

The sample was an aluminum plate with the thickness of 16 mm and artificial defect was made on it. The depth of the defect was 15 mm. The AMR sensor was fixed with an X-Y stage and the scanning was done by moving the AMR sensor. The scanning speed was about 5 mm/s. The penetration depth of eddy current can be expressed by $\delta = 1/\sqrt{\pi f \mu \sigma}$, where, f is the excitation frequency, μ is the permeability of the material and σ is the conductivity of the material. For aluminum, $\mu = \mu_0$; $\sigma = 3.5 \times 10^7$ S/m. If δ is 15 mm, the calculated excitation frequency is about 32 Hz. Figure 3 shows scanning results for the excitation frequency of 16 Hz, 32 Hz, 64 Hz, and 128 Hz. The amplitude of the excitation current was about 20 mA. The signal-to-noise ratio was best at the frequency of 32 Hz.

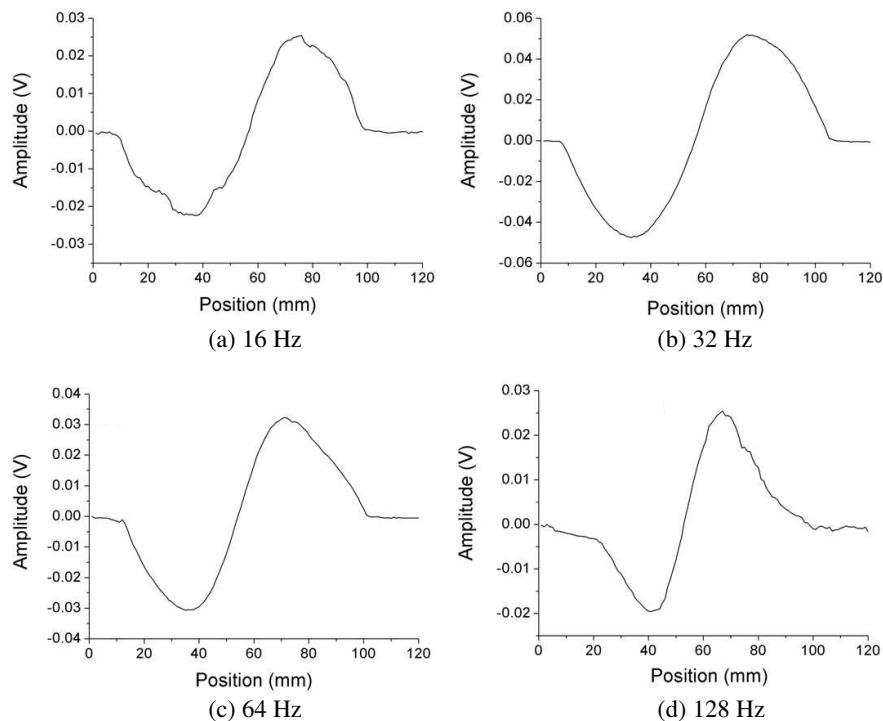


Figure 3: ECT signal of defect with the depth of 15 mm. (a) Excitation frequency was 16 Hz. (b) Excitation frequency was 32 Hz. (c) Excitation frequency was 64 Hz. (d) Excitation frequency was 128 Hz.

4. SUMMARY

Using ECT with AMR sensor, the defect with the depth of 15 mm was detected and the signal-to-noise ratio was best for the excitation frequency of 32 Hz.

REFERENCES

1. Ditchburn, R. J. and S. K. Burke, "Planar rectangular spiral coils in eddy-current non-destructive inspection," *NDT & E International*, Vol. 38, No. 8, 690–700, 2005.
2. Smith, R. A. and D. J. Harrison, "Hall sensor arrays for rapid large-area transient eddy current inspection," *Insight*, Vol. 46, No. 3, 142–146, 2004.
3. Carr, C. and J. C. Macfarlane, "The performance of fluxgate magnetometers for non-destructive evaluation," *Insight*, Vol. 41, No. 1, 20–24, 1999.
4. Vacher, F., F. Alves, and C. Gilles-Pascaud, "Eddy current nondestructive testing with giant magneto-impedance sensor," *NDT & E International*, Vol. 40, No. 6, 439–442, 2007.
5. Allweins, K., M. von Kreutzbruck and G. Gierelt, "Defect detection in aluminum laser welds using an anisotropic magnetoresistive sensor array," *Journal of Applied Physics*, Vol. 97, No. 10, 10Q102, 2005.
6. Jenks, W. G., S. S. H. Sadeghi, and J. P. Wikswo, "SQUIDS for nondestructive evaluation," *Journal of Applied Physics*, Vol. 30, No. 3, 293–323, 1997.
7. Hohmann, R., D. Lomparski, and H. J. Krause, "Aircraft wheel testing with remote eddy current technique using a HTS SQUID magnetometer," *IEEE Trans. Appl. Supercond.*, Vol. 11, No. 1, 1279–1282, 2001.
8. He, D. F., M. Tachiki, and H. Itozaki, "Highly sensitive anisotropic magnetoresistance magnetometer for Eddy-current nondestructive evaluation," *Rev. Sci. Instru.*, Vol. 80, 036102, 2009.
9. He, D. F., M. Shiwa, J. P. Jia, et al., "Multi-frequency ECT with AMR sensor," *NDT & E International*, Vol. 44, No. 5, 438–441, 2011.
10. He, D. F., Y. Z. Zhang, M. Shiwa, and S. Moriya, "Development of eddy current testing system for inspection of combustion," *Rev. Sci. Instru.*, Vol. 84, 014701, 2013.

Comparative Study on Graphene-based Artificial Magnetic Conductor (AMC)

X.-C. Wang, W.-Y. Li, W.-S. Zhao, J. Hu,
Y.-Y. Xu, and Q.-S. Huang

Centre for Opt. & EM Research, Zhejiang Provincial Key Lab for Sensing Technologies
State Key Lab of MOI, Zhejiang University, Hangzhou 310058, China

Abstract— In this work, the graphene is employed to form artificial magnetic conductor (AMC). It is demonstrated that both the resonance frequency and bandwidth of graphene-based AMC can be tuned by varying the external electric field. Additionally, the graphene-based AMC with different geometries are examined and compared.

1. INTRODUCTION

Over the past several years, artificial magnetic conductor (AMC) has attracted much attention due to its ability in suppressing the surface wave and realizing the low-profile antennas. The AMC ground plane has a perfect magnetic conductor characteristic over a certain frequency range, i.e., the so-called AMC operation bandwidth, which is defined in the frequency range corresponding to the reflection phases between $+90^\circ$ and -90° . The main drawback of AMC ground plane is its relatively narrow bandwidth compared to the ultra-wide bandwidth (UWB) antenna. Typically, there are two ways to broaden the bandwidth of the AMC: 1) Modify the AMC structure such as using via holes or adopting multilayer FSS over a grounded substrate; and 2) insert varactor diodes between the AMC units [1, 2]. However, these methods become tough at very high frequencies (for example THz) due to the complex manufacturing process.

Recently, tunable THz devices based on graphene are successfully developed by using the variable surface conductivity of graphene under different applied voltages. In this work, the tunability of graphene-based AMC is proposed and demonstrated. By varying the chemical potential, the equivalent capacitance of the graphene-based AMC can be dramatically tuned, and the resonance frequency shifts subsequently.

2. GRAPHENE-BASED AMC

Figure 1(a) shows the graphene-based AMC unit, and the surface impedance can be characterized as a capacitor. The equivalent circuit model is shown in Fig. 1(b).

2.1. Analytical Modeling

The reflection phase R for graphene-based AMC with square patches is obtained by [3]

$$R = \frac{Z_s - \eta_0}{Z_s + \eta_0} \quad (1)$$

where $\eta_0 (= 377 \Omega)$ is the intrinsic impedance of free-space, and Z_s is the total surface impedance, which can be calculated by (see Fig. 1(b))

$$Z_s = \left(\frac{1}{Z_d} + \frac{1}{Z_g} \right)^{-1} \quad (2)$$

where Z_d and Z_g are the impedances of grounded substrate and patch array, respectively [3, 4]

$$Z_d = j \frac{\eta_0}{\sqrt{\epsilon_r}} \tan(k_d t) = j\omega L_d(\omega) \quad (3a)$$

$$Z_g^{\omega\tau \gg 1} = j \left(\frac{D}{D-g} \frac{\omega\tau}{\sigma_0(\mu_c)} - \frac{1}{\omega C_{eff}} \right) = \frac{1}{j\omega C_g(\omega)} \quad (3b)$$

$$\sigma_0(\mu_c) = \frac{e^2 k_B T \tau}{\pi \hbar^2} \left[\frac{\mu_c}{k_B T} + 2 \ln \left(1 + e^{-\mu_c/k_B T} \right) \right] \quad (3c)$$

where D is the patch width, g is the patch gap, ϵ_r is the relative permittivity of the substrate, t is the substrate thickness, k_B is the Boltzmann's constant, h is the reduced Planck's constant, T is the temperature, τ and μ_c are the relaxation time and chemical potential of the graphene sheet, respectively. The comparison between the analytical and simulated results is plotted in Fig. 2, and good agreement can be observed. Here, $D = 8 \mu\text{m}$, $g = 1 \mu\text{m}$, $t = 8 \mu\text{m}$, and the substrate is selected as SiO_2 . By changing μ_c from 0.4 eV to 1 eV (see Fig. 3), the resonance frequency can be tuned from 1.88 THz to 2.67 THz, and a wide bandwidth of 50% can be achieved. It is also found that the graphene-based AMC will be more reflective as μ_c increases.

The performance of AMC ground plane will be affected in the case of oblique incidence, as shown in Fig. 4(a). With the increase of incident wave angle θ , the resonance frequency increases, while the bandwidth decreases, as shown in Fig. 4(b). Also, the impact of θ becomes more significant as it increases.

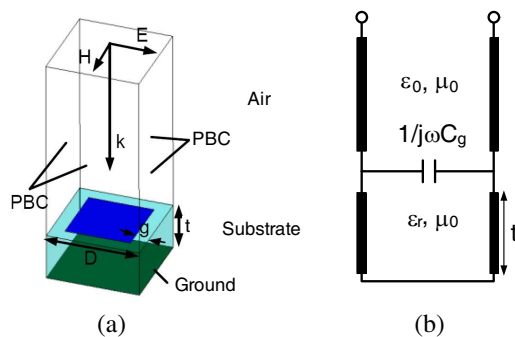


Figure 1: (a) Schematic of the AMC unit, and (b) its equivalent circuit model.

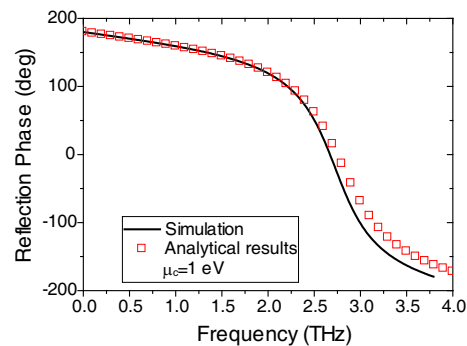


Figure 2: Comparison between analytical and simulated results for graphene-based AMC with square patches.

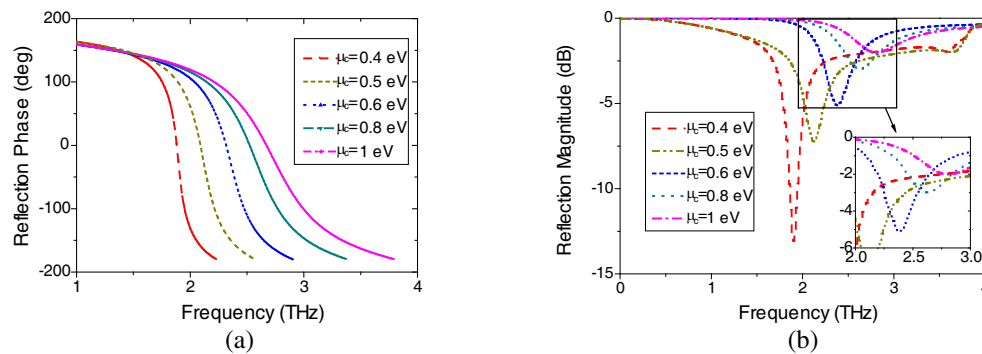


Figure 3: The reflection of the graphene-based AMC with different chemical potentials. (a) Phase; (b) Magnitude.

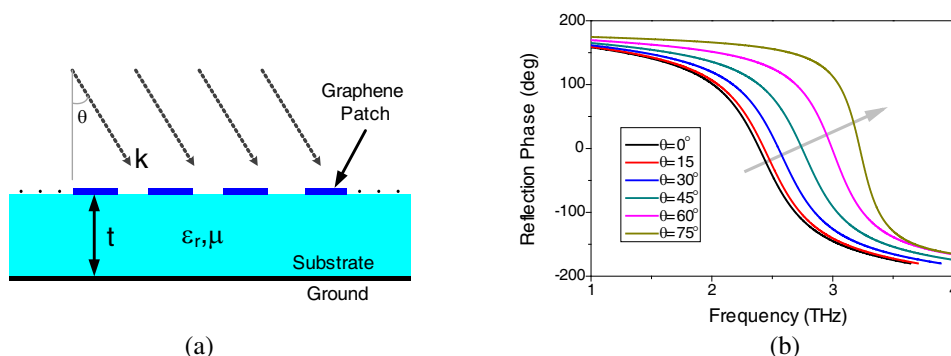


Figure 4: (a) Schematic view and (b) reflection phase of graphene-based AMC. The angle of incident wave is denoted by θ .

2.2. Parametric Study

The impacts of D , g , t , and ϵ_r are investigated for graphene-based AMC ground plane, as shown in Figs. 5(a)–(d). Both the bandwidth and resonance frequency of the graphene-based AMC can be decreased by increasing either D or ϵ_r , while g has an opposite effect. By employing thicker substrate, the resonance frequency decreases but the bandwidth increases greatly. Nevertheless, it is impossible to thicken the substrate unlimitedly since the grounded substrate will not exhibit an inductor when t becomes larger than $\lambda/4$.

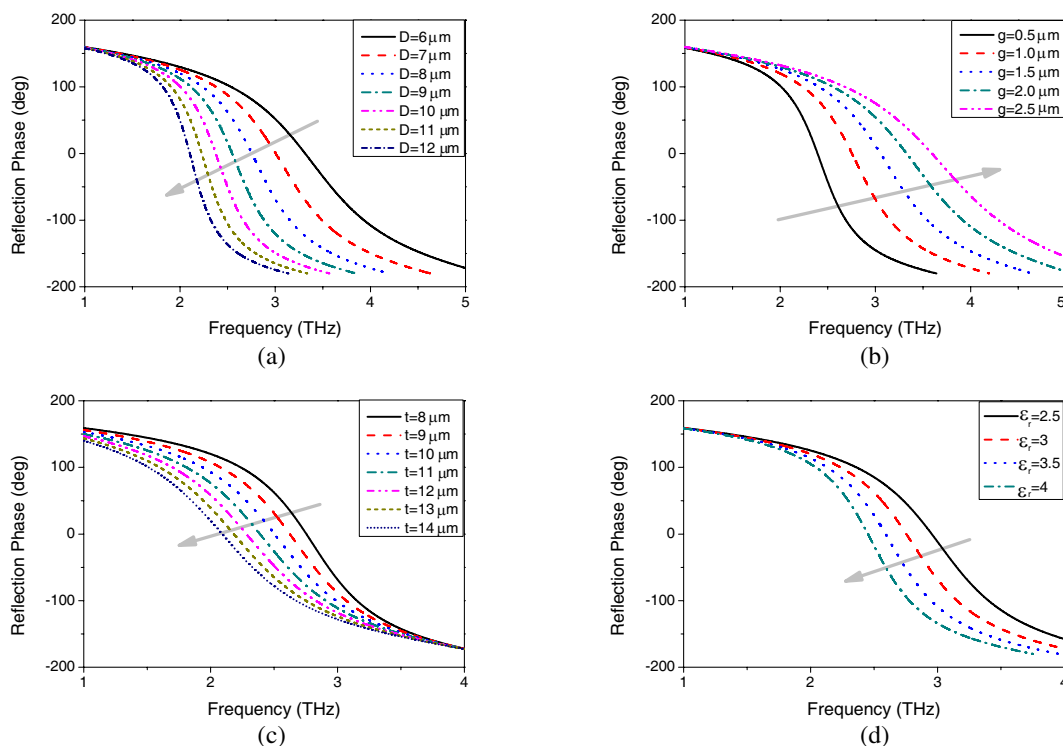


Figure 5: Parametric study of graphene-based AMC with (a) D , (b) g , (c) t , and (d) ϵ_r , respectively.

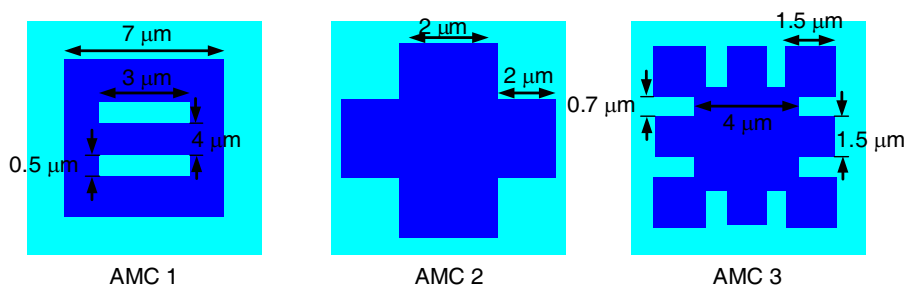


Figure 6: Schematic of different geometric shapes for graphene-based AMC.

3. DIFFERENT GEOMETRIC SHAPES

Various kinds of geometric shapes are studied in this section (see Fig. 6), and their reflection phases under different chemical potentials are plotted in Figs. 7(a)–(c). It is shown that the AMC 2 exhibits larger bandwidth while the AMC 1 shows a better tunability, as shown in Fig. 7(d). This phenomenon is because that the tunability of AMC ground plane is mainly determined by the equivalent capacitance contributed by the structure. From (3b), it is found that a larger C_{eff} means that the reactance contributed by graphene dominates thus a better tunability can be attained.

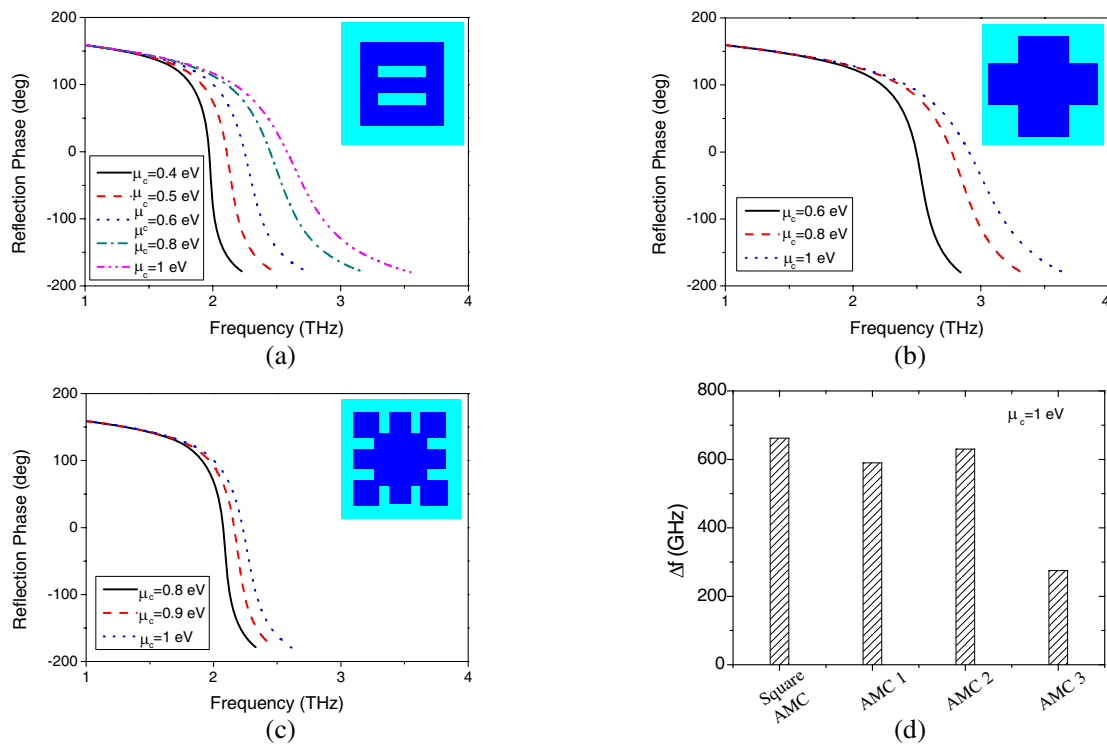


Figure 7: (a)–(c) The reflection phases of graphene-based AMC with different geometric shapes, and (d) comparison of frequency range Δf .

4. CONCLUSION

In this work, a tunable graphene-based AMC is proposed and demonstrated. It is shown that both the resonance frequency and bandwidth increases with the increase of applied voltage. Parametric study is also conducted to find the influence of the geometries on the AMC performance. Finally, the bandwidth and tunability of different AMC shapes are examined and compared.

ACKNOWLEDGMENT

This research was supported by the Program of Zhejiang Leading Team of Science and Technology Innovation.

REFERENCES

1. Veysi, M. and M. Shafee, “EBG frequency response tuning using an adjustable air-gap,” *Progress In Electromagnetics Research Letters*, Vol. 19, 31–39, 2010.
2. Costa, F., S. Talarico, A. Monorchio, and M. F. Valeri, “An active AMC ground plane for tunable low-profile antenna,” *Int. Symp. Antennas Propag.*, 1–4, San Diego, CA, 2008.
3. Tretyakov, S., *Analytical Modeling in Applied Electromagnetics*, Artech House, Inc., 2003.
4. Hanson, G. W., “Dyadic Green’s functions for an anisotropic non-local model of biased graphene,” *IEEE Trans. on Antennas Propag.*, Vol. 56, No. 3, 747–757, 2008.